ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
UNIVERSITY OF GOTHENBURG
225

Economic Determinants and Consequences of Political Institutions

Oana Borcan

UNIVERSITY OF GOTHENBURG

ISBN 978-91-88199-05-8 (printed)
ISBN 978-91-88199-06-5 (pdf)
ISSN 1651-4289 print
ISSN 1651-4297 online

Printed in Sweden,
Gothenburg University 2015

To Amrish, my family
and my mentors

Acknowledgements

First I want to thank my thesis advisors Andreea Mitrut and Ola Olsson. Words cannot do justice to how much help I have received from Andreea, who has been my guide, mentor and friend in this process and beyond. How can I ever repay you? Ola opened my eyes to new ways of thinking, sparked many new interests and showed great trust and optimism in my work. Whether through our joint work or their kind advice, the amount they taught me makes me feel like one exceptionally lucky PhD student.

I am deeply indebted to my co-authors Mikael Lindahl and Louis Putterman who have shared their passion and their knowledge and have taken our work to new heights. Mikael has taught me a lot about the value of being a great teacher. Louis has been very generous with his time and with facilitating my visit at Brown University.

A big "Thank you!" goes to all the teachers I had through the years here at the University of Gothenburg. Måns Söderbom deserves a special acknowledgement for his meticulous and inspiring teaching method, which I hope to take with me further, and also for his support at critical points. I was also fortunate to receive insightful feedback from Randi Hjalmarsson, who, in addition, was very kind to facilitate my visit at Queen Mary University.

This thesis had a great deal of valuable input of my licentiate and higher seminar discussants - Maria Perrotta Berlin and Andrea Tesei. I am thankful to them and also to the many discussants at international conferences, workshops and department seminars at Gothenburg, Brown and Queen Mary.

Importantly, the process of building further upon this thesis will be enhanced owing to the fruitful defense discussions. For this I want extend my most sincere thanks to my discussant Sasha Becker and the committee members Katarina Nordblom, Carl-Johan Dalgaard and Nicholas Charron.

The wonderful people that work hard to have the department run smoothly and who never spared efforts to assist me also have my utmost admiration and gratitude. To Econ-alla: I have a hard time believing any other economics crowd will ever match your friendliness and kind smiles.

I had a wonderful experience in Gothenburg being surrounded by lovely colleagues and friends. I will miss all the friends in my PhD generation, from my "PhD Sisters": Simona, Marcela, Anja and Diem, to Reda, Joakim, Emil, Remidius, Sied, Xiao-Bing and Hang. A special place in my heart is of course reserved for Simona, who has been my "partner in crime". We shared the good and the less glamorous PhD and life experiences and the friendship we forged is one of the best things in my life. Speaking of life outside PhD, I am happy to have made some fantastic friends here in Gothenburg, people that inspire me every time: Vali, Lavinia, Costel, Andrea, Efi, Mitesh, Verena, Simon, Vlad, Andrei. I thank you all and I will also forever be thankful to my lovely lifelong friends Flavia, Veronica and Corina, who have been by my side through thick and thin.

Going back in time, I nostalgically remember two teachers who have laid the foundation of my work: Valeria Suster and Marian Ursarescu have given me the words and the maths to get where I am today. I will never forget you!

It should be the easiest to talk about my family's input in this achievement. Yet words cannot express how much of this is their work. Besides their love, they instilled in me values essential to an academic career: My mother has been a wonderful, kind and dedicated teacher in just about everything. My father has, from a very young age, inspired me to read, ask, travel, search and discover. I've also learnt a thing or two about "constructive criticism" from them...

And the award for most patience, understanding, tireless support and unyielding belief that all will be well goes to my life companion - Amrish. I cannot imagine what the past five years would have been without your love and wisdom. This virtually became your second PhD, so enjoy the fruits of your work! *Iti multumesc pentru tot si te iubesc mult!*

To all of you I promise to do my best to give back and pass along to others all the good you gave me!

Contents

Introduction

Political institutions are inextricably linked with economic performance, a relationship that goes both ways. Institutions drive growth by, for instance, securing property rights and enforcing contracts. Institutions themselves are determined by the dynamics of economic growth since, for instance, contract enforcement depends on wealth.

The prominence of institutions in understanding economic performance dates to the seminal work of North (1990). His work spurred a new empirical tradition that broadly links country-level aggregates of growth, like income per capita, with measures of institutional quality, like the rule of law (e.g. Knack and Keefer 1995; Rodrik et al. 2004). While the data revealed that institutions may play an even bigger role in a nation's prosperity than initial conditions, like biogeography, it remained silent on how to achieve good institutions. In response another research tradition developed concurrently, focusing on specific cases of institutional failure and corrective policy interventions. The realization that bad governance and corruption were severely undermining poverty alleviation efforts was a major motivation for this approach.

Our enhanced understanding of institutions has no doubt contributed to the large reductions in poverty (from 50% to 21% of citizens in developing countries since 1981). However, challenges remain. Inequality is rising fast (OECD 2011); in 2000 the income levels of the richest and poorest countries differed by a factor of 200, in 2014 that factor was nearly 250.[1] In light of these dynamics, it is of critical importance to examine and reexamine where and why institutions fail. Are today's development challenges inherited from history or artefacts of the modern world? What are the implications of dysfunctional institutions for key aspects of socio-economic development, like education and democracy?

This thesis answers these questions with four empirical essays on the economic determinants and consequences of political institutions. It opens with a broad perspective on the link between the state throughout history and the past and current economic performance of nations. Remaining chapters focus on a particular form of institutional failure "endemic corruption" with illustrations of its harmful impact on key development areas: human capital formation and democracy. More specifically, chapters two and three identify causal links between public sector wages, monitoring

[1] In 2000, the GDP per capita of the richest country in the world, Qatar, was 112,283 USD and that of the poorest, The Democratic Republic of Congo, was 518 USD. In 2014 Qatar's per capita GDP was 139,456 USD and the Central African Republic's was 566 USD. [World Bank data: GDP in constant 2011 international dollars using PPP rates].

and incentives and corruption in education. The final chapter examines the clientelistic structure of electoral politics showing that local politicians influence national elections by vote buying and electoral fraud.

The historical role of institutions

The experience with state institutions has been put forth as one of the important correlates of the current wealth distribution in the world. In order to capture the modern day nations' accumulated experience of large-scale political organization, Bockstette, Chanda and Putterman (2002) have created the *State antiquity index*. This index reflected all forms and changes of government above tribal level, between 1-1950 CE, for each of 159 countries defined by modern-day borders. The authors showed that this indicator of state experience was positively associated with 1995 income and with the 1960-1995 GDP growth rate. Other studies confirmed this empirical pattern, besides demonstrating that this "persistence of fortune" is not a localized phenomenon, but is reinforced by migration between countries (Putterman and Weil, 2010).

The examples of former colonies that have gained independence late in the 20th century and are therefore much younger and inexperienced states than European states like the United Kingdom and France are in line with these findings. However, there are states that do not fit this pattern: countries with old institutions that consolidated as far back as antiquity (e.g. Iraq, Syria, India) perform much more poorly today than the much younger European states.

In **Chapter 1**, "State History and Economic Development: Evidence from Six Millennia" (with Ola Olsson and Louis Putterman), we investigate this issue and we attempt to make two distinct contributions to the literature. First, we provide a complete state history index from its first origin around 3500 BCE up until the present day. We extend the index from 1 CE backwards in time to the first origins of states around 3500 BCE and also code the 1950-2000 CE period. We follow the methodology in the original effort by Bockstette el al. (2002). This combines three dimensions of state development: 1) The existence of a state above tribal level; 2) Whether rule is internally or externally based; 3) The territorial coverage of the state in relation to current national borders. The three indicators were coded for each of the 159 countries in our sample and for each 50- year period from the origin of the first states around 3500 BCE, yielding a panel data set with 17,490 country-period observations.

Second, we investigate how our extended state history index is related to indicators of long-run economic development. Our theoretical foundation is a modified

version of the Malthusian growth model where we include the rise of a state that taxes its population and provides public goods that enhance productivity. The key hypotheses from the model are that the levels of productivity and population density have a concave and possibly an inverted u-shaped relationship with the extent of state history. This is particularly true in the pre-industrial era, but the non-linear effects of state experience are predicted to persist to this day. In the empirical section, we then confirm that the relationship between our state history index and levels of economic development has the shape of an inverted u, implying that the very young and very old states have the least developed economies whereas the richest countries have intermediate state history scores.

These results are only suggestive of the positive role of fiscal capacity in economic growth and the hindering role of early centralized power giving unlimited prerogatives to rent-seeking elites. Despite this limitation, the additional data and more flexible econometric specifications enhance the ability of state institutions to predict past and current economic performance.

Economics determinants and consequences of corruption

The relevance of historical approaches is naturally limited for today's policy challenges. One of the most difficult and pervasive current institutional failures that require urgent policy action is corruption. The most widely used definition of corruption is "the abuse of private office for personal gain". This includes, but is not limited to bribery; trading favours and gaining non-meritorious benefits through personal connections are common manifestations of corruption. According to Transparency International's Corruption Perception Index, more than two thirds of the world struggle with endemic corruption in most areas, from education and health, to business and elections. With few exceptions, corruption is bad for economic development, because it creates inefficiencies in the allocation of resources and talent, and, to the extent that it prevents the poor from accessing public services (OECD, 2014) and is developing countries, it deepens inequality. While it is most prevalent in developing countries, corruption also poses threats in richer countries, for instance, through the capture of the democratic process by vote buying or lobbying. Yet, despite intensified policy importance and academic study, corruption remains poorly understood. The biggest progress has been made in the measurement of corruption and its costs, ranging from perception–based surveys, to more objective metrics like audit-revealed missing funds. But all the political and academic efforts to understand and to reduce corruption have so far yielded more questions than answers (Olken and Pande, 2009).

Among these, how corruption responds to change in the bureaucrats' wage, and what are the distributional consequences of fighting corruption are some underexplored topics of particular interest to policy makers.

Corruption in Education

Chapter 2, "The Impact of an Unexpected Wage Cut on Corruption: Evidence from a "Xeroxed" Exam"(with Mikael Lindahl and Andreea Mitrut) investigates the effects of wages on corruption in the public sector, exploring a quasi-natural experiment generated by an unexpected 25% wage cut incurred by the public sector employees in Romania in 2010. The cut was announced in May, just one month ahead of the national high-school exit exam (the Baccalaureate) and it affected all public sector employees, including teachers. The ensuing question is whether the wage drop generated an increase in corruption for exam grades. The prevalence of corruption at the Baccalaureate exams was notorious and was attributed to the high-stakes character of the exam (it accounts for up to 100% of the university/college admission score) and the poor remuneration of teachers in general.

We do not observe the corrupt transactions at this exam, but we know that scores are corruption-inflated. Hence, our strategy is to use a Difference-in-Differences strategy to compare the change in the Baccalaureate exam outcomes (mainly the school level average grades and passing rates of the standardized Romanian language exam) from 2007 to 2010 between public and private schools. We choose private schools as a comparison group, as the latter category was not affected by the policy.

Our results show a positive and significant change in the exam outcomes between public and private schools, which we attribute to an increase in incentives to engage in corrupt activities in 2010 relative to previous years. In particular, we find a wage cut-driven effect equivalent to a 0.26 S.D. increase in exam scores and an increase in school-level Romanian exam pass rates by 3.3 percentage points. The estimated effects are equivalent to a nearly 4% increase in both exam outcomes. We employ different falsification tests and sensitivity analysis to lend further credibility to our results. The findings are also in line with the post–2010 exam unprecedentedly high number of allegations and trials for fraud and bribery by school principals and teachers connected with the Baccalaureate.

To our knowledge, this is the first paper that identifies a causal relationship between a wage cut in the public sector and corruption activities.

In **Chapter 3**, "Fighting Corruption in Education: What Works and Who Benefits", (with Mikael Lindahl and Andreea Mitrut), we follow-up on the exam corruption, by exploiting another policy change that came in 2011. Following these

trials and the accompanying media scandals, the Romanian Ministry of Education launched a large corruption-fighting campaign. The campaign consisted of two distinct components: 1) increasing the threat of punishment for teachers and students caught taking/giving bribes and 2) closed-circuit TV (CCTV) monitoring of the exam centers to prevent mass-cheating in the exam rooms. The CCTV monitoring introduction was gradual, with 25 counties installing cameras in 2011 and the remainder 17 counties in 2012. Hence, we use the quasi-experimental variation of the camera introduction and the characteristics of the Baccalaureate exam to answer two questions: 1) Was the campaign an effective means to reduce corruption? 2) Were students from different backgrounds differentially affected by the campaign?

Firstly, we find that the monitoring and punishment worked (as the pass rates nearly halved between 2010 and 2011). Comparing score changes between counties that installed camera early and those that installed it later, we find that the additional effect of the cameras was an 8% drop in exam passing rates. Secondly, we also compare the drop in exam performance between students from poor families with those from non-poor backgrounds. Our findings contradict our original expectation that fighting corruption should close the score gap between poor and non-poor students. The results indicate that the anti-corruption measures made the already underperforming poor students relatively worse off than non-poor students. The campaign induced an increase in achievement gaps also between low and high- ability student, as well as between male and female student. Thirdly, using data from an elite university, we show that the monitoring significantly reduced the chances of admission for poor students, hence confirming most of the results found for the Baccalaureate.

This paper contributes additional evidence to the literature on anti-corruption policies (e.g. Ferraz and Finan, 2011; Di Tella and Schargrodsky, 2003; Duflo et al., 2012). Our perhaps most important contribution is the estimated impact of fighting corruption on equality of educational opportunity, a topic much less explored.

Corruption in Elections

The last part of the thesis shifts focus on elite's capture of the democratic freedoms.

Chapter IV, "The Benefits of Local Party Alignment in National Elections", provides robust evidence that local officials deliver votes for their parties in national elections, quite plausibly through corrupt means.

In this paper I compare electoral outcomes between localities aligned with the party in government and those unaligned. Much of the elections literature has found larger intergovernmental transfers for aligned localities, but no significant benefits

for national politicians associated with these favours (e.g. Brollo and Nannicini, 2012).

To overcome the issue of endogenous alignment, I use a Regression Discontinuity Design with closely-contested Romanian local elections in June 2012. In terms of outcomes, I examine the turnout and vote shares at the July 2012 referendum launched by the governing coalition to dismiss the president, and the November 2012 legislative elections.

I find up to 5.4 percentage points increased turnout in government-aligned localities at the referendum. Turnout was crucial in the first polls, as a minimum participation of 50% of all voters was required for validating the referendum. By contrast, I find no electoral alignment advantage in turnout or vote shares in subsequent parliamentary elections.

The referendum alignment effect is driven by rural areas, with less educated and more manipulable voters. This along with the contrasting results at legislative elections, and extra heterogeneity tests suggest that local politicians mobilize voters successfully when: i) the voter commitment problem is overcome (unlike the vote, turnout is observable); ii) vote buying is common; iii) there is weak local competition and monitoring of incumbents. I also show suggestive evidence that after the referendum, government transfers increase in aligned localities and higher referendum turnout also drives higher legislative elections turnout and vote shares for the government coalition.

References

[1] Bockstette, V., Chanda, A. and Putterman, L. (2002). "States and Markets: The Advantage of an Early Start," Journal of Economic Growth, Springer, 7(4): 347 – 69, December.

[2] Brollo, F. and Nannicini T. (2012). "Tying your enemy's hands in close races: The politics of federal transfers in brazil." American Political Science Review 106: 742–761.

[3] Di Tella, R. and Schargrodsky, E. (2003). "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires," Journal of Law and Economics, University of Chicago Press, 46(1): 269-92.

[4] Duflo, E., Hanna, R. and Ryan, S. (2012). "Incentives Work: Getting Teachers to Come to School," American Economic Review, 102(4): 1241–78.

[5] Ferraz, C. and Finan, F. (2011). "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments," American Economic Review, American Economic Association, 101(4): 1274–1311, June.

[6] Knack, S. and Keefer, P. (1995). "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." Economics & Politics, 7: 207-227.

[7] North C. D. (1990). "Institutions, Institutional Change and Economic Performance," New York, Cambridge University Press.

[8] Rodrik, D., Subramanian, A. and Trebbi, F. (2004). "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development." Journal of Economic Growth, 9(2):131–165

[9] OECD (2011). "An Overview of Growing Income Inequalities in OECD Countries: Main Findings." http://www.oecd.org/els/soc/49499779.pdf

[10] OECD (2014). "The rationale for fighting corruption." http://www.oecd.org/cleangovbiz/49693613.pdf.

[11] Olken, B. and Pande, R. (2012). "Corruption in Developing Countries," Annual Review of Economics, 4: 479–505.

[12] Putterman, L. and Weil, D. N. (2010). "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," The Quarterly Journal of Economics, MIT Press, 125(4): 1627 – 1682, November.

Chapter I

State History and Economic Development: Evidence from Six Millennia*

Oana Borcan
University of Gothenburg

Ola Olsson†
University of Gothenburg

Louis Putterman
Brown University

Abstract

The presence of a state is one of the most reliable historical predictors of social and economic development. In this article, we complete the coding of an extant indicator of state presence from 3500 BCE forward for all countries in the world. We outline a theoretical framework, set in a Malthusian growth model, where accumulated state experience increases fiscal capacity but might have a negative effect on productivity when centralized power becomes excessive. The predictions of the model are tested in an empirical analysis where we introduce our extended state history variable. Our key finding is that, both as early as 1500 and today, the level of economic development is a concave not a linear function of accumulated state history.

JEL Codes: O11, O43, O50, N00

Keywords: State history, comparative development

*We are grateful for useful comments from Carl-Johan Dalgaard, Jakob Gerner Hariri, and from seminar participants at University of Copenhagen and Brown University. We also thank Taewan Roh and Nicholas Carter for valuable research assistance.

†Corresponding author: Ola Olsson, Dept Economics, University of Gothenburg, Box 630, 40530 Gothenburg, Sweden. ola.olsson@economics.gu.se.

1 Introduction

History has shown that economic development often thrives in states where governments guarantee the rule of law and provide public goods for their citizens. In order to reach a deeper understanding of why some countries have good government and others do not, social scientists have become increasingly interested in studying the long-run patterns of institutional development within states. The roots of countries' contemporary failures or successes have often been traced back to "critical junctures" far back in history.[1]

In this paper, we analyze how state development has interacted with economic development. More specifically, we attempt to make two distinct contributions to the literature. First, we provide a complete state history index from its first origin around 3500 BCE up until the present day. Initially developed by Bockstette, Chanda and Putterman (2002) for 159 countries, the index covered the period 1-1950 CE. We extend the index from 1 CE backwards in time to the first origins of states around 3500 BCE and also code the 1950-2000 CE period, which was previously missing from the time series.

Second, we investigate how our extended state history index is related to indicators of long-run economic development. Our theoretical foundation is a modified version of the Malthusian growth model where we include the rise of a state that taxes its population and provides public goods that enhance productivity. The key hypotheses from the model are that the levels of productivity and population density have a concave and possibly an inverted u-shaped relationship with the extent of state history. This is particularly true in the pre-industrial era, but the non-linear effects of state experience are predicted to persist to this day. In the empirical section, we then confirm that the relationship between our state history index and levels of economic development has the shape of an inverted u, implying that the very young and very old states have the least developed economies whereas the richest countries have intermediate state history scores.

For the first of these objectives - the creation of a state history index for the BCE-period, we follow the methodology in the original effort by Bockstette el al. (2002). This combines three dimensions of state development: 1) The existence of a state above tribal level; 2) Whether rule is internally or externally based; 3) The territorial coverage of the state in relation to current national borders. The three indicators were coded for each of the 159 countries in our sample and for each 50-year period from the origin of the first states around 3500 BCE, yielding a panel

[1] See for instance North (1990), Acemoglu at al (2005 and 2012), and Besley and Persson (2009).

data set with 17,490 country-period observations. The details of the sources for and construction of the index are described further below.

Our second objective hinged crucially on extending the state history data initially compiled by Bockstette et al. (2002). Their study was the first to show a significant correlation between state history and recent growth rate and between state history and income level. The numerous studies that followed strengthened the evidence that current development is positively related to state experience. Although subsequent versions of the index used in these papers expanded the set of countries, none coded the history of states BCE.

With these developments in mind and with the new data on the extended state history index, we revisit the relationship between the degree of exposure to state institutions and current output. We show that the relationship between state history and current income per capita is concave rather than linear, and that this is due to the inclusion of state experience before the Common Era. Thus, in addition to young, inexperienced states, very old states also incur economic disadvantages relative to states with around 2000 years of state experience.

Our inquiry is supported by the empirical observation that old states like Iraq, Turkey and China are poorer today than younger states like Britain, Denmark and Japan, a fact that remained unexplained in previous work. The early experience of the former was uncoded in the previous data, which effectively forced very old states to take similar values with intermediate states, such as England (the U.K.).

Building on previous literature, we contribute additional knowledge about the influence of early political and societal development on modern economic development. We show that along with young states, a very long state experience also comes with economic disadvantages relative to countries with intermediate state experience. We show that this more complex relationship primarily underlies the economic development indicators (population density and urbanization) and technology adoption in 1500 CE, but it also feeds into current economic performance. Moreover, the relationship for current outcomes is robust to adjusting the index for the ancestral lines of post-1500 migrant populations.[2]

The work clearly involves several methodological challenges. For instance, how should a state be defined? In this regard, we follow the tradition of Service (1962), Carneiro (1981), Johnson and Earle (2000) and others, distinguishing between bands, chiefdoms, and full-fledged states. Unlike the other forms of governments, states are further characterized by a centralized government with the ability to collect taxes,

[2]Olsson and Paik (2013) hint at this relationship, showing a "reversal of fortune" of countries that made an early Neolithic transition. This idea of reversal was also discussed by Acemoglu et al. (2001) and Hariri (2012).

enforce laws, and mobilize forces for war. Using this definition, most sources seem to be in rough agreement about the time when states arise in different countries. Accompanying this paper is an extensive online data appendix where we motivate the coding for each country-period observation.

Another issue concerns the unit of analysis, which is the territory delimited by modern-day country borders, for 159 contemporary countries in the sample. It is a well-known fact that the borders of current countries sometimes have very little resemblance with the geopolitical logic in ancient times.[3] However, to the extent that researchers are interested in tracking the histories of countries in order to understand contemporary levels of development, the modern configuration of countries is still a natural point of departure.[4]

The paper is organized as follows: In section 2, we provide an overview of the literature on the definition of a state and the relationship of state history to economic development. In section 3, we present our theoretical framework. In section 4 we present the new data and the principles guiding its construction. In section 5, we carry out an econometric analysis of the relationship between economic development and state history. Section 6 concludes.

2 Literature review

2.1 State history and economic development

It is a well established empirical fact that history has shaped the contemporary economic development of nations in numerous ways. Whether initial biogeographic endowment and transition to agriculture (e.g. Hibbs and Olsson, 2004; Olsson and Hibbs, 2005, Galor and Moav, 2007) or past technology adoption (Comin et al. 2006, 2009), early and productive starts have been typically shown to translate into better income and institutions in present times.

The experience with state institutions has been put forth as one of the important correlates of the current wealth distribution in the world. Specifically, from its original development, the State antiquity index of Bockstette et al. (2002) has been shown to be positively associated with 1995 income and with the 1960-1995 GDP

[3] Although this is a valid critique of the approach used here, there are also instances of countries where states have evolved in close proximity to current borders, at least for some periods of time (e.g. Norway, Sweden, and Japan).

[4] A potential alternative to using country borders could have been to divide the world into equal-sized grid cells and then study the history of states in each such cell. State history has been coded at the grid-cell level for sub-Saharan Africa after 1000 CE by Depetris-Chauvin (2014). For the present study, this would entail a very different type of analysis with its own methodological challenges. We leave this for future work.

growth rate. Bockstette et al.'s aim was to use presence and duration of experience with macro polities as one of several potential indicators of societal complexity and level of technological advancement. The authors were interested in investigating the effect of early social and technological development on post-WW2 economic growth rates, and they assumed that the impact of very early experience would decay over time, so they did not attempt to code information on state presence before 1 CE or after 1950. They coded all countries with substantial populations for which relevant economic growth and other indicators were available, resulting in a sample of 104 countries, of which their analysis focused especially on 70 non-OECD member countries.

Roughly the same data set was also used by Chanda and Putterman (2005), and Chanda and Putterman (2007). Bockstette et al.'s data were subsequently expanded to include more ex-Communist countries (Iliev and Putterman, 2007), more African countries (Cinyabuguma and Putterman, 2011), and a few other countries for which complementary income or other required data had initially been viewed as unreliable. Based on this extended dataset, Putterman and Weil (2010) demonstrated that the ability of state history to predict current levels of development is greatly strengthened by replacing the state history that transpired on a given country's territory by the weighted average state history of the places in which current residents' ancestors lived in the past. This adjustment was motivated by the large movements of populations especially from "Old World" continents to the Americas, Australia and New Zealand after 1500. Chanda, Cook and Putterman (2014) apply the same procedure to demonstrate "persistence of fortune" of ancestral lines in former colonies that display a "reversal of fortune" (Acemoglu, Johnson and Robinson, 2002) in the absence of such ancestry and migration accounting.[5]

In short, previous work has largely agreed on a positive association between long-run state history and current development. However, as scholars have acknowledged, the present shares complex links with the past. For instance, pre-1500 economic advantages seem to have become relative disadvantages among colonized countries during the colonial era (Acemoglu, Johnson and Robinson, 2001, 2002). As of late, this idea of reversal has been revisited in two studies that are particularly relevant to our paper: Hariri (2012) presents compelling evidence that early (precolonial) experience of state institutions in countries outside Europe prevented them from

[5]The state history data have also been employed in a number of other studies, receiving focal attention in Ang (2013a, 2013b), playing important roles in Ahlerup and Olsson (2012), Hariri (2012), Ertan, Putterman and Fiszbein (2012), and Daniele (2013), and being included as a control in a number of other studies. None of the above studies attempts to extend the information on states to include the BCE years or fill in the last half of the 20th Century.

transplanting democratic institutions brought by European colonizers, leaving instead an "autocratic legacy" in these countries. Olsson and Paik (2013) reveal a negative association between the time from Neolithic transition and current income levels in the Western agricultural core - Europe, North Africa and Southwestern Asia.

Furthermore, the long-run persistence literature has begun to reveal nonlinearities in how events in the very distant past affect economic development. For instance, the migration out of Africa is argued to have generated a wide array of genetic diversity levels in human populations around the world. In turn, predicted genetic diversity displays an inverted-u shape relationship with indicators of economic development, including per capita income in 2000 (Ashraf and Galor, 2013).

Thus, in light of these recent developments, allowing for a more flexible relationship between state history and contemporaneous levels of development is a natural extension to the literature. In the theoretical section below, we present a formal model of how states affect and are affected by economic development during the agricultural era.

2.2 Defining the "state"

How do we know when a state has emerged? The first challenge stems from the question of how to define the state, hardly a novel dilemma in social sciences. The classical understanding of the "state" comes from Weber (1919), who defined it as an entity which "upholds the claim to the monopoly of the legitimate use of physical force in the enforcement of its order" (Weber, 1978, p. 54). This implies that we should be looking for evidence of the initial monopolization of power within a certain territory.

However, there is also the question of the extent of this original jurisdiction: how large is the population or the territory subject to the power monopoly? Is, for instance, a village with 100 tribesmen, led by a chief, large enough to classify as "state"? It appears that we need to find an appropriate threshold to distinguish between small and large scale political organization. Therefore we adopt the convention that, although simple chiefdoms fall short of being states, a paramount chiefdom which incorporates multiple individually substantial chiefdoms can be understood as a form of incipient state. Hence we decided to begin according partial weight when a polity reaches this level. By this convention, for instance, the land of what is today Belgium came under large-scale political organization for the first time between 59 and 52 BCE, when it was integrated in the Roman Empire.

This agrees with established sociological and anthropological taxonomies of human societies throughout their evolution. For instance, Johnson and Earle (2000) proposed a division of societies into small-scale local group (further divided into family, village and the Big Man group) and larger-scale regional polity, which can be a chiefdom or a state. This distinction goes back even earlier, to Charles Tilly: "the term [state] includes city-states, empires, theocracies, and many other forms of government, but excludes tribes," (Tilly, 1990, p. 1) and to Service's (1962) proposed typology of bands/tribes/chiefdoms/states.[6]

2.3 State history, fiscal capacity and the economy

The key assumption underlying our story of reversal of fortune is that up to a point, accumulated state history favors capacity building, taxation and the provision of public goods, which in turn spur economic growth. But beyond a certain level, state experience is conducive to the build-up of extractive institutions and the rise of powerful elites that appropriate tax revenue rather than turn it into public goods and thereby undermine the entire economy. This assumption has three distinct implications that need to hold for it to be valid: 1) that fiscal capacity is conducive to economic growth; 2) that young, inexperienced states lack the ability to build a solid tax infrastructure; 3) that leaders in old, very experienced states often tend to misuse the tax revenue at the expense of the economy, despite having access to a solid fiscal capacity. We discuss these elements in turn below:

Firstly, there is increasing evidence that a consolidated bureaucracy enables financing large public projects and technological innovation, and investments in effective warfare, thus spurring economic growth. This is empirically supported by recent studies on the historical role of state capacity, revealing strong correlations and potential causal links between administrative infrastructure, high taxes and economic prosperity (Besley and Persson, 2013; Dincecco and Katz, 2014).

Secondly, there is widespread agreement that a short state history generally implies weaker fiscal capacity. This has recently been discussed among others by Tilly (1990), Collier (2009), and Besley and Persson (2013).

However, the third supposition that older and more autonomous states are more predisposed to develop and maintain abusive power structures, while intuitively appealing, is less evident. The idea that states naturally develop from a basic need to sustain collective action in large communities, particularly in response to attacks by predators, is advanced by Tilly (1990) and Olson (1993). Olson goes on to argue that in the face of theft from "roving bandits", it is welfare enhancing to have one

[6]We thank Jacob Gerner Hariri for useful references on the matter of state definition.

member of the community set himself up as a dictator that collects taxes which he uses partly for defense, but mostly as selfish rent extraction. This autocrat then becomes a "stationary bandit", interested to advance productivity and income only to the extent that he can extract more rents. This would work in the long run if the autocrat was able to commit to always provide a certain level of public goods like defense, private property and insurance against the risk of expropriation of his subjects. However, by nature of dictatorships, there are no commitment devices for an autocrat. Thus, the autocrat is susceptible to breaking his promises when he takes a short-term view of his reign when, for instance, his position is uncertain, or when there is no clear heir to his throne. The autocrat then extracts even more rents, behaving like a "roving bandit". Hence, concludes Olson, in an autocracy good economic performance is unsustainable in the long run.

Complementary data-based evidence comes from Hariri (2012) who presents causal estimates of the effect of state history on autocracy. Using an instrumental variable approach, he shows that older indigenously formed states are more likely to develop autocratic institutions than later states, his main proposed channel being that older states more successfully fend off attempts at colonization, and hence do not transplant democratic institutions from western colonizing powers.

Besley and Persson (2013) and Dincecco and Katz (2014) take the view that political regimes characterized by non-inclusive institutions and regime instability do not even manage to build fiscal capacity.[7] In a similar spirit, Gennaioli and Voth (2014) argue that success in warfare in early history was not necessarily contingent on military investment, and therefore did not stimulate tax collection as much as it did in the past five centuries (with the advent of the military revolution).[8]

The oldest states were indeed subject to numerous regime changes due to raids, internal strife and shifts in the locus of power.[9] However, we take a slightly different angle, in the spirit of Olson (1993), and argue that even if older states had built bureaucratic infrastructure, their propensity to be autocracies under regime instability led to over-extraction of tax rents to the detriment of economic activity.

Similar conclusions stem from the finding that economic development in countries with old civilizations typically lag behind the countries with an intermediate length of state history like the United Kingdom and the Scandinavian countries, discussed

[7] Using 1600-1913 data from 11 European countries, Dincecco and Katz (2014) present causal evidence that historical reforms of tax centralization and, to some extent, constraints on the executive led to higher tax revenue, improved infrastructural power, which then stimulated economic growth. By contrast, states with more autonomous regional authorities generated smaller revenues.

[8] A classic example is that of China's "Warring States" period, 475 - 221 BCE)

[9] We capture regime instability and power fragmentation in our *Statehist* index through a downgrade in the origin of the rule and territorial components of our index.

in recent work by Olsson and Paik (2013).[10] According to the authors, the main reasons for this "Western reversal of fortune" since the onset of agriculture were institutional: the old civilizations developed autocratic, hierarchical societies that were not conducive to the emergence of democracy and innovation, which became critical factors for economic growth during the modern era.[11] The more peripheral regions, which were slower to develop state institutions, were furthermore less exposed to raids by roaming armies and to incursions by migrating peoples.

In what follows, we build this assumption into a new version of the Malthusian growth model of Ashraf and Galor (2011) and present supporting empirical evidence for the model's predictions.

3 Theoretical Framework

In this section, we present the key features of our theoretical framework. The basic setting is the Malthusian growth model of Ashraf and Galor (2011). We assume a geographically well defined region where a population has made the transition to an agricultural, sedentary society. In the section below, we only outline in detail the new assumptions. A full characterization of the behavior and dynamics of the model are presented in the Appendix.

The specific aim of the model is to propose a mechanism for how states interact with economic development during the agricultural era. The key novel features of the model are that we introduce the rise of an early state which taxes individuals but also provides public goods. Both fiscal capacity as well as the centralization of power increase with accumulated state experience. The centralization of power initially has a positive impact on the effective provision of public goods and economic growth but might eventually transform into having a negative impact when the constraints and checks against governments become too weak.

3.1 State history

Let us begin by specifying that states only come into existence when population density has passed a critical level $\tilde{L}/X = \tilde{P} > 0$, where L is the adult population size, X is the amount of land and P is the population density. This assumption reflects

[10] Olsson and Paik (2013) present preliminary evidence showing that similar reversals appear to have been in place also in East Asia and in Sub-Saharan Africa.

[11] Wittfogel's (1957) "hydraulic hypothesis" makes the related argument that the old riverine civilizations were autocratic due to the technological nature of large-scale irrigation. See also Acemoglu and Robinson (2012) for an analysis of how countries with inclusive, democratic institutions eventually tend to dominate countries with extractive, autocratic institutions.

the stylized fact that the first state formations emerged in areas with densely populated (and sometimes environmentally circumscribed) agricultural populations. An organization in states with a king, a government and centrally provided public goods was indeed most often a natural adaptation to Malthusian population pressures.[12]

Let s_t be an indicator of whether a state exists or not at time t such that we can also define a stock of cumulative state history index S_τ as:

$$s_t = \begin{cases} 1 \text{ if } P_t \geq \tilde{P} \\ 0 \text{ if } P_t < \tilde{P} \end{cases} \quad ; \quad S_\tau = \sum_{t=0}^{\tau} (1+\rho)^{t-\tau} \cdot s_t$$

τ is thus the contemporary time period and the starting date is $t = 0$ when a first state was founded. The scores on the s_t-indicator are depreciated by a time discount factor $\rho \geq 0$ where time t ranges from the initial period 0 to the current date τ: $t \in \{0, ...\tau\}$. The depreciation process implies that state experience closer to the current period τ is more important for the ability to provide public goods and broadcast centralized power than experiences further back in time.

3.2 Behavior

Let us imagine an overlapping generations framework where a representative individual lives for two periods, *childhood* and *adulthood.* All key choices are made in adulthood. The individual has a utility function given by $u_t = c_t^{\gamma} \cdot n_t^{1-\gamma}$ where c_t is the individual's level of consumption at time t with an associated preference parameter γ and n_t is the (continuous) number of children of each adult person. One unit of time might be thought of as a generation.

The adult individual farmer earns an income y_t which can be used for either child-rearing, consumption, or paying a tax to a state. The budget constraint is $c_t + \mu n_t \leq y_t - s_t\theta(S_t)$. The parameter $\mu > 0$ is the cost of rearing one child and $s_t\theta(S_t) \geq 0$ is a lump-sum tax where s_t is the binary indicator defined above, describing whether there is a state or not during the particular period t. The tax level $\theta(S_t)$ is a positive function of the historical accumulated experience of a state $S_t \geq 0$. We will define taxation and the involvement of a state in the economy further below.

[12] In a related paper, we show empirically that the timing of Neolithic transition is the most important factor in the emergence of states and strongly influences the accumulation of state experience (Olsson et al., 2015). In this model we also assume that X is above some reasonable minimum scale, i.e. more than a hectare or two. For simplicity, we do not explicitly formalize this sensible but minor detail in the notation. Please see previous sections for an overview of this literature.

As shown in the Appendix, the utility-maximizing level of fertility at t is $n_t^* = (1-\gamma)(y_t - s_t\theta(S_t))/\mu$. Fertility thus increases with after-tax income $(y_t - s_t\theta(S_t))$ and decreases with cost per child μ.

Aggregate production is given by a standard production function $Y_t = A_t X^{1-\alpha} L_t^{\alpha}$ where Y_t is output, X is the amount of land, A_t is total factor productivity, L_t is the size of the labor force (equal to the number of live adults in period t), and $\alpha \in (0,1)$ is the output elasticity of labor. If we define population density as $P_t = L_t/X$, we can express output per adult individual as $Y_t/L_t = y_t = A_t P_t^{\alpha-1}$.

3.3 Public goods

Taxes are used for the provision of public goods. Public goods are large-scale utilities that only regions with an existing state can provide such as defensive walls, large-scale irrigation, infrastructure, property rights, rule of law, science, money, and safety along trade routes. Publicly provided goods like these obviously enhance productivity for farming populations by lowering transaction costs of production and trade. The effective provision of public goods $G_t \geq 0$ requires the collection of taxes $\theta(S_t)$ which we assume are a positive function of state experience S_t. More precisely, we assume that $\theta(0) = 0$, $\lim_{S_t \to \infty} \theta(S_t) = \bar{\theta} < 1$, and that $\theta'(S_t) > 0$ at all $S_t > 0$. The longer the state experience, the greater the fiscal capacity to collect taxes. In a sense, there is thus a learning-by-doing effect from state history.[13]

The effective provision of public goods not only depends on the physical resources $\theta(S_t)$ extracted through taxation but also on the effective level of *centralized coordination* $Z(S_t)$. This refers to the institutions in place to initiate, coordinate and enforce collective action in large public good projects such as the construction of a defensive tower or the digging of irrigation canals. $Z(S_t)$ might thus be thought of as reflecting how efficiently physical resources from taxation $\theta(S_t)$ are transformed into actual public goods. In the early stages of development, centralized power made possible the creation of public goods previously unseen and most likely enhanced productivity greatly. Centralization also provided sufficient law and order (or suppression of conflict) among immediate neighbors so that people could settle thickly, and invest their energies in production while benefiting from specialization and trade. $Z(0) = 1$ means that when a state first arises, the government manages to transform $\theta(S_t)$ into G_t in a one-to-one manner.

[13] Although we do not model it here, one might argue that there is potentially also a reverse causality in the sense that the ability to tax the population is an important factor behind the emergence and sustainability of states.

However, centralized power without constraints might also lead to the enrichment of a small elite at the expense of the masses and the stifling of individual initiative. Extremely centralized political and economic power implies very weak incentives for individual effort and a culture of counter-productive rent seeking and corruption. Hence, we assume that $Z(S_t)$ is at first an increasing and then a decreasing, concave function of state experience such that $Z'(S_t) > 0$ at low levels of S_t and that $Z'(S_t) < 0$ at higher levels.

We assume that the effective provision of public goods G_t is a multiplicative function of physical tax resources $\theta(S_t)$, centralized coordination $Z_t(S_t)$, and the binary indicator s_t describing whether there is a state or not during the particular period t:

$$G_t = s_t \cdot \theta(S_t) \cdot Z(S_t)$$

Here we assume that public goods are provided only when state institutions are in place, i.e. $s_t = 1$. We will henceforth present all dynamics under this assumption. This is a convenient simplification, but in the Appendix we also present a version where there is taxation even in the absence of a state.[14]

A $Z(S_t) < 1$ implies that a fraction $(1 - Z(S_t)) \cdot \theta(S_t) > 0$ of public resources is wasted or counterproductive due to mismanagement, corruption, or crowding out of private activity.

3.4 Productivity

We assume that during the Malthusian era, the total productivity or technology variable A_t in the aggregate production function depends on two key factors; the quality of the *natural environment* for agriculture N and the region's effective provision of *public goods* G_t.

The natural environment N includes biogeographical factors such as the access to suitable plants and animals for domestication (Olsson and Hibbs, 2005) but also the quality of the soil, the annual patterns of precipitation, the prevalence of frosts,

[14]There might be instances when old states with a lot of state history and a high S_t collapse so that $s_t = 0$. Despite the region's long history of governments, our assumption implies that public goods are provided neither when a state collapses, nor before it emerges. If a state should arise again so that $s_t = 1$, that new state organization can benefit from the accumulated (time discounted) experience S_t from previous governments. An example is the Indus Valley, where the early Harappan states (emerged cca 2000 BCE) broke down around the 17th century BCE, with no evidence of political organization, urban developments, or any public good provision for an entire millennium. Yet, the Harappan civilization has a bearing in the total accumulated state experience in the territory of India and Pakistan, since its technological and political innovations influenced subsequent development.

etc. For simplicity, we imagine that N is a constant and that A_t increases linearly with N.[15]

State-provided public goods like security, roads, irrigation, and market places greatly enhanced the aggregate level of productivity in agricultural societies. We propose a very simple specification for productivity:

$$A_t = A(N, S_t) = N + G_t(S_t) = N + \theta(S_t) \cdot Z(S_t) \tag{1}$$

Hence, a key feature of the model is that

$$\frac{\partial A_t}{\partial S_t} = A_S = \theta'(S_t) \cdot Z(S_t) + \theta(S_t) \cdot Z'(S_t) \gtreqless 0.$$

Given the concavity of $Z(S_t)$, we can define the productivity maximizing level of state history as $S_t^* = \arg\max [\theta(S_t) \cdot Z(S_t)]$. When $S_t = S_t^*$, further state experience is going to induce a net negative impact on productivity through the negative effects of too much centralized power. Of course, if the link between state history and the degree of centralized coordination $Z(S_t)$ is very weak, S_t^* would be very high and potentially outside the empirically relevant range. In that case, we would only observe a positive relationship between A and S. We will return to this issue in the empirical section.

3.5 Dynamics

If a state exists so that $s_t = 1$, it is demonstrated in the Appendix that the equilibrium population density will converge towards a level

$$\bar{P}_t^s = \left(\frac{(1-\gamma)\left(N + \theta(S_t) \cdot Z(S_t)\right)}{\mu + (1-\gamma)\,\theta(S_t)} \right)^{\frac{1}{1-\alpha}}. \tag{2}$$

This expression has taxes $\theta(S_t)$ both in the numerator and in the denominator. The intuition is that taxation, on the one hand, decreases optimal fertility since taxes mean that less resources are available for raising children. This is the effect in the denominator. On the other hand, taxes also increase the provision of public goods and hence productivity and income per person, which has a positive impact on the number of children (in the numerator). It should also be remembered that the level of taxes $\theta(S_t)$ increases monotonically with S_t.

The level of centralized coordination $Z(S_t)$ only enters in the numerator. This stems from the impact of $Z(S_t)$ on aggregate productivity and on income levels per

[15] In reality, we know that the natural environment will be affected over the long run by human exploitation, climate change, etc. We abstract from these aspects below.

adult.[16] Due to the specific nature of $Z(S_t)$, centralized coordination will have a positive impact on population density during a first phase of state experience. As state history accumulates, government may tend to become more and more exploitative and tax resources may be wasted in corruption and embezzlement. Moreover, the effective level of public goods includes effectiveness at enhancing the population's productivity, and effectiveness can go down from overmeddling or regulating (like China banning overseas commerce in 15th century or Louis XIV over-regulating France). Eventually, aggregate production will start to fall and population density will decline in response.

3.6 Industrial era

In this section, we will briefly explore the channels through which state history might have an impact even on post-Malthusian, modern economies.

In the industrial era, land is no longer a central factor of production and it is standard to model aggregate output as a function of productivity or technology A_t, labor L_t, and physical capital K_t such that $Y_t = F(A_t, L_t, K_t)$. If we assume a labor-augmenting technology so that $Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}$, then output per worker (or per capita) can be written as $y_t = Y_t/L_t = A_t k_t^{\alpha}$ where $k_t = K_t/A_t L_t$ is capital per unit of effective labor.

Another standard assumption is that technological progress during the industrial era grows in each period at a percentage rate $g > 0$ which we assume to be exogenously given and not a function of the natural environment or of state-provided public goods.[17] In line with the expression in (1), let us define the level of productivity at the end of the Malthusian era $t_M > 0$ to be $A_{t_M} = N + \theta(S_{t_M}) \cdot Z(S_{t_M})$. The level of productivity in some industrial era time period $\tau > t_M$ is then

$$A_\tau = A_{t_M} \cdot (1+g)^{\tau - t_M} = (N + \theta(S_{t_M}) \cdot Z(S_{t_M})) \cdot (1+g)^{\tau - t_M}$$

As demonstrated in the Appendix, the equilibrium level of log output per capita in this economy is:[18]

$$\ln \bar{y}_\tau = \ln(N + \theta(S_{t_M}) \cdot Z(S_{t_M})) + (\tau - t_M)\ln(1+g) + \frac{\alpha}{1-\alpha} \ln\left(\frac{s}{g+n+\delta+ng}\right). \tag{3}$$

[16] Note that $A_t = N + \theta(S_t) \cdot Z(S_t)$.

[17] In a globalized industrial economy, important innovations tend to spread geographically through technological diffusion from the technological frontier.

[18] For an analysis along the same lines, see Mankiw et al (1992).

Thus, in this model we make the simplifying assumption that state history solely affects current levels of prosperity through its impact on the level of productivity at the time of the country's industrial revolution.[19]

Comparative statics show that

$$\frac{\partial \ln \bar{y}_\tau}{\partial S_{t_M}} = \frac{A_S}{A_{t_M}} = \frac{\theta'(S_{t_M}) \cdot Z(S_{t_M}) + \theta(S_{t_M}) \cdot Z'(S_{t_M})}{N + \theta(S_{t_M}) \cdot Z(S_{t_M})}$$

Since the denominator of this expression is always positive, the sign is determined by whether the numerator is positive or negative. As was discussed above, the fact that $Z(S_t)$ is a concave function of S_t means that, once again, the derivative will be positive at low levels of S_{t_M} and negative at high levels of S_{t_M}.

3.7 Key predictions

On the basis of the model, we have derived three key hypotheses:

First, total factor productivity A_t during the Malthusian era should have a concave, inverted u-shaped relationship with accumulated state history S_t. Given the key role of our conjecture that levels of centralization $Z(S_t)$ eventually had a negative impact on the provision of productivity-enhancing public goods (e.g. education or secure property rights, necessary to ensure a favorable climate for technological innovation), we start by testing this relationship in the empirical section.

Second, the Malthusian era population density $\bar{P}_t^s$ in (2) will also have a concave, inverted u-shaped relationship with S_t. The existence of a state means the individual's resources are taxed, which leaves less for children. However, the positive effect of a state on productivity will dominate initially but then turn negative when centralization lowers the effective provision of public goods and strangles individual effort, which leads to a stagnation or even decrease in fertility.

Third, since the starting level of productivity on the eve of the industrial revolution is a concave function of state history, a final prediction from the model is that there should be a persistent, non-linear effect of state history on contemporary levels of prosperity.

[19]This is a very strong assumption, and in a richer model, we might also have assumed that the actual timing of the industrial revolution t_M or the level of g was affected by state history, as suggested by previous research. Empirical observation seems to suggest that it was countries with intermediate levels of state history like United Kingdom and France that made the transition first.

4 Data

In this section, we will outline how the existing index of state history has been extended to cover the BCE period. We will also briefly present some of the key tendencies in the new data series.

4.1 Constructing the index

The construction of the index for the BCE period follows the principles developed by Bockstette et al (2002), applied here to 159 modern-day countries.[20] We use evidence of written records where available. Where not extant, we rely on archaeological data, following a "diagnostic traits" approach: we consider material manifestations of the monopolization of power, as an "archaeological confirmation of the process of state formation" (Jones and Kautz, 1981, pp. 16-17). These can be monumental structures, such as palaces, temples or large urban settlements etc. In the case of Iraq, for instance, there is the transition from small to large urban centers with grand architectural structures such as Uruk in the middle of the 4th millennium BCE.[21]

The second task is to mark the transition from chiefdom to fully-fledged state. Following the paradigm of the evolution of pristine states from chiefdoms (see e.g. Carneiro 1981, Earle 1987, Flannery 1995, Marcus 1992, Spencer 1990, Spencer and Redmond 2004), we mark this distinction in our data by assigning the following values: Band/tribe is marked by a rule score of 0, paramount chiefdom is assigned 0.75 and fully-fledged state receives the value 1. Robert Carneiro emphasizes that the paramount chiefdom is the evolutionary link between autonomous bands or tribes and the state.[22]

While it is difficult to know exactly where the chiefdom ends and where the state begins in pre-history, we have made efforts to draw a sensible line where the evidence suggests a noteworthy evolution in socio-political organization.[23] While this

[20]The reader is referred to the online Appendix C for a more detailed discussion of the coding procedures and exceptions.

[21]Admittedly, the drawback of this "symptomatic" approach is that it blurs the boundary between state and civilization and it is susceptible to misclassifying an emerging or transient civilization into a state in the Weberian sense.

[22]In his definition, the paramount chiefdom is "an autonomous political unit comprising a number of villages or communities under the permanent control of a paramount chief" (Carneiro, 1981, p. 45), while the state is "an autonomous political unit, encompassing many communities within its territory and having a centralized government with the power to collect taxes, draft men for work or war, and decree and enforce laws" (Carneiro, 1970, p. 733).

[23]Such is the case of Mexico, where we assign a score of 0.75 to the period 450 - 100 BCE for the early urban settlements at Chiapas and Oaxaca. We then raise this score to 1 in 100 BCE when

approach is not uncontroversial, it is the most feasible given limited documentary resources. We further detail the assignment of scores in a later section.

For each country, the time of emergence of the first state institutions on its territory is identified, as defined above. *State age* is defined as the time elapsed from this date until 2000 CE. The oldest state was established on the land of today's Iraq around 3500 BCE. Hence, the time for all countries is divided into 110 periods of 50-years.

For each country i and half century t, scores are assigned to reflect three dimensions of state presence, based on the following questions:[24]

1. Is there a government above the tribal level? Score component z_{it}^1 receives 1 point if yes, 0.75 if the government can at best be described as a paramount chiefdom and 0 points if no government is present.

2. Is this government foreign or locally based? z_{it}^2 is 1 if the rule is locally based, 0.5 if externally based, and 0.75 for local government with substantial foreign oversight.[25]

3. How much of the territory of the modern country was ruled by this government? z_{it}^3 reflects the proportions of the territory under some rule: 1 (over 50 percent), 0.75 (25-50 percent), 0.5 (10-25 percent), 0.3 (under 10 percent).[26]

Time is indexed by t and refers to a 50-year period ranging from $t = 0$ for 3500-3451 BCE when the first state arose, to $t = 109$ for 1950-2000 CE). For every such time interval, we compute a composite *State index* score by multiplying the three components by one another and by 50:[27]

$$s_{it} = z_{it}^1 \cdot z_{it}^2 \cdot z_{it}^3 \cdot 50 \tag{4}$$

Finally, joining the BCE- with the preexisting CE-era series, we aggregate all "flow" scores s_{it} into *Statehist* - the comprehensive index of the cumulative state

large-scale urban growth at Teotihuacan and the development of previously missing institutions such as a standing army warrant the status of fully-fledged state.

[24]Each dimension is denoted by z_{it}^c, which is the score for component c in country i for period t.

[25]If there were multiple polities within a present country's borders, its state score for the period is coded as a simple average of their respective scores.

[26]For multiple contemporaneous states within what is now a single country z_{it}^3 is adjusted down one category, because centralized coordination is assumed to decrease.

[27]Within period changes in z_{it}^c require averaging the scores over subperiods, using as weights the number of years in each sub-period θ divided by 50:
$s_{it} = 50 \cdot \left[(z_{it1}^1 \cdot z_{it1}^2 \cdot z_{it1}^3) \cdot w_{it1} + (z_{it2}^1 \cdot z_{it2}^2 \cdot z_{it2}^3) \cdot w_{it2} + \cdots \right]$.

history.[28] The index is normalized by the score of a hypothetical state with full discounted scores between 3500 BCE and the period of interest τ:

$$S_{i\tau} = \frac{\sum_{t=0}^{\tau}(1+\rho)^{t-\tau} \cdot s_{it}}{\sum_{t=0}^{\tau}(1+\rho)^{t-\tau} \cdot 50} \tag{5}$$

This cumulative *Statehist* index $S_{i\tau}$ ranges from 0 to 1 and can be calculated at virtually any point in history $\tau = \{0, 1, ...109\}$. We calculate it mainly for 1500 CE (at $\tau = 99$) and for 2000 CE. We can use various rates $\rho \geq 0$ for discounting historical scores. The previous literature has set the convention at $\rho = 0.05$, in light of the reasonable assumption that the more distant past matters less today than recent history. With the additional data, however, this rate gives insufficient weight to the long stream of s_{it}-scores before 1 CE.[29] While it of course remains to be seen below just how useful placing weight on the distant past will be, our convention is to employ the 1 percent discount factor of the normalized *Statehist* score.[30]

To answer the three questions (a-c) above in a manner that is consistent across periods, we relied mainly on information in the Encyclopedia Britannica Online. We provide additional detail on our data sources and illustrate the coding process and further data aggregations in the online Appendix C. Accompanying this paper is also an online Data Coding Appendix, which provides a comprehensive list of coding decisions for all country-period observations.

4.2 A brief look at the data

In this section we present some patterns that arise from the complete state history time series and the data used in forthcoming analyses.

Firstly, we note that the evolution of state institutions in the world follows approximately an exponential upward trend with periods of rapid growth punctuated by periods of stagnation (Figure 1). The graph shows the log of the aggregated percentage score for all contemporary countries in our sample at each 50-year period on the vertical axis and year on the horizontal axis. The percentage score in period t is calculated as $\textit{State index world}\ \ (t) = 100 \cdot \sum_{i=1}^{N} s_{it} / (N \cdot 50)$ where $N = 159$ is the number of included countries and where $s_{it} \in [0, 50]$ is the state history score for

[28] Some minor adjustments were made to the original CE index, but the correlation with the initial index, considering year 1 to 1950 CE periods only, is 99 percent.

[29] The extended *Statehist* score (for the 3500 BCE to 2000 CE period) has a correlation of up to 99.3 percent with the 5 percent discounted 1 - 1950 CE score.

[30] The 1 percent discounted Statehist index at 2000 CE has a 0.93 correlation with the 1 - 1950 CE 1 percent discounted Statehist index and 0.86 correlation with the 1 - 1950 CE 5 percent discounted Statehist index.

country i during 50-year interval t, as described above.[31] A value close to 0 percent in this world index indicates that there is no sign of state presence in any of the included countries in period t whereas a score of 100 means that all 159 countries reach the maximum value $s_{it} = 50$ in our state measure during that period.[32]

Several periods are characterized by rapid state evolution whereas other periods are marked by a general decline. The first boom in state emergence appears already in 3500-2300 BCE, which then ends with a long period of stagnation. The other major stagnations in the figure happened around 1750 BCE, 1200 BCE, and 400 CE. A second period of rapid growth was 850 BCE-1 CE during the Iron Age. From just after the collapse of the Roman empire around 450 CE, aggregate state emergence shows a steady upward trend.

This pattern is also visible in Figure 2, which in addition shows the regional aggregated percentage score for all contemporary countries in our sample (this figure displays world and regional averages of the *Statehist* index, rather than the natural log of that average as in Figure 1, allowing the reader to get a sense of the trend in a form some may find more intuitive). We disaggregate the evolution of state history into the four main agricultural core areas: Western, Eastern Asia, Sub-Saharan Africa, and the Americas.[33] These four areas are created on the basis of how Neolithic agriculture and civilization spread during early historical times.

When we divide up the world in this way, some striking historical differences between the regions appear: State evolution started earliest in the Western area, with Eastern Asia lagging behind until rough convergence (indeed, initially overtaking) around 500 CE, with the other regions gaining steam later and all converging only toward the end of the era of European colonialism. State emergence was earliest in Eastern Asia and in the Western region. Interestingly, both of these early civilizations took off on a more rapid path after 850 BCE. By the time of the Western Roman collapse after 450 CE, Asian state development overtakes the Western one for the first time.[34]

[31] Note that *State index world(t)* describes the "flow" level of state development in the world in period t and not the cumulative "stock" of state experience.

[32] Since many modern-day countries did not have full states in the spirit of our definition during the entirety of last time period 1950-2000, the aggregate percentage in the graph is about 88 percent at the end of the time series. Many states were de-colonized part way through the period, a number emerged from the Soviet Union and Yugoslavia, others experienced contending governments or state failure, etc.

[33] The division into agricultural core areas follows the practice in Morris (2010) and Olsson and Paik (2013) (see also Diamond and Bellwood, 2003). Combining the two or three distinct agricultural cores of the Americas identified by some writers is a convenient simplification.

[34] See Morris (2010) for a detailed comparative analysis of Western and Eastern history since the Neolithic.

The other two regions, the Americas and Sub-Saharan Africa, clearly lag behind, in particular after the Eurasian turning point 850 BCE. From about 500 CE, the pace of state emergence starts to increase in Sub-Saharan Africa. When the colonial era starts in the late 15th century CE, the lagging regions experience a dramatic increase in the *State index*. This increase is of course to a great extent driven by the emergence of colonial states, created by European powers. By the final period of observation (1951-2000), the Americas has the highest score on state presence among all regions in the world.[35]

The *Statehist* index and other variables related to state experience, as well as outcomes and control variables used in all forthcoming analyses are summarized in Table 1 below. Full definitions of each variable are given in the online appendix.

5 State History and Economic Development

We now proceed to analyze the relationship between state history and pre-industrial as well as current economic development.

5.1 State history and pre-industrial economic development

State history and productivity in 1500

We begin with the empirical question of the relationship between state history and productivity in the Malthusian era. The first prediction of our model was that total factor productivity should increase up to some level of accumulated state history, beyond which it may decrease, owing to the negative effects of centralized power on the coordination of economic activity and public good provision. Since accumulated state history may have adversely affected productivity by stifling innovation, we proxy productivity with the average index of technology adoption constructed by Comin, Easterly and Gong (2010). Using various data sources on the presence and complexity of various technologies, the country-level index captures advances in five sectors: agriculture, transportation, communications, writing, and military. The index is computed for 1000 BCE, 1 CE, 1500 CE and 2000 CE, using slightly different approaches, which we describe in some detail in Appendix B.

[35] In Figure D1 of the online appendix, we zoom in on the last 550 years of state history and show trends for Western Offshoots including the U.S. and Canada (along with Australia and New Zealand) and for the rest of the Americas (Latin America and the Caribbean). The latter two regions are shown to come from behind to overtake even Europe in internally controlled state presence by the mid-1800s.

In order to test this prediction, we set up the following model:

$$Technology1500_i = \beta_0 + \beta_1 \cdot Statehist1500_i + \beta_2 \cdot Statehist1500_i^2 + \beta_j' \cdot Z_i + \beta_k' \cdot X_i + \lambda_c + \epsilon_i \quad (6)$$

On the left hand-side of equation (6) we have the average technology adoption index in 1500 CE. On the right-hand side we include our main independent variable, *Statehist* (the cumulative index shown in equation (5) accumulated in 1500 CE), both linear and squared, to account for the potentially quadratic relationship. The *Statehist* index is normalized with respect to 3500 B.C.E - 1500 CE and computed using a 1 percent discount rate per period. Z_i is a vector of historical controls including: $Agyears_i$, the time before present since the Neolithic transition in the country-area in question, a variable taken from Putterman and Trainor (2006); $Origtime_i$ - the approximate time since the first settlement on the territory of the modern-day country by anatomically modern humans, a variable introduced by Ahlerup and Olsson (2012) as a determinant of the variation in levels of ethnic diversity across the world. In a more flexible specification, we include the square of $Origtime_i$ and a linear control *State age*$_i$. X_i is a vector containing geographic controls. These include: latitude of the centroid of the modern-day country i, whether the country is landlocked, its distance to coast or ocean-navigable river, average elevation, the land suitability for agriculture, climatic variables for temperature and precipitation, and the risk of malaria.[36] λ_c is a vector of continent fixed effects. The results are displayed in Table 2. Columns (1)-(4) present the results without controlling for geographic characteristics. In columns (6)-(7) we present the results using continent fixed effects.

Our main coefficients of interest are β_1 and β_2, which estimate the relationship between pre-industrial level of technology adoption and state experience. In column (1) we display the simple association between technology adoption and *Statehist,* which is positive and significant. In column (2) we add the squared *Statehist*, and the estimates of the *Statehist parameters* display the concave pattern predicted by the model: both coefficients are significant at 1 percent, β_1 is positive, while β_2 is negative.

We move directly to column (4) in panel A, where we add to the model the first historical control - *Agyears* (shown to be positively significantly correlated with the dependent variable in column 3, for comparison purposes). Its inclusion only slightly changes the signs and the magnitude of the coefficients of the *Statehist* terms.

[36]These variables are taken from the Portland Physical Geography dataset and from the dataset compiled from various other sources by Ashraf and Galor (2013). See the online appendix for more details on variables' construction and collection.

Moreover, the effect of the time from transition to agriculture is reduced relative to column (3).[37] When we also add *Origtime* and geographical controls in column (5), the magnitude of the estimates changes slightly, but the relationship remains concave. In columns (6) and (7) we learn that the quadratic relationship holds after controlling for continent fixed effects as well. The introduction of continent fixed effects wipes out the estimate of *Agyears*, while the estimate on *Statehist* squared remains significant, albeit diminished.

The last column accounts for the age of states and also for recent developments in the literature postulating that the patterns of human settlement in prehistory may have complex effects on later economic development (Ashraf and Galor, 2013). By introducing the squared *Origtime* variable, we allow for a nonlinear relationship in the time since first human settlement. However, the coefficients of the terms containing *Origtime* are insignificant, while the *State Age* control has a negative and significant, albeit small effect.

State history, population and urbanization in 1500 CE

Having established that state history is related to pre-industrial levels of productivity in a non-monotonic fashion, we also inquire whether this pattern is reflected in population density at 1500 CE, the second prediction of our model. We also ask whether state experience made a difference for the level of economic development in 1500 CE, measured through urbanization rates.

In Table 3 we display the results from the model explaining population density in 1500 CE (panel A) and urbanization rate in 1500 CE (panel B). All specifications are analogous to those in Table 2.[38]

In the first column in both panels we see that extended *Statehist* is positively and significantly correlated with past population density and urbanization. Interestingly, in the second column, where we introduce *Statehist* squared, both coefficients are highly significant, displaying the same quadratic relationship with the left-hand side variable as uncovered in Table 2. These unconditional estimates convey that the positive impact of an increase in state experience on population density diminishes up to a point where it becomes null. Beyond this point, increased experience with state institutions impacts negatively on population density. For historical urban-

[37] As a robustness check, we have redone the estimations using the overall technology adoption index excluding the agriculture components; the results are very similar. See Table E5 in the appendix.

[38] We also fitted multiple regression models where the economic outcomes depend linearly on *Statehist*. The results displayed in Tables E6-E9 in the appendix reveal that these models have slightly lower explanatory power than the models allowing for quadratic *Statehist* effects.

ization rates, we observe the same quadratic pattern displayed by the coefficients of *Statehist* and *Statehist* squared, which are significant at conventional levels for outcomes in 1500 CE. The non-linear pattern is robust to all changes in specification and it holds even with continent fixed effects. We note also that the inverted-u relationship for these outcomes as well as technology adoption is also observed for all outcomes in year 1 CE, but it becomes insignificant when we gradually introduce controls in the regressions (see online appendix Table E3 - panels A and B, and Table E4, panel A). The estimated state history effects are net of the contribution of early transition to agriculture, which was still influencing positively the population density in 1500 CE.

5.2 State history and current economic development

The third prediction of our model was that accumulated state history should affect current economic development non-linearly, due to the persistence of its effects (positive at low levels, negative at high ones) on productivity in the pre-industrial era. To investigate whether this is indeed the case, we start by estimating our model with technology adoption in 2000 CE as a quadratic function of state history. The results are displayed in Table 4 below. In panel A, we regress technology adoption on *Statehist* in 2000 CE.

However, when analyzing the current levels of technological sophistication, using the raw *Statehist* data means that we only account for the history within the territories of modern-day countries. This ignores the state history of other territories from which people migrated in recent centuries to settle in new territories. Population flows after 1500, when the era of colonization began, are instrumental in mapping the impact of historical events to today's economic performance. This is because the ancestors of today's population have evidently brought with them the history, the know-how and the experience with state institutions from their places of origin (Putterman and Weil, 2010; Comin et al, 2010; Ashraf and Galor, 2013).

We therefore also use an alternative measure of state history which is obtained by adjusting the 1500 CE *Statehist* index with the migration matrix developed by Putterman and Weil (2010). We then re-estimate our model using this new measure - *the ancestry-adjusted Statehist* - which, for each country, represents the average pre-industrial *Statehist* of its year 2000 population's ancestors, with the weights for each source country being the share of then-living ancestors estimated to have lived on its present-day territory. These alternative results are displayed in Table 4, panel B.

Remarkably, the same concave relationship emerges when the dependent variable is the average index of technology adoption in 2000 CE. Furthermore, using the ancestry-adjusted *Statehist* in 1500 CE to explain the differences in average technology adoption in 2000 yields significant estimates in all specifications, with larger magnitudes and higher R-squared statistics than using the *Statehist* in 2000 CE. Thus, the relationship between state experience and technology was indeed concave in the Malthusian era, and it was transmitted, albeit in weakened form, all the way into modern-day levels of technology adoption (compare Table 4 to Table 2). This result is so far consistent with the predictions of our model.

The last prediction of our model was that there should also be a persistent non-linear effect of state history on contemporary economic prosperity, owing to its non-linear effect on productivity before the industrial revolution. We test this prediction by comparing per capita GDP levels in 2000, as a function of accumulated state history.

Figure 3 illustrates the essence of our findings. On the Y-axis we have the logarithm of GDP per capita in 2000 and on the X-axis we have the extended *Statehist* (normalized with respect to 3500 B.C.E - 2000 CE and computed using a 1 percent discount rate per period).

The figure displays a scatter plot of all countries in the sample, while also allowing for a *quadratic* fit of the relationship between output and *Statehist.* A hump-shaped relationship emerges when using the extended *Statehist.*[39] The immediate implication is that states with extreme values of *Statehist* fare worse in terms of per capita GDP in 2000 than states with intermediate levels of *Statehist*, as measured by the extended index. In the online appendix, we show that the relationship between year 2000 income and the state history index based on 1 - 1950 CE data does not display the downward sloping portion of the inverted U seen in Figure 3.

However, Figure 3 provides only an initial impression. The inverted U pattern is not especially apparent in the scatter of points, and in any case the relationship plotted is not conditioned on the influence of other variables. The question arises whether the apparent relationship between *Statehist* and income mainly reflects other historical forces at play or natural conditions which may have shaped both the history of state institutions and current wealth. In order to investigate this issue, we estimate the model in (6) with the logarithm of GDP per capita in 2000 CE as the dependent variable. The results are displayed in Table 5 below. In panel

[39]This quadratic relationship is also suggested by the scatter plots displayed separately for internally- and externally- originated states (i.e. the rule was imposed from within the state's territory and from without, respectively) and when we use the ancestry-adjusted *Statehist* index. See Figures D3-D5 in the appendix.

A, we use the new *Statehist* index, while in panel B, the *Statehist 1 -1950 CE* data are used for purposes of comparison. All specifications are analogous to the ones in the previous tables.

As before, the main coefficients of interest are those of *Statehist* and *Statehist squared*, which estimate the relationship between current per capita income and state experience in 2000 CE. The unconditional correlation between per capita income and *Statehist* is positive and similar in magnitude across the two panels, but slightly less precisely estimated when the independent variable is (the new, extended) *Statehist*. In column (2) we add the squared *Statehist*, and the results mirror the pattern conveyed by Figure 3: In panel A, both coefficients are significant at 1 percent, and their signs confirm the concave relationship between log per capita GDP and state history. By contrast, in panel B, the counterpart of this specification using *Statehist 1 -1950 CE* displays coefficients with the same signs but much smaller and insignificant (the coefficient of the quadratic term even turns positive when controls are included).[40]

While *Agyears* is significantly positively correlated with modern-day GDP (column 3), when we control for it alongside the linear and quadratic *Statehist*, its inclusion hardly changes the signs and the magnitudes of the coefficients of the *Statehist* terms. Moreover, the effect of the time from transition to agriculture is insignificant. As with previous estimations, the results are robust to the inclusion of *Origtime*, as well as geographical controls and continent fixed effects.

When we control for the age of states and also introduce the squared *Origtime*, the coefficient on the latter is insignificant and state age displays a significant but small coefficient.[41] The introduction of state age diminishes the estimate on *Statehist* squared, indicating that the right extreme of Figure 3 is explained by the length of state existence (the extensive margin of state history), in addition to the overall degree of autonomy or territory considerations (the intensive margin). We note that in panel B, the main estimates when using quadratic specification of the old *Statehist* are neither significant, nor similar in terms of signs with the estimates in panel A. This speaks to the added value of the extended *Statehist* data.

Lastly, from Table 5, based on the estimates of our coefficients of interest, we can infer that the optimal predicted level of *Statehist* is reached at 0.356, which is very

[40]Note that we obtain similar estimates if we use the 1-2000 CE *Statehist* index instead, meaning that the 1950-2000 CE period is not what is driving the quadratic relationship documented in panel A.

[41]We explore alternative specifications in Tables E1 and E2 in the appendix, where we include linear and squared variables such as the time since transition to agriculture, state age, absolute latitude, migratory distance from Addis Ababa, and predicted genetic diversity (where the latter two are taken from Ashraf and Galor, 2013). Our main coefficients of interest are robust.

close to that of the United Kingdom (0.357), and most countries in Western Europe. The effects' magnitudes are not straightforward to assess from the tables. However, some numerical examples may show more clearly how the impact of an increase in *Statehist* depends on the original level of state experience. Take for instance the case of Indonesia, which has 1350 years of state existence and a *Statehist* score of 0.254. If we could hypothetically increase the *Statehist* score by 0.1 (reaching the level of the UK score), the implied approximate effect on per capita GDP in 2000 would be roughly a 20 percent increase, from USD 773 to USD 944 in 2000.[42] The opposite would happen if we were to increase the value of the *Statehist* score by 0.1 for China, which starts off with a value of 0.582: the approximate effect would be a drop in per capita GDP in 2000 by 44.4 percent.

Taken together, our estimation results so far are consistent with the last prediction of the model. Moreover, we also showed that this evidence only comes to light when we employ the new extended *Statehist* index. While these results cannot necessarily be read as causal, we take an additional step in that direction, by also estimating the model for *per capita GDP* above using the *ancestry-adjusted Stathist* index. The results, using two alternative adjustment methods, are displayed in Table 6. In panel A, we use the *Statehist* index in 1500 adjusted by the migration matrix (as in previous studies, but for the first time including full state history before 1 CE). In panel B, we use a composite index obtained by adding the raw 1500 - 2000 *Statehist* to the ancestry-adjusted *Statehist* index at 1500, which is then normalized by the full discounted score for 3500 BCE - 2000 CE. The 1500 - 2000 CE part is added in order to account for the places' histories in the past five centuries.[43]

We find that the inverted-u shape relationship between per capita income and the ancestry-adjusted *Statehist* is robust to all specifications and that the coefficients of interest are significant at 1 percent level in all columns in panel A. Moreover, the explanatory power of the model when we introduce only the ancestry-adjusted *Statehist* terms (column 2) is now 20.9 percent vs 5.2 for unadjusted *Statehist*. The results using the measure used in panel B, look reassuringly similar to those in Table 2, panel A. The fact that the column (2) R-squared is much lower when we account for the 1500-2000 portion of *Statehist* tells us that, in line with our model,

[42]The exact calculation based on estimates in column 2 of panel A is $[(7.010-2\cdot 9.842\cdot 0.254)/10]\cdot 100\% = 20.1\%$.

[43]Conceptually, the first part of the component index represents the history non-indigenous populations brought with them to their new homes in 1500 (or after), the second part the political experience they (and indigenous descendants, if any) experienced there since that time. Such a composite gives only a rough accounting for actual experience insofar as many migrants arrived long after 1500, and the timing of migration differs considerably both by receiving and by source country.

the most relevant variation for economic performance today comes from the pre-industrial era state experience. We note also that the fact that the main estimates are unaffected by *State age*, indicates that the intensive margin of the index, given by the autonomy, coherence and territorial extent of the centralized rule, is the main driver of the results. All of these elements are crucial for amassing and maintaining centralized power.

The interpretation of these results is similar, but more nuanced than when we use the raw data: territories which accumulated limited or extensive, solid and unchallenged state experience by 1500 CE, either locally or through an inflow of knowledge from migrant populations, have a lower per capita GDP in 2000 CE than those with an intermediate level of state experience. This result is also consistent with our theoretical framework where we argued that the link between state history and current levels of development should mainly be driven by what happened during the Malthusian era (i.e. prior to 1500 CE).

5.3 Discussion

To sum up, we have presented a model of the role of state experience in economic development, through its non-linear effect on productivity in the Malthusian era, which carries over into modern day productivity and output. The accompanying empirical analysis revealed consistent reduced-form regressions, where a robust concave relationship is confirmed between extended *Statehist* and technology and economic development in 1500 and 2000 CE.

Our central assumption was that the non-linearities stem from the use of taxes and provision of public goods, which are higher the stronger and more experienced a state becomes, but which can be undermined when a high level of centralized power is attained, that gives rise to provision of less efficiency-enhancing public goods and crowding out of productive enterprise. Although an extensive analysis of the causal mechanisms is beyond the scope of this paper, we believe we have offered support from the literature that this is a very plausible channel. However, we do not exclude complementary channels which could partly drive the concave relationship we uncovered on which we offer some reflections.

First, our finding appears to be consistent with the fact that while there is indeed a great deal of persistence of early societal advantages, it is also the case that the technological and institutional know-how of societies can slowly diffuse to neighbouring societies through migration or trade. These societies with younger states can then pick the best practices of the older societies and potentially avoid some of the pitfalls that might have become a drag for the old civilizations.

State capacity might be one example of such institutional transfer across state borders. The ability to levy taxes and to consolidate an administrative infrastructure has recently been shown to produce regional spill-overs to neighbouring areas' economic performance (Acemoglu, Garcia-Jimeno and Johnson, 2014). The argument is that state capacity may be more easily built around pre-existing bureaucracies (which, in this context, younger states would naturally have had access to).

Other factors that have been proposed for explaining the reversal in the Western core include environmental degradation in the Fertile Crescent and in parts of the Mediterranean region. Once agriculture spread out of the Fertile Crescent, the more robust loess soils of northern Europe, combined with a reliance on rain rather than irrigation for cultivation, proved to be an advantage in the long run (Jones, 1981). It has also been suggested that the rise and fall of dominant empires of the Western core followed cycles of expansion, over-extension, and eventually decline, with a gradual shift of power towards the northwest (Kennedy, 1989). Acemoglu et al (2005) show that the emergence of Atlantic trade after 1500 CE had a major impact on the rise of for instance Spain and the United Kingdom.[44]

A similar process can potentially explain comparative development in East Asia. Japan's less powerful central court and greater perceived vulnerability to potential Western colonizers led it to undertake decisive modernization measures almost a century before China. This development had spillover effects on Korea and Taiwan, all young states in comparison with China.

In summary, the new pattern uncovered by the extended *Statehist* shows that, beyond a certain point, a longer enduring state history is associated with economic disadvantages. While we leave it for future work to attempt to identify the exact causal mechanisms behind this pattern, we believe caution is recommended against the interpretation of these disadvantages as fully automatic and insurmountable consequences of long state histories. Our view is not that a long uninterrupted state history is always bad for economic development and as such undesirable. We believe this is a story of moderation in the exercise of centralized power and adaptability of the state institutions to the ever-changing economic realities. While those in the middle range of state history have thus far exhibited such moderation and adaptability more effectively, on average, there remains a considerable space of indeterminacy within which political actors may still exert influence over their countries' fates.

[44] Acemoglu et al (2001) argue that there was also a reversal among former colonies such that relatively less advanced pre-colonial societies had an inflow of European migrants who installed strong institutions that still persist today. Hariri (2012) argues that non-European countries with older states that resisted European colonization had worse economic outcomes in the modern era due to the persistently autocratic nature of their states.

6 Conclusions

We coded and assembled a comprehensive data series on state history from state emergence (which often occurred before the Common Era) to 2000 CE for a sample of 159 countries, building on the previously constructed *State antiquity index* of Bockstette et al (2002). Grounding our definition of state in the anthropological and political literatures, we coded three components that make up the state history index: 1) Existence of a state, 2) whether the state is home-based or imposed from without; and 3) territorial coverage of the state relative to the land areas defined by modern country borders. We obtained three overarching measures of state presence and evolution: 1) a cumulative *Statehist* index (as in Bockstette et al.), 2) *State age* (time since state emergence) and 3) a contemporaneous *State index* capturing the level of state presence at different points in time. Moreover, the availability of state history information at various levels of spatial and temporal aggregation render our data particularly versatile for a large variety of comparative analyses.

We derived a model of economic growth in the Malthusian era, where the new key element is the presence of a state that taxes its population and provides public goods. The model predicts that up to a certain point, accumulated state experience is beneficial for productivity, income and population density. However, beyond a certain level, it can have adverse effects, yielding a concave relationship between state experience and economic performance, which persists to the modern era.

In our regression analysis, we confirmed the predictions about the relationships between state history and early historical proxies for income (population density and urbanization), and technology in 1500 CE. Previous estimates using data for the period 1 to 1950 CE only had suggested a linear relationship between state experience and contemporaneous levels of economic development. Contrary to this, we showed that cumulative state history from earliest emergence to 2000 CE has an inverted-u shaped relationship with current income. We confirmed that inclusion of the BCE period is crucial to this result. Countries with extreme *Statehist* scores are worse off in terms of both current and historical economic development than countries with intermediate values of *Statehist*. The optimal level of state history as defined here is estimated to be that of modern-day United Kingdom.

References

[1] Acemoglu, D., Garcia-Jimeno, C. and Robinson, J. (2014). "State Capacity and Economic Development: A Network Approach," NBER Working Paper No. 19813.

[2] Acemoglu, D., Johnson, S., and Robinson, J. (2001). "The colonial origins of comparative development: An empirical investigation," American Economic Review 91 (5), pp. 1369 – 1401.

[3] Acemoglu, D., Johnson, S., and Robinson, J. (2002). "Reversal of fortunes: Geography and institutions in the making of the modern world income distribution," Quarterly Journal of Economics 117 (4), pp. 1231 – 1294.

[4] Acemoglu, D., Johnson, S. and Robinson, J. (2005). "Institutions As the Fundamental Cause of Long Run Economic Growth," in Handbook of Economic Growth, Aghion, P. and Durlauf, S. eds., North-Holland, Amsterdam.

[5] Acemoglu, D. and Robinson, J. (2012)"Why Nations Fail: The Origins of Power, Prosperity and Poverty," Crown Business, New York.

[6] Ahlerup, P. and Olsson, O. (2012). "The roots of ethnic diversity," Journal of Economic Growth, Springer, vol. 17(2), pp. 71 – 102, June.

[7] Ang, J. (2013a). "Are Modern Financial Systems Shaped by State Antiquity?" Journal of Banking and Finance 37 (11), pp. 4038 – 4058.

[8] Ang, J. (2013b). "Institutions and the Long-run Impact of Early Development," Journal of Development Economics 108m, pp. 1 – 18.

[9] Ashraf, Q. and Galor, O. (2011). "Dynamics and Stagnation in the Malthusian Epoch" American Economic Review, 101(5), pp. 2003 – 2041.

[10] Ashraf, Q. and Galor, O. (2013). "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development," American Economic Review, 103(1), pp. 1 – 46.

[11] Besley, T. and T. Persson (2009). "The Origins of State Capacity: Property Rights, Taxation and Politics," American Economic Review 99(4), pp. 1218 – 1244.

[12] Besley, T. and T. Persson (2013). "Taxation and development," in "Handbook of Public Economics " ed. A. Auerbach, R. Chetty, M. Feldstein, and E. Saez, vol. 5(2), pp. 51 – 110.

[13] Bockstette, V., Chanda, A. and Putterman, L. (2002). "States and Markets: The Advantage of an Early Start," Journal of Economic Growth, Springer, vol. 7(4), pp. 347 – 69, December.

[14] Carneiro, R.L. (1970). "A theory of the origin of the state," Science 169, pp. 733 – 38.

[15] Carneiro, R.L. (1981). "The chiefdom: precursor of the state," in "The Transition to Statehood in the New World," ed. GD Jones, RR Kautz, Cambridge: Cambridge Univ. Press, pp. 37 – 79.

[16] Chanda, A. and Putterman, L. (2005). "State Effectiveness, Economic Growth, and the Age of States," in Matthew Lange and Dietrich Rueschemeyer, eds., "States

and Development: Historical Antecedents of Stagnation and Advance," Basingstoke, England, Palgrave MacMillan, pp. 69 – 91.

[17] Chanda, A. and Putterman, L. (2007). "Early Starts, Reversals and Catch-up in the Process of Economic Development," Scandinavian Journal of Economics, Wiley Blackwell, vol. 109(2), pp. 387 – 413.

[18] Chanda, A., Cook, A. and Putterman, L. (2014). "Persistence of Fortune: Accounting for Population Movements, There was No Post-Columbian Reversal," American Economic Journal – Macroeconomics (in press, August issue).

[19] Cinyabuguma, M. and Putterman, L. (2011). "Sub-Saharan Growth Surprises: Being Heterogeneous, Inland and Close to the Equator Does not Slow Growth within Africa," Journal of African Economies 20(2), pp. 217 – 262.

[20] Collier, P. (2009). "The Political Economy of State Failure," Oxford Review of Economic Policy, Volume 25, Number 2, 2009, pp.219 – 240.

[21] Comin, D., Easterly, W. and Gong, E. (2010). "Was the Wealth of Nations Determined in 1000 BC?" American Economic Journal: Macroeconomics, American Economic Association, vol. 2(3), pp. 65 – 97, July.

[22] Daniele, V. (2013). "Does the Intelligence of Populations Determine the Wealth of Nations?" Journal of Socio-Economics 46, pp. 27 – 37.

[23] Depetris-Chauvin, E. (2014) "State history and contemporary conflict: Evidence from Sub-Saharan Africa," unpublished working paper, Brown University.

[24] Diamond, J. and Bellwood, P. (2003), "Farmers and their languages: The first expansions," Science, vol. 300, no. 5619, pp. 597 – 603.

[25] Dincecco, M and Katz, G. (2014). "State capacity and long-run economic performance," The Economic Journal doi: 10.1111/ecoj.12161.

[26] Earle, T.K. (1987). "Chiefdoms in archaeological and ethnohistorical perspective," Annual Review of Anthropology 16, pp. 279 – 308.

[27] Ertan, A., Putterman, L. and Fiszbein, M. (2012). "Determinants and economic consequences of colonization: A global analysis," Brown University Department of Economics Working Paper, August, 2012.

[28] Flannery, K.V. (1995). "Prehistoric social evolution". In "Research Frontiers in Anthropology," ed. M Ember, Englewood Cliffs, NJ: Prentice-Hall, pp. 1 – 26.

[29] Gennaioli, N. and Voth, H.J. (2014). "State Capacity and Military Conflict," Working Papers 593, Barcelona Graduate School of Economics.

[30] Galor, O. and Moav, O. (2007). "The Neolithic Revolution and Contemporary Variations in Life Expectancy," Working Papers 2007-14, Brown University, Department of Economics.

[31] Hariri, J.G. (2012). "The Autocratic Legacy of Early Statehood," American Political Science Review, vol. 106(3), pp. 471 – 494.

[32] Hibbs, D. A., and Olsson, O. (2004). "Geography, biogeography, and why some countries are rich and others are poor," Proceedings of the National Academy of Sciences, 101, pp. 3715 – 3720.

[33] Iliev, P. and Putterman, L. (2007). "Social Capability, History and the Economies of Communist and Post-Communist States," Studies in Comparative International Development 42 (1), pp. 36 – 66.

[34] Johnson, A. and Earle, T. (2000). "The evolution of human societies: from foraging group to agrarian state," Stanford University Press, 2nd edition, December.

[35] Jones, E.L. (1981). "The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia," Cambridge: Cambridge University Press.

[36] Jones, G. D., and Kautz, R.R. (1981). "Issues in the Study of New World State Formation," in "The Transition to Statehood in the New World," Grant Jones and Robert Kautz, eds.,Cambridge: Cambridge University Press, Pp. 3 – 34.

[37] Kennedy, P. (1989). "The Rise and Fall of the Great Powers," Vintage Books, New York.

[38] Mankiw et al (1992) "A Contribution to the Empirics of Economic Growth" *Quarterly Journal of Economics* (May), 408 – 437.

[39] Marcus, J. (1992). "Dynamic cycles of Mesoamerican states: political fluctuations in Mesoamerica," National Geographic Research and Explorations. 8:, pp. 392 – 411.

[40] McEvedy, C., and Jones, R. (1978). "Atlas of World Population History," New York, NY:Penguin Books Ltd.

[41] Morris, I. (2010). "Why the West Rules—For Now: The Patterns of History, and What They Reveal About the Future," London: Profile Books.

[42] North, D. (1990) "Institutions, Institutional Change and Economic Performance," Cambridge: Cambridge University Press.

[43] Olson, M. (1993) "Dictatorship, Democracy, and Development" *American Political Science Review* 87(3), 567 – 576.

[44] Olsson, O. and Hibbs, D. (2005). "Biogeography and Long -run Economic Development," European Economic Review, Vol. 49, No. 4, pp. 909 – 938.

[45] Olsson, O. and C. Paik (2013). "A Western Reversal Since the Neolithic: The Long-Run Impact of Early Agriculture," University of Gothenburg, mimeo.

[46] Olsson, O., Borcan, O. and Putterman, L. (2015), "The Neolithic Transition and State Formation", University of Gothenburg, mimeo.

[47] Peregrine, P.N., Ember, C.R. and Ember M. (2007). "Modeling State Origins Using Cross-Cultural Data," Cross-Cultural Research, Vol. 41, No. 1, pp. 75 – 86.

[48] Peregrine, P.N. and Ember, M., eds. (2001). "Encyclopedia of Prehistory," New York: Kluwer Academic/Plenum Publishers.

[49] Peregrine, P. N. (2003). "Atlas of Cultural Evolution," World Cultures: Journal of Comparative and Cross-Cultural Research, 14(1): 1 – 75.

[50] Putterman, L. with C.A Trainor (2006) "Agricultural Transition Year - Country Data Set," http://www.econ.brown.edu/fac/louis_putterman/.

[51] Putterman, L. and Weil, D. N. (2010). "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," The Quarterly Journal of Economics, MIT Press, vol. 125(4), pp. 1627 – 1682, November.

[52] Service, E. R. (1962). "Primitive social organization; an evolutionary perspective," Random House, New York.

[53] Spencer, C.S. (1990). "On the tempo and mode of state formation: Neoevolutionism reconsidered," Journal of Anthropological Archaeology 9, pp. 1 – 30.

[54] Spencer, C.S and Redmond, E.M. (2004). "Primary state formation in Mesoamerica," Annual Review of Anthropology 33, pp. 173 – 99.

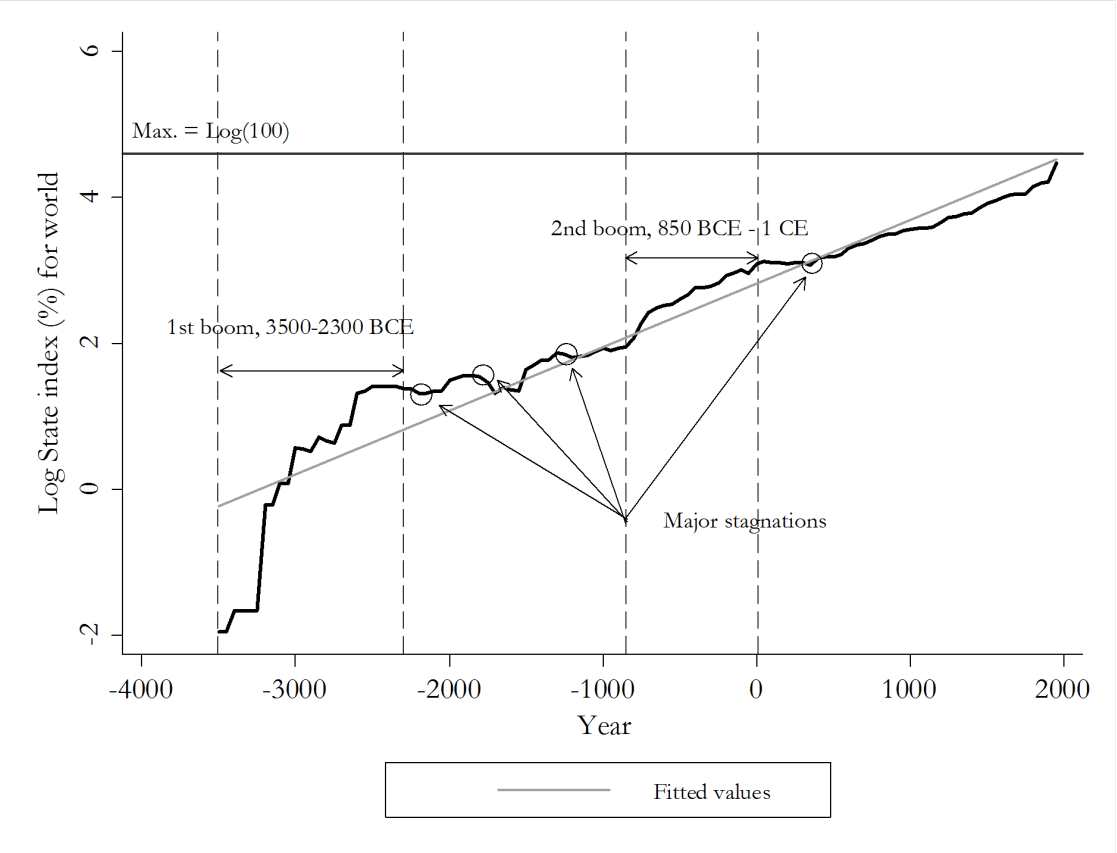
[55] Tilly, C. (1990). "Coercion, Capital and European States," AD 990-1990, Oxford.

[56] Weber, M. (1919). "Politics as a vocation."

[57] Wittfogel, K.A. (1957). "Oriental Despotism; a Comparative Study of Total Power," Yale University Press.

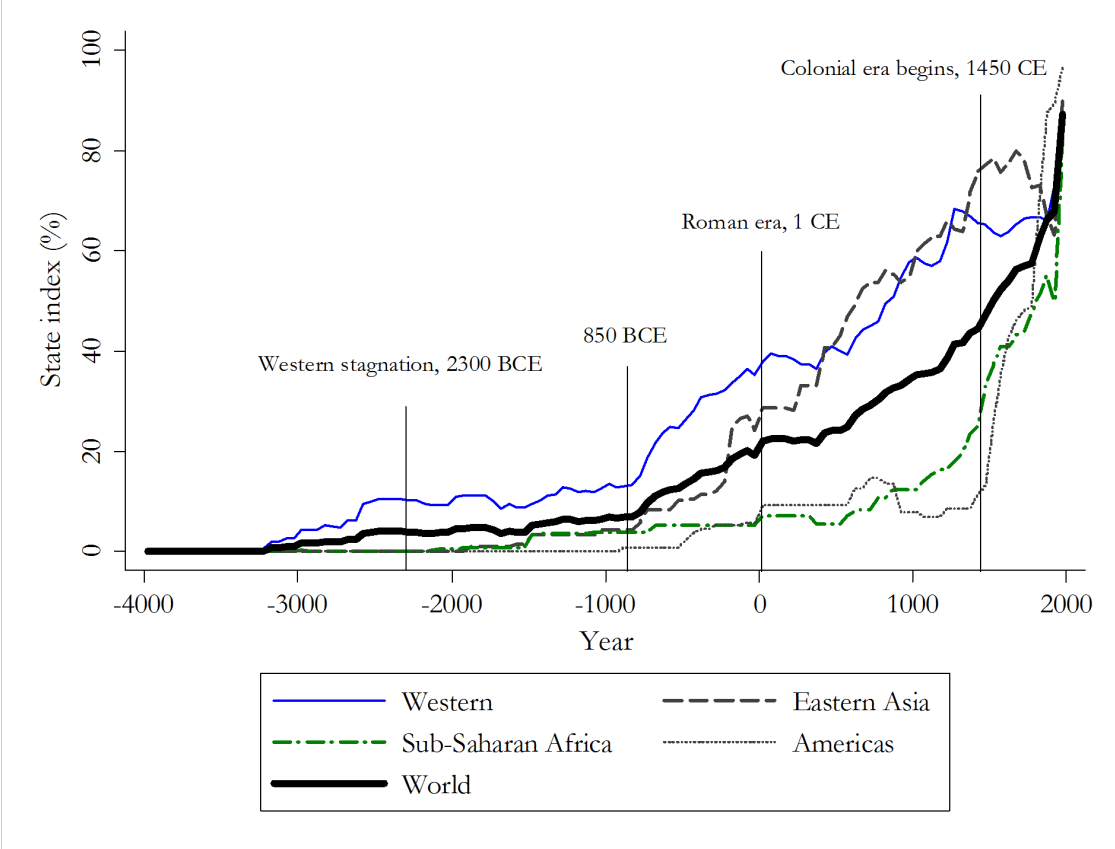
Figures and Tables

Figure 1: Emergence of states in the world 3500 BCE-2000 CE



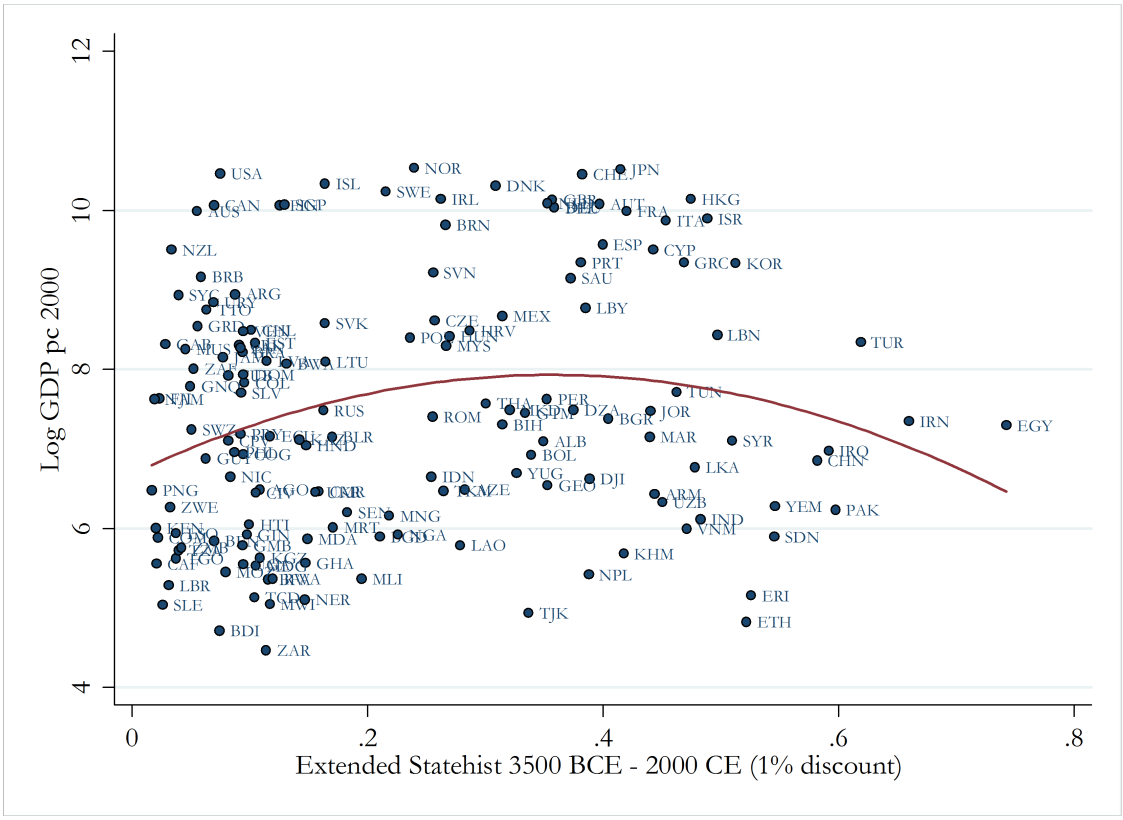
Note: The graph shows the logged value of the aggregate State index for 159 countries identified during 110 50-year intervals between 3500 BCE and 2000 CE. The value 100 is equivalent to all 159 countries in our sample being full states, as defined in the text. On the horizontal axis, negative values imply years BCE whereas positive values show the CE-period. A linear fitted regression line has been included. The State index is calculated as described in text.

Figure 2: Emergence of states in four agricultural core areas and in the world as a whole 3500 BCE- 2000 CE



Note: The figure shows the development of the aggregated State index in the Western agricultural zone (including 62 current countries in Europe, North Africa, the Middle East, as well as Afghanistan, Armenia, Azerbaijan, Georgia, India, Iran, Kazakhstan, Kyrgyzstan, Pakistan, Russia, Tajikistan, Turkmenistan, Uzbekistan), Eastern Asia (20 countries), Americas (including 27 countries in North and South America and in the Caribbean), and Sub-Saharan Africa (47 countries). Oceania (including 3 countries) is omitted. It also shows the aggregate index for the 159 countries in the world as a whole (solid black line). On the horizontal axis, negative values imply years BCE whereas positive values show the CE period. Particular years with trend breaks are marked.

Figure 3: Non-linear relationship between Log GDP per capita in 2000 and Statehist index



Note: The figure shows a fitted quadratic regression line corresponding to the estimates in Table 5, Panel A, column 2, with 154 country observations distinguished by 3-letter country isocodes.

Table 1: Summary Statistics

Panel A State history indicators	**N**	**Mean**	**SD**	**Min**	**Max**
Statehist	159	0.234	0.172	0.017	0.743
Statehist 1 – 1950 CE	159	0.386	0.261	0.012	0.978
Ancestry – Adjusted Statehist in 2000 CE	154	0.252	0.189	0.017	0.811
Ancestry – Adjusted Statehist of 1500 CE	154	0.218	0.167	0.000	0.747
State Age (millenia)	159	1.639	1.430	0.100	5.500
Internally – originated	159	0.490	0.501	0	1
Panel B Outcome Variables					
Average Technology Adoption in 1500 CE	112	0.487	0.317	0.000	1.000
(Log) Population Density in 1500 CE	154	0.905	1.461	-3.817	3.842
Urbanization Rate in 1500 CE	83	7.278	5.134	0.000	28.000
Average Technology Adoption in 2000 CE	130	0.451	0.198	0.174	1.012
(Log) GDP pc in 2000	154	7.488	1.606	4.463	10.531
Panel C Covariates					
Agyears (millenia)	151	4.717	2.442	0.362	10.500
Origtime (millenia)	158	58.917	49.958	0.200	160.000
Absolute centroid latitude	159	26.368	17.704	0.422	67.469
Landlocked	134	0.224	0.418	0.000	1.000
Distance to coast and rivers	149	374.333	457.408	7.952	2385.58
Mean Elevation	149	637.715	551.281	9.167	3185.920
Land Suitability	145	0.378	0.248	0.000	0.960
Percentage Arable Land	156	15.852	14.001	0.040	62.100
Temperature	158	18.226	8.350	-7.929	28.639
Precipitation	158	92.959	61.700	2.911	259.952
Malaria (percentage population at risk)	151	0.316	0.426	0.000	1.000

Note: The table summarizes all variables used in the analysis, as follows: 1) Panel A describes the State history variables created by us. Note that *Ancestry – Adjusted statehist of 1500* is the average accumulated state history to 1500 CE of the year 1500 ancestors (by territory of residence at that time) of the year 2000 population of each country; 2) Panel B outlines some historical and economic variables which are used as dependent variables in the regression analysis. The data for historical population density is based on population data from McEvedy and Jones(1978) and land data from World Bank World Development Indicators. The data for urbanization rate in 1 CE is taken from Comin, Easterly and Gong (2010) and is based on Peregrine (2003). The data for urbanization rate in 1500 CE is that reported by Acemoglu, Johnson and Robinson (2005). The Average Technology Adoption indices in 1 CE, 1500 CE and 2000 CE are constructed by Comin, Easterly and Gong (2010). Per capita GDP is expressed in current US dollars, as provided by the World Bank; 3) Panel C details the covariates included in the regressions. Agyears was assembled by Putterman with Trainor (2006) and it records the number of millennia elapsed in 2000 C.E. since the Neolithic transition took place. Origtime was coded by Ahlerup and Olsson (2012) and it represents the time since initial uninterrupted settlement by modern humans (before 2000 CE). The geographic and climatic controls are retrieved from various sources. The variables' construction is detailed in Appendix B.

Table 2: State history and average technology adoption in 1500 CE

	Technology Adoption in 1500 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist in 1500 CE	1.227***	2.841***		1.695***	1.727***	1.068***	1.782***
	(0.157)	(0.350)		(0.444)	(0.338)	(0.229)	(0.416)
Statehist in 1500 CE squared		-3.359***		-2.587***	-1.855***	-0.743**	-0.943***
		(0.738)		(0.887)	(0.624)	(0.346)	(0.319)
Agyears in 1500 CE			0.104***	0.073***	0.038***	0.004	0.012
			(0.008)	(0.014)	(0.013)	(0.010)	(0.011)
Origtime in 1500 CE					0.001	-0.001**	0.000
					(0.001)	(0.001)	(0.001)
Origtime in 1500 CE squared							-0.000
							(0.000)
State Age in 1500 CE							-0.092*
							(0.049)
Observations	112	112	110	110	107	107	107
R-squared	0.389	0.521	0.532	0.616	0.809	0.902	0.911
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The dependent variable is the technology adoption index in 1500 CE and the main independent variables are the *extended statehist index* between 3500 BCE and 1500 CE, linear and squared. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table 3: State history, Log Population Density and Urbanization in 1500 CE

Panel A	Log Population Density in 1500 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist in 1500 CE	3.883***	9.559***		6.184***	7.473***	6.129***	11.077***
	(0.670)	(1.666)		(2.119)	(1.802)	(1.709)	(3.426)
Statehist in 1500 CE squared		-12.324***		-9.893***	-7.339**	-4.894**	-7.326**
		(3.098)		(3.498)	(3.169)	(2.253)	(2.905)
Agyears in 1500 CE			0.315***	0.211***	0.157**	0.131*	0.217***
			(0.042)	(0.067)	(0.065)	(0.067)	(0.068)
Origtime in 1500 CE					0.005**	-0.003	-0.020
					(0.003)	(0.004)	(0.014)
Origtime in 1500 CE squared							0.000
							(0.000)
State Age in 1500 CE							-0.509**
							(0.214)
Observations	154	154	147	147	128	128	128
R-squared	0.184	0.254	0.269	0.314	0.709	0.767	0.786
Panel B	Urbanization in 1500 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist in 1500 CE	16.678***	29.429***		35.364***	48.134***	41.542***	69.670***
	(2.384)	(7.323)		(8.662)	(12.193)	(12.707)	(22.835)
Statehist in 1500 CE squared		-25.531**		-29.862**	-43.924***	-35.621**	-49.359***
		(12.514)		(12.662)	(15.010)	(14.076)	(17.606)
Agyears in 1500 CE			0.761***	-0.382	-0.206	-0.323	0.152
			(0.177)	(0.244)	(0.325)	(0.417)	(0.481)
Origtime in 1500 CE					-0.076**	-0.082*	-0.216
					(0.037)	(0.042)	(0.146)
Origtime in 1500 CE squared							0.002
							(0.002)
State Age in 1500 CE							-2.831
							(1.750)
Observations	83	83	83	83	76	76	76
R-squared	0.278	0.311	0.111	0.324	0.459	0.498	0.532
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: In panel A, the dep. var. is log population density in 1500 CE and the main independent variables are the *extended statehist index* between 3500 BCE and 1500 CE, linear and squared. In panel B, the dep. var. is the urbanization rate in 1500 CE and the main independent variables are the *extended statehist index* between 3500 BCE and 1500 CE, linear and squared. The data for historical population density is based on population data from McEvedy and Jones(1978) and land data from *World Bank World Development Indicators*. The data for 1500 CE urbanization rate is reported by Acemoglu, Johnson and Robinson (2005), defined as the percentage of a country's total population residing in urban areas (each with a city population size of at least 5,000). The controls include: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses.***p<0.01, **p<0.05, *p<0.1

Table 4: State history and average technology adoption 2000 CE

Panel A	Technology Adoption in 2000 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	0.086	0.842***		0.664**	0.302	0.461*	0.604**
	(0.095)	(0.318)		(0.332)	(0.239)	(0.244)	(0.272)
Statehist squared		-1.285***		-1.192**	-0.405	-0.554*	-0.531*
		(0.452)		(0.459)	(0.347)	(0.322)	(0.319)
Agyears			0.011	0.011	-0.007	-0.004	0.002
			(0.007)	(0.010)	(0.008)	(0.012)	(0.012)
Origtime					0.000	0.001**	0.001
					(0.000)	(0.001)	(0.002)
Origtime squared							0.000
							(0.000)
State Age							-0.030
							(0.021)
Observations	130	130	129	129	125	125	125
R-squared	0.006	0.044	0.016	0.050	0.643	0.683	0.688

Panel B	Technology Adoption in 2000 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry – Adjusted	0.233**	1.332***		1.329***	0.794***	0.672***	0.787***
Statehist in 1500 CE	(0.092)	(0.275)		(0.309)	(0.215)	(0.219)	(0.223)
Ancestry – Adjusted		-2.088***		-2.085***	-1.125***	-0.938***	-0.887***
Statehist in 1500 CE sqr.		(0.513)		(0.514)	(0.287)	(0.317)	(0.303)
Agyears			0.011	-0.000	-0.010	-0.005	0.001
			(0.007)	(0.011)	(0.009)	(0.011)	(0.012)
Origtime					0.001*	0.001**	0.001
					(0.000)	(0.001)	(0.002)
Origtime squared							0.000
							(0.000)
State Age							-0.029
							(0.018)
Observations	130	130	129	129	125	125	125
R-squared	0.037	0.151	0.016	0.150	0.675	0.698	0.704
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: In both panels the dependent variable is the technology adoption index in 2000 CE. In panel A the main independent variables are the *extended statehist index* between 3500 BCE and 2000 CE, linear and squared. In panel B, the main independent variables are the *extended statehist index* between 3500 BCE and 1500 CE, ancestry-adjusted, linear and squared. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table 5: Statehist vs. Statehist 1-1950 CE and (Log) GDP pc 2000. Nonlinear relationship

Panel A	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	1.326*	7.010***		7.337***	3.869**	4.530**	6.790***
	(0.723)	(2.291)		(2.658)	(1.921)	(2.057)	(2.496)
Statehist squared		-9.842***		-9.832***	-4.718	-4.970*	-4.657*
		(3.529)		(3.549)	(2.854)	(2.793)	(2.776)
Agyears			0.105**	0.004	-0.071	-0.087	0.010
			(0.048)	(0.079)	(0.063)	(0.079)	(0.081)
Origtime					0.002	0.008**	0.010
					(0.003)	(0.004)	(0.013)
Origtime							-0.000
squared							(0.000)
State Age							-0.460**
							(0.183)
Observations	154	154	147	147	125	125	125
R-squared	0.020	0.052	0.026	0.064	0.702	0.719	0.734
Panel B	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist	1.277**	1.940		2.200	0.066	0.251	1.267
1-1950 CE	(0.531)	(2.049)		(2.278)	(1.441)	(1.597)	(1.667)
Statehist 1-1950		-0.783		-0.748	0.942	0.962	0.453
CE squared		(2.518)		(2.625)	(1.608)	(1.811)	(1.776)
Agyears			0.105**	-0.011	-0.069	-0.080	0.012
			(0.048)	(0.068)	(0.055)	(0.072)	(0.081)
Origtime					0.001	0.007*	0.011
					(0.003)	(0.004)	(0.013)
Origtime							-0.000
squared							(0.000)
State Age							-0.267**
							(0.127)
Observations	154	154	147	147	125	125	125
R-squared	0.043	0.044	0.026	0.058	0.704	0.722	0.730
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The dependent variable is Log per capita GDP in 2000. In panel A the main independent variables are extended *Statehist index* linear and squared. In panel B the main independent variables are the *Statehist index 1-1950 CE*, linear and squared. The list of controls includes: absolute latitude, an indicator of whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table 6: Ancestry – Adjusted Statehist and (Log) GDP pc 2000. Nonlinear relationship

Panel A	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry – Adjusted	2.778***	12.213***		13.110***	6.068***	5.347***	6.041***
Statehist of 1500	(0.794)	(2.263)		(2.100)	(1.574)	(1.647)	(1.761)
Ancestry – Adjusted		-18.218***		-18.636***	-8.984***	-7.519***	-6.794***
Statehist of 1500 sqr.		(4.326)		(4.076)	(2.176)	(2.324)	(2.317)
Agyears			0.105**	-0.025	-0.056	-0.075	-0.027
			(0.048)	(0.059)	(0.056)	(0.078)	(0.082)
Origtime					0.003	0.006*	0.008
					(0.003)	(0.004)	(0.013)
Origtime sqr.							-0.000
							(0.000)
State Age							-0.233
							(0.146)
Observations	149	149	147	144	125	125	125
R-squared	0.083	0.209	0.026	0.243	0.722	0.727	0.733

Panel B	Log GDP pc 2000						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry – Adjusted	1.389**	7.074***		6.661***	3.514**	4.123**	6.268***
Statehist in 2000	(0.670)	(2.113)		(2.426)	(1.727)	(1.849)	(2.286)
Ancestry – Adjusted		-9.085***		-8.378***	-4.034*	-4.250*	-4.033*
Statehist in 2000 sqr.		(3.021)		(3.000)	(2.364)	(2.308)	(2.289)
Agyears			0.105**	0.021	-0.069	-0.085	0.010
			(0.048)	(0.080)	(0.063)	(0.079)	(0.081)
Origtime					0.002	0.008**	0.010
					(0.003)	(0.004)	(0.013)
Origtime sqr.							-0.000
							(0.000)
State Age							-0.463**
							(0.186)
Observations	149	149	147	144	125	125	125
R-squared	0.027	0.066	0.026	0.071	0.702	0.719	0.734
Controls	No	No	No	No	Yes	Yes	Yes
Continent FE	No	No	No	No	No	Yes	Yes

Note: The dependent variable is Log per capita GDP in 2000. In panel A the main independent variables are the *ancestry-adjusted extended Statehist index* between 3500 BCE and 1500 CE, linear and squared. In panel B the main independent variables are the *composite ancestry-adjusted Statehist index* (where the discounted ancestry-adjusted scores between 3500 BCE and 1500 CE are added to the raw discounted scores between 1500 and 2000 CE, and the final score is normalized by the sum of discounted full scores between 3500 BCE and 2000 CE), linear and squared. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Appendix

Appendix A - The Model

In this Appendix, we outline the full version of the extended Malthusian growth model of Ashraf and Galor (2011). The basic setting is a geographically well defined region where a population has made the transition to an agricultural, sedentary society. The specific aim of the model is to propose a mechanism for how states interact with economic development during the agricultural era. The key novel features of the model are that we introduce the endogenous rise of a state which taxes individuals but also provides public goods. Both fiscal capacity as well as the centralization of power increases with accumulated state experience. The centralization of power initially has a positive impact on the effective provision of public goods and economic growth but might eventually transform into having a negative impact when the constraints and checks against governments become too weak.

A1 Individuals

Let us imagine an overlapping generations framework with a representative (hermaphroditic) individual who lives for two periods, *childhood* and *adulthood.* All key choices are made in adulthood. The individual has a utility function given by

$$u_t = c_t^{\gamma} \cdot n_t^{1-\gamma} \tag{1}$$

where c_t is the individual's level of consumption at time t with an associated preference parameter γ and n_t is the (continuous) number of children of each adult person. One unit of time might be thought of as a generation.

The adult individual farmer earns an income y_t which can be used for either child-rearing, consumption, or paying a tax. The budget constraint is therefore

$$c_t + \mu n_t \leq y_t - s_t\theta(S_t) \tag{2}$$

where $\mu > 0$ is the cost of rearing one child and $s_t\theta(S_t) \geq 0$ is a lump-sum tax where s_t is a binary indicator $s_t \in \{0, 1\}$ describing whether there is a state or not

during the particular period t and where the tax level $\theta(S_t)$ is a function of the historical accumulated experience of a state $S_t \geq 0$. We will define taxation and the involvement of a state in the economy further below.

By maximizing (1) with respect to the budget constraint in (2), it is straightforward to show that the utility-maximizing quantity of children is

$$n_t^* = \frac{(1-\gamma)\left(y_t - s_t\theta(S_t)\right)}{\mu}. \tag{3}$$

Since the adult population in the next generation is given by $L_{t+1} = n_t^* L_t$, we can express the law of motion for labor as:

$$L_{t+1} = n_t^* L_t = \frac{(1-\gamma)\left(y_t - s_t\theta(S_t)\right) L_t}{\mu} \tag{4}$$

A2 Production and population density

The aggregate production function during the Malthusian era is given by

$$Y_t = A_t X^{1-\alpha} L_t^{\alpha}.$$

In this function, Y_t is output, X is the amount of land and A_t is total factor productivity, L_t is the size of the labor force (equal to the number of live adults in period t), and $\alpha \in (0,1)$ is the output elasticity of labor. There are no property rights to land so holders of land receive no compensation.

Workers are paid their average product, which will also be equal to the gross income per adult person:

$$y_t = \frac{Y_t}{L_t} = A_t \left(\frac{X}{L_t}\right)^{1-\alpha} = A_t P_t^{\alpha-1} \tag{5}$$

In this expression, $P_t = L_t/X$ is population density. Income per capita thus increases with productivity and decreases with population density at time t. Combining this expression with (4) and dividing by X means that we can express the evolution of population density as

$$P_{t+1} = \frac{(1-\gamma)\left(A_t P_t^{\alpha-1} - s_t\theta(S_t)\right)}{\mu} \cdot P_t$$

P_{t+1} is thus a concave function of P_t and a linear, positive function of A_t.

A3 Public goods, productivity, and state history

The equations specifying the level of public goods $G_t = s_t\theta(S_t) \cdot Z(S_t)$, total factor productivity $A(N, S_t)$, and levels of state history S_τ are displayed and explained in sections 3.2-3.4 in the main paper.

A4 Dynamics

Combining the components above with the equations in the paper describing public goods, productivity and state history, population density in the Malthusian era can be written as

$$P_{t+1} = \frac{(1-\gamma)\left(P_t^{\alpha-1}(N + s_t\theta(S_t)\left(Z(S_t)-1\right)\right)}{\mu} \cdot P_t \tag{6}$$

In the pre-state economy when $s_t = 0$, the size of the population will tend to converge towards a steady state when $P_{t+1} = P_t = \bar{P}$ and $\bar{n} = 1$. Combining (4) and (5), we can deduce that the equilibrium level of population density is

$$\bar{P} = \frac{\bar{L}}{X} = \left(\frac{(1-\gamma)\,N}{\mu}\right)^{\frac{1}{1-\alpha}}. \tag{7}$$

Population density thus increases with productivity (i.e. the natural quality of the environment) N and decreases with the cost of raising children μ. If this equilibrium level $\bar{P}$ is lower than the critical state viability level $\tilde{P}$, there will be no state. However, we assume that the normal scenario for a region that made the transition to agriculture is that $\tilde{P} < \bar{P}$. Hence, the region will naturally converge towards a level of population where a state is eventually formed. The regions with the highest level of N, i.e. the highest land quality for agriculture, will reach the critical level first and hence have the oldest states.

If a state exists so that $s_t = 1$, the equilibrium population density will converge towards a level

$$\bar{P}_t^s = \left(\frac{(1-\gamma)\left(N + \theta(S_t)\cdot Z(S_t)\right)}{\mu + (1-\gamma)\,\theta(S_t)}\right)^{\frac{1}{1-\alpha}}. \tag{8}$$

This expression has taxes $\theta(S_t)$ both in the numerator and in the denominator.

A key feature is further that state history S_t will at first have a net positive effect on equilibrium population due to the beneficial impact of early centralization $Z(S_t)$ but will eventually have a net negative impact when a lower effective public goods provision leads to lower income levels, leaving less resources for raising children.

Note also that unlike (7), the equilibrium level of population density during states will not be a steady, constant equilibrium since the stock of state history keeps changing with time.

Inserting (8) level into (5), we find that the equilibrium disposable income per person in the Malthusian economy with states is:

$$\bar{y}_t - \theta(S_t) = \frac{\mu}{(1-\gamma)} + \theta(S_t) - \theta(S_t) = \frac{\mu}{(1-\gamma)}$$

The expression shows that that income per capita $\bar{y}_t$ is unaffected by total productivity A_t but increases with state presence and with state history through fiscal capacity $\theta(S_t)$. It also shows that net income per capita is not a function of state history S_t since the increase in income per capita due to a lower fertility is exactly offset by the decrease in disposable income due to taxation.

Equilibrium levels of consumption remains stable throughout and can be shown to be[1]

$$\bar{c} = \frac{\gamma\mu}{(1-\gamma)}.$$

A5 Industrial era

The Malthusian era comes to an end at the time of the industrial revolution t_M. The nature of the economy is then fundamentally changed from being dependent on land to being driven by capital accumulation. Aggregate output is now a function of productivity or technology A_t, labor L_t, and physical capital K_t such that $Y_t = F(A_t, L_t, K_t) = K_t^{\alpha}(A_t L_t)^{1-\alpha}$. Output per worker (or per capita) is $y_t = Y_t/L_t = A_t k_t^{\alpha}$ where $k_t = K_t/A_t L_t$ is capital per unit of effective worker.

As in Mankiw et al (1992) and many other studies, we assume that dynamics of the capital stock are given by

$$K_{t+1} - K_t = sY_t - \delta K_t.$$

The two other factors of production A_t and L_t have exogenous growth rates $g, n > 0$:

$$A_{t+1} = A_t\,(1+g) = A_{t_M}\,(1+g)^{t+1-t_M}\,; \quad L_{t+1} = L_t\,(1+n)\,.$$

Unlike during the Malthusian era, n is now independent of y_t.

[1] Also indirect utility is constant with and without a state and is $v = \gamma\left(\frac{\mu}{(1-\gamma)}\right)^{\gamma}$.

Another standard assumption is that technological progress during the industrial era grows in each period at a percentage rate $g > 0$ which we assume to be exogenously given and not a function of state-provided public goods.[2] In line with the expression for A_t in the main paper, let us define the level of productivity at the end of the Malthusian era $t_M > 0$ to be $A_{t_M} = N + \theta(S_{t_M}) \cdot Z(S_{t_M})$. The level of productivity in some industrial era time period $\tau > t_M$ is then

$$A_\tau = A_{t_M} \cdot (1+g)^{\tau - t_M} = (N + z\theta(S_{t_M}) \cdot Z(S_{t_M})) \cdot (1+g)^{\tau - t_M}$$

The dynamics of the capital stock per unit of effective worker can be rewritten as

$$\begin{aligned} k_{t+1} &= \frac{K_{t+1}}{A_{t+1}L_{t+1}} = \frac{sY_t + (1-\delta)K_t}{A_{t+1}L_{t+1}} = \frac{sY_t + (1-\delta)K_t}{A_t(1+g)L_t(1+n)} = \\ &= \frac{sk_t^\alpha + (1-\delta)k_t}{(1+g)(1+n)} \end{aligned}$$

In a steady state, it will be the case that $k_{t+1} = k_t = k^*$. Solving for the equilibrium level k^* yields

$$k^* = \left(\frac{s}{g + n + \delta + ng} \right)^{\frac{1}{1-\alpha}}$$

The equilibrium level of output per capita at some time τ is thus:

$$\begin{aligned} \ln \bar{y}_\tau &= \ln A_\tau + \alpha \ln k^* = \\ &= \ln(N + \theta(S_{t_M}) \cdot Z(S_{t_M})) + (t_M - \tau)\ln(1+g) + \frac{\alpha}{1-\alpha} \ln \left(\frac{s}{g + n + \delta + ng} \right). \end{aligned} \tag{9}$$

[2] In a globalized industrial economy, important innovations tend to spread geographically through technological diffusion from the technological frontier.

Appendix B - Variables Description and Data Sources

B1 State history variables

Statehist. The extended statehist is the normalized aggregate index of state history. This index is defined as the sum of all 50-year period state history scores, adjusted by a discount factor, divided by the maximum value of a discounted index, corresponding to a state with a score of 50 in every half century between 3500 B.C.E. and 2000 C.E. The index can be calculated using various discount rates to put more weight on recent history than on the more distant past. Throughout the paper we use the 1% discount rate.

Statehist 1-1950 C.E. This is the statehist computed according to the initial version of the index in Bockstette et al. (2002), considering only the period 1 – 1950 C.E. This is also a normalized index (with respect to a virtual state that would have full scores for every half century between 1 and 1950 C.E.). In this paper we use a discount factor of 1% for this index.

Ancestry – Adjusted Statehist of 1500 C.E. This is the average accumulated state history to 1500 C.E. of the year 1500 ancestors (by territory of residence at that time) of the year 2000 population of each country. It is obtained by adjusting the extended statehist index at 1500 C.E. by the migration matrix of Putterman and Weil (2010), as follows: for each country i in the sample, we use the matrix to identify the share of the current population that has ancestry that can be traced to the territory of country j; for each country j we multiply its statehist score accumulated at 1500 C.E. (discounted by 1% and normalized with respect to a state with full scores from 3500 B.C.E. until 1500 C.E.) by a weight which is the share of the current population in country i identified to have roots in country j; the sum of all weighted statehist scores thus obtained across all j is the ancestry-adjusted statehist index at 1500 C.E. Using this adjustment, we account for the state experience prior to 1500 C.E. of other territories, brought by post-1500 migrants into a land area defined by modern country borders, in addition to the state history of the country's own territory.

Ancestry – Adjusted Statehist in 2000 C.E. This is the extended statehist index at 2000 C.E., adjusted for post-1500 migrations, as follows: we compute the statehist index between 1500 and 2000 C.E. (discounted by 1% and normalized with respect to the same period); we add this score to the ancestry –‘ adjusted statehist of 1500 C.E. defined as above.

State Age. This variable, measured in millennia, represents the total amount of time elapsed from the first date (exact or approximate) when state experience is assigned a positive scores (the first date when the component S1 pertaining to the existence of a rule above tribal level is positive) until 2000 C.E. State age does not discount the periods of state collapse (when scores revert to 0) incurred after the state emergence date.

B2 Historical controls

Agyears. This variable assembled by Putterman with Trainor (2006) records for each present-day country in a sample of 170 countries the number of millennia elapsed in 2000 C.E. since the Neolithic transition of populations that lived on the territory of that country. The year of transition is assigned by cross-referencing expert opinions about the time when the population is a particular region covered more than half of their calorie intake from agriculture.

Origtime. This variable coded by Ahlerup and Olsson (2012) represents the time since initial uninterrupted settlement of anatomically modern humans (before 2000 C.E.) on the territories that now belong to modern-day countries. The variable was coded for 191 countries and the coding was based on Oppenheimer (2003) and Bradshaw Foundation (2007), as well as Encyclopedia Britannica (2007) for the island cases. Since the original settlements follow the paths of the migration routes out of Africa, the variable is correlated with the migration distance, and can therefore also be employed as a proxy for the latter.[3]

The Matrix of Migration since 1500 C.E. This matrix was developed by Putterman and Weil (2010) to describe the composition of the populations of modern-day countries in terms of ancestry at 1500 C.E., before the migration flows of the colonial era. The matrix contains 165 rows (each row corresponding to a present-day country) and 165 columns (representing the same countries), where every cell records the percentage of current population in country on row i that traces its ancestry to the population in the source country on column j, such that the sum of all cells on each row is 1. In their paper, Putterman and Weil (2010) obtained ancestry-adjusted measures of Statehist (1-1950 C.E.) and Agyears, by multiplying each row to each one of the vectors containing the values of their variables of interest at 1500 C.E. for each country in their sample (which amounted to a sum over the values of the

[3]In our sample the correlation coefficient between the time since original human settlement and the migration distance of Ashraf and Galor (2013) is -0.51, which indicates that the shorter the migration distance to a particular territory, the earlier the first human settlement.

variables of interest of each source country by the corresponding share of the population with ancestry in other countries). We follow the same procedure in order to obtain the extended ancestry-adjusted statehist.

B3 Outcomes variables

GDP per capita in 2000. Data in current US dollars, as reported by the World Bank.

Population density in 1 C.E. and 1500 C.E.. This variable is measured in number of individuals per square kilometer. The variable is retrieved from Ashraf and Galor (2013), who employ the population size data from McEvedy and Jones (1978), and the land area from World Development Indicators. Since the territorial unit employed in McEvedy and Jones (1978) is based on 1975 country borders, in some cases, the same value of the population density is assigned to contiguous present-day countries (that may have been part of the same constituency in 1975, such as Yugoslavia).

Urbanization in 1 C.E. This measure of urbanization was computed by Peregrine (2003) in the Atlas of Cultural Evolution (ACE). The variable takes three values: 1 if the largest settlement on the territory defined by the borders of a given modern-day country was under 100 persons, 2 if the largest settlement was between 100 and 399 persons and 3 for settlements larger than 400 persons. We retrieved this variable from Comin, Easterly and Gong (2010), where it also used previously.

Urbanization in 1500 C.E. The urbanization rate for 1500 A.D. comes from Acemoglu, Johnson and Robinson (2005) and it is calculated as the percentage of a country's urban area population (for cities with at least 5,000 inhabitants).

Technology Adoption in 1 C.E. These variables are three indices created by Comin, Easterly and Gong (2010), henceforth CEG. The index in 1 C.E. is based on data from Peregrine's (2003) "ACE" in which various cultural traits of 289 prehistoric cultures are evaluated: writing, agriculture, transportation, urbanization. CEG used this and additional data to code country - level data on technology adoption in five sectors: agriculture, transportation, communications, writing, and military. The authors structured the information in "ACE" into indicators that denoted the presence or absence of a technology within a certain sector and territory, which they then averaged over to create the sector specific technology adoption index between 0 and 1 (e.g. where "ACE" codes "technological specialization" by 1

for none, 2 for pottery and 3 for metalworks, CEG marked pottery and metalwork as the two potential technologies within the "industry sector" at the time, which they coded using a binary convention 1 – if technology is present and 0 if not; the average over all these dummies within every sector is the value of the technology adoption index for that sector; this average for the industry sector in this case would be 0 if neither technology was present, 0.5 if only one was present and 1 if both were present). Then, the overall adoption level, the variable that we use in this paper, for each country, was calculated as the average of the adoption levels across sectors.

Technology Adoption in 1500 C.E. For the average technology adoption measure in 1500 C.E., CEG (2010) used many different sources to summarize information on 20 technologies across 4 sectors excluding agriculture (for instance, for "Industry", the two possible technologies are "presence of iron" and "presence of steel"). For the latter they used a proxy based not on technology presence, but rather on which type of agriculture was the primary source on a particular territory – e.g. pastoralism, hand or plough cultivation, or none). As with overall technology adoption in 1 C.E., the overall measure in 1500 C.E. is obtained by averaging over the scores for each sector.

Technology in 2000 C.E. The technology measure in 2000 C.E. is constructed in CEG(2010), based on Comin, Hobijn and Rovito (2008) and it captures the gap in the intensity of technology adoption for every country with respect to the US (in terms of years of usage of each technology relative to the number of years since the invention of that technology) for ten technologies: electricity, internet, PC's, cell phones, telephones, cargo and passenger aviation, trucks, cars and tractors, in per capita terms. The average across the technologies' scores is subtracted from 1 (the level of US, by construction) to obtain the country-level technology adoption gap measure. This measure is different from the measures for 1 and 1500 C.E., since it also measures adoption along the intensive margin.

B4 Geographical variables

Absolute latitude. This is the absolute value of the country's centroid latitude. The variable was retrieved from the Portland Physical Geography dataset.

Distance to coast and river. This variable represents the mean distance to the nearest coastline or sea-navigable river, measured in kilometers. The variable was retrieved from the Portland Physical Geography dataset.

Mean elevation. The mean elevation above sea level is measured in meters. The variable was retrieved from the Portland Physical Geography dataset. The original source is NOAA's National Geophysical Data Center.

Land suitability. This is a measure of land suitability for agriculture, computed at country level by Michalopoulos (2012), based on grid-cell data reported by Ramankutty et al. (2002). For details on the construction of the original index, the reader is referred to Ramankutty et al (2002). The index includes information on ecological indicators of climate and soil suitability for agriculture (such as drowing degree days, evapotranspiration, soil carbon density and soil pH).

Percentage arable land. This measures the percentage of a modern-day country's area that is arable. The source is World Bank's World Development Indicators.

Temperature. This is a mean across the average monthly temperature over time (1961-1990) in 1-degree resolution grids within a country. This variable was retrieved from Ashraf and Galor (2013), whose source is the G-ECON project (Nordhaus 2006).

Precipitation. This is a mean across the average monthly precipitation over time (1961-1990) in 1-degree resolution grids within a country. This variable was retrieved from Ashraf and Galor (2013), whose source is the G-ECON project (Nordhaus 2006).

Malaria (percentage population at risk). This variable represents the level of risk of contracting malaria (measured by the percentage population in 1994 in areas of high risk of contracting malaria, times the share of cases in the country involving fatal species of P. Falciparum). The original data was constructed by Gallup and Sachs (2001).

Landlocked. This is a dummy variable equal to 1 if the country is landlocked.

Appendix C - Additional *Statehist* Coding Information and Illustrations

To code all components of the index in a manner that is consistent across periods, we relied mainly on information in the Encyclopedia Britannica Online. We surveyed the main articles on the history of the modern-day country (e.g. "History of Azerbaijan"), but also articles connected to events in its history (e.g. "Azerbaijan-

historical region", "Ancient Iran: The Sasanian period"). There were a number of instances where information in Britannica was sparse, in which cases we surveyed alternative sources, such as books or journal articles treating individual cases. For instance, in the case of Afghanistan, in addition to the Encyclopedia Britannica entries, we consulted two books: "Ancient Cities of the Indus Valley Civilization," by Jonathan Kenoyer and "Bactria: An Ancient Oasis Civilization from the Sands of Afghanistan" by Giancarlo Ligabue and Sandro Salvatori. The complete description of coding decisions and sources for all country-period observations can be found in the online Data Coding Appendix accompanying the paper.

Table C1 illustrates an example of coding based on information from Encyclopedia Britannica, covering the period 450 BCE - 1 CE for the territory of modern-day Bosnia and Herzegovina.

Table C1: Coding example - the case of Bosnia and Herzegovina, 400 BCE – 1 CE

Year BCE	Government is above tribe level? (z^1_{it})	Government domestic? (z^2_{it})	What percentage of the territory is ruled? (z^3_{it})	Weight* ($w_{it\theta}$)	s_{it}
400-201	0.75	1	0.3	1	11.25
200-151	0.75	1	0.3	0.9	12.63
(cutoff at 155)	1	0.5	1	0.1	
150-1	1	0.5	1	0.72	25

Note: $w_{it\theta}$ = number of years between period ends and cutoff, or between two cutoffs, divided by 50.

CODING INFORMATION

400-200: (0.75, 1, .3). From the 4th century BCE, along with the coming of Celtic tribes in the area, the Illyrian tribes became gradually more politically cohesive. Sources recall the existence of early indigenous petty kingdoms in Illyria on the territory of present-day Albania only. We therefore mark the occasional Illyrian tribe alliances by z^1_{it}=.75.

200-151: (0.75, 1, .3) until 155 and (1, .5, 1) from 155 onwards. Delminium (on modern-day Bosnian territory) was taken by the Romans in 155 BCE, hence $z^2_{it} = .5$. Most of the area of Bosnia was integrated in the Roman province Dalmatia, hence z^3_{it}). The score is [0.9(0.75, 1, .3) + 0.1(1, .5, 1)]*50.

150-1: (1, .5, 1). Bosnia was under Roman occupation.

The extended Statehist includes some coding exceptions not present in the 1-1950 CE Statehist:

1) The z_{it}^1 component (existence of a supra-tribal rule) may also take the value 0.5 to indicate radical uncertainty with respect to the existence of a supra-tribal rule on the territory of country i in period t. An example is the case of Somalia, which receives a score z_{it}^1 of 0.5 between 1500 BCE and 1 CE, when its territory is believed to have been part of the Kingdom of Punt.

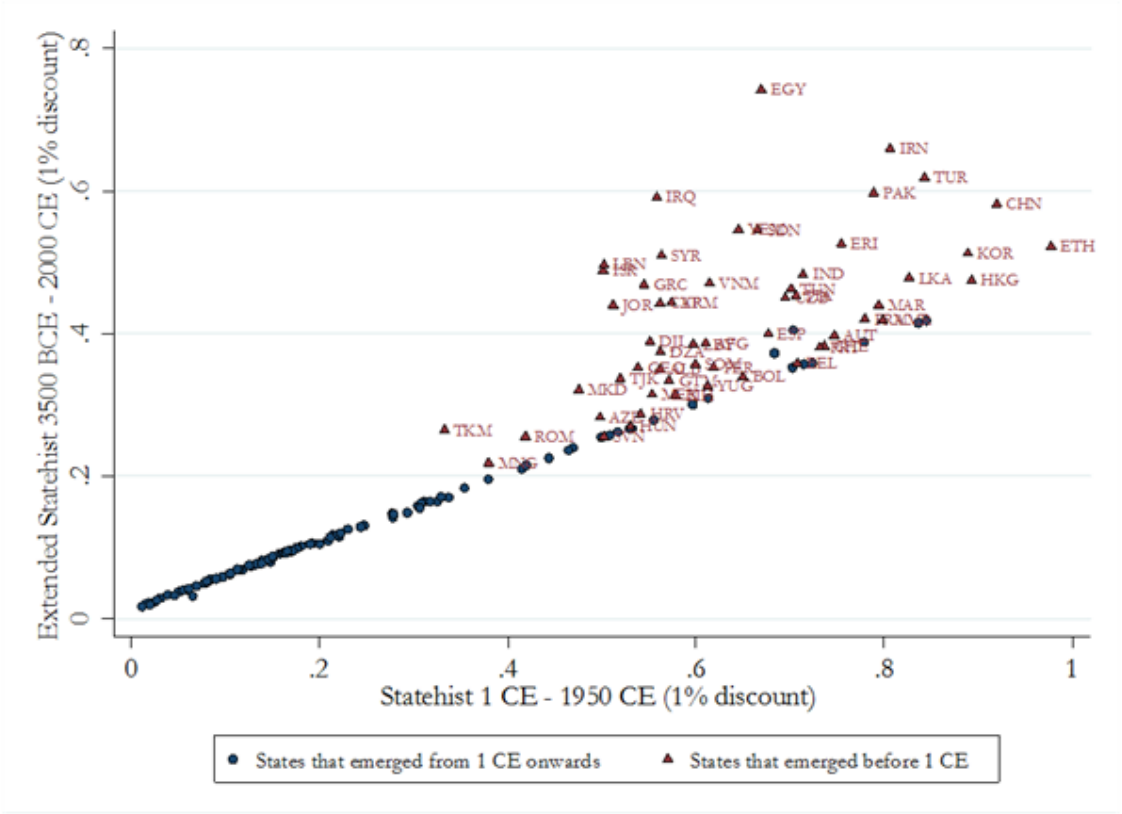
2) The z_{it}^1 component (existence of a supra-tribal rule) may also take the value 0.875 (an average of 0.75 and 1) to indicate the joint presence of a paramount chiefdom and a full-fledged state on the territory of country i. An example is the case of Armenia between 1275 and 840 BCE, when the Urartu (indigenous paramount chiefdom) and the Assyrian rule (full state) coexist on its territory.

3) The z_{it}^2 component (is the government imposed from within or from without) may take the value 0.875 (an average of 1 and 0.75) to indicate that the state is largely self-governed, with some foreign oversight. For instance, Austria qualifies for this score between 1945 and 1955, when although largely independent, it was closely observed by the Allied powers, which were still occupying some territories.

4) The z_{it}^2 component (is the government imposed from within or from without) may take the value 0.9375 (an average of 1 and 0.875) to indicate that most of the territory is governed from within, but some part of the territory is influenced or supported (at best nominally) by an external government. For instance, Cyprus receives this score from 1960 onwards, as the northern part of the island was declared the independent Turkish Republic of Northern Cyprus, recognized only by Turkey.

Figure C1 shows the comparison between the original old Statehist index for the CE-period and the new one presented in this paper. A notable feature is that the previous index failed to reflect the state history of several ancient civilizations like Egypt and Iran, countries that now receive a substantially higher score. Ethiopia, which had a full state by 1 CE but was a relative newcomer compared to those just mentioned, has the highest state history value going by the old measure, but now loses in relative terms to the older civilizations.

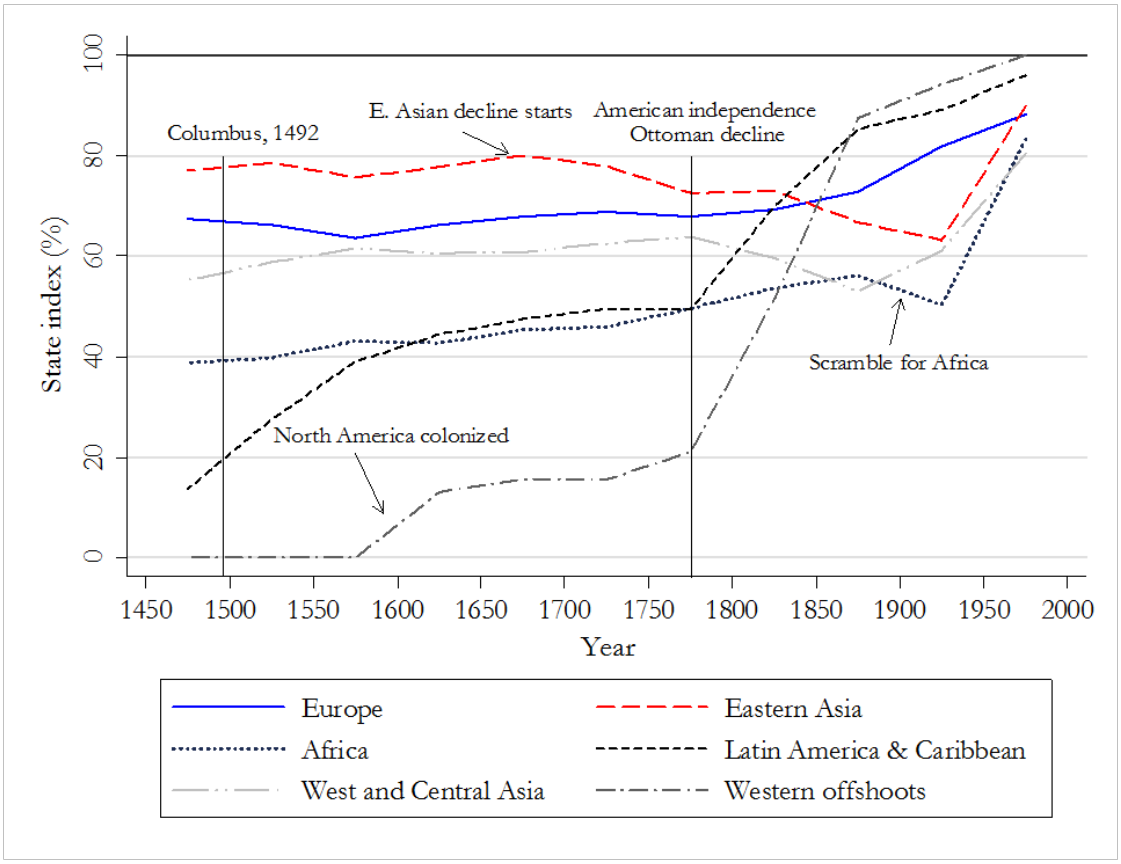
Figure C1: Extended Statehist (3500 BCE-2000 CE) vs Statehist 1-1950 CE



Note: The red, triangular observations with 3-letter isocodes show country observations where states emerged before 1 CE and whose index score changes considerably with the extended coding. Both variables use a 1 percent discount rate.

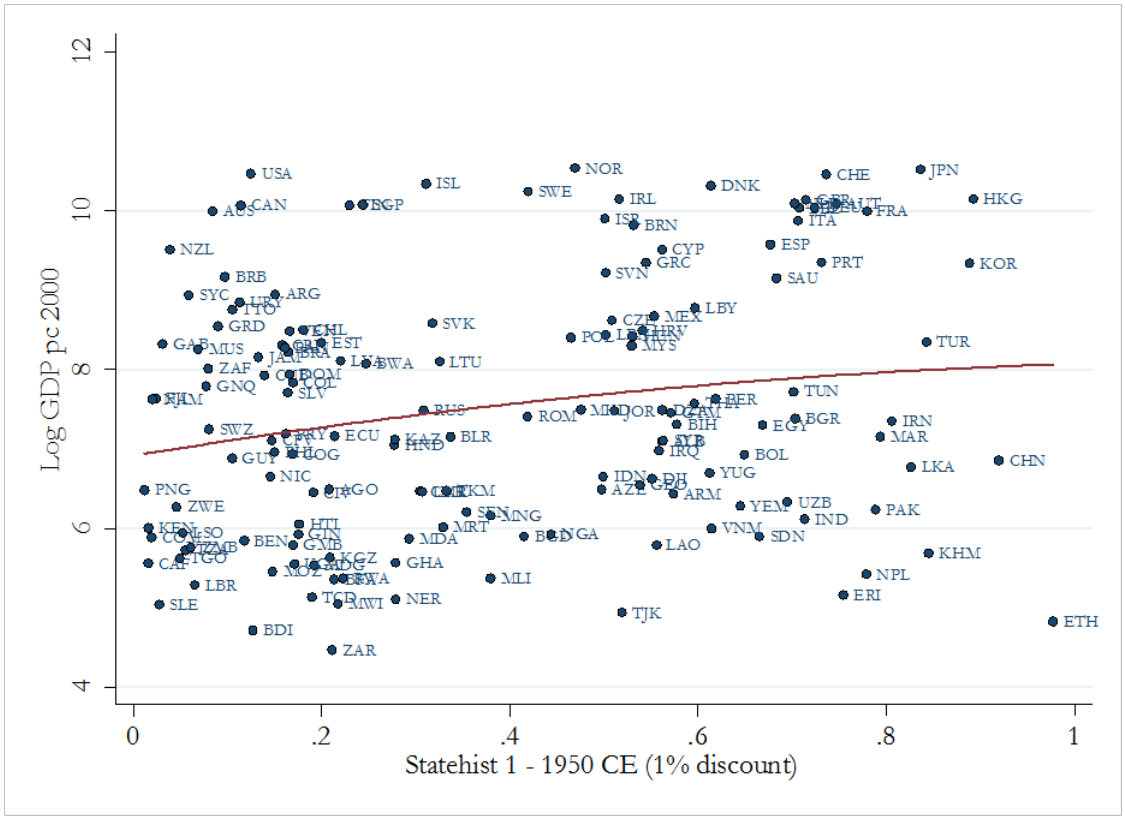
Appendix D - Supplementary Figures

Figure D1: Emergence of states in six world regions during the colonial era, 1450-2000 CE



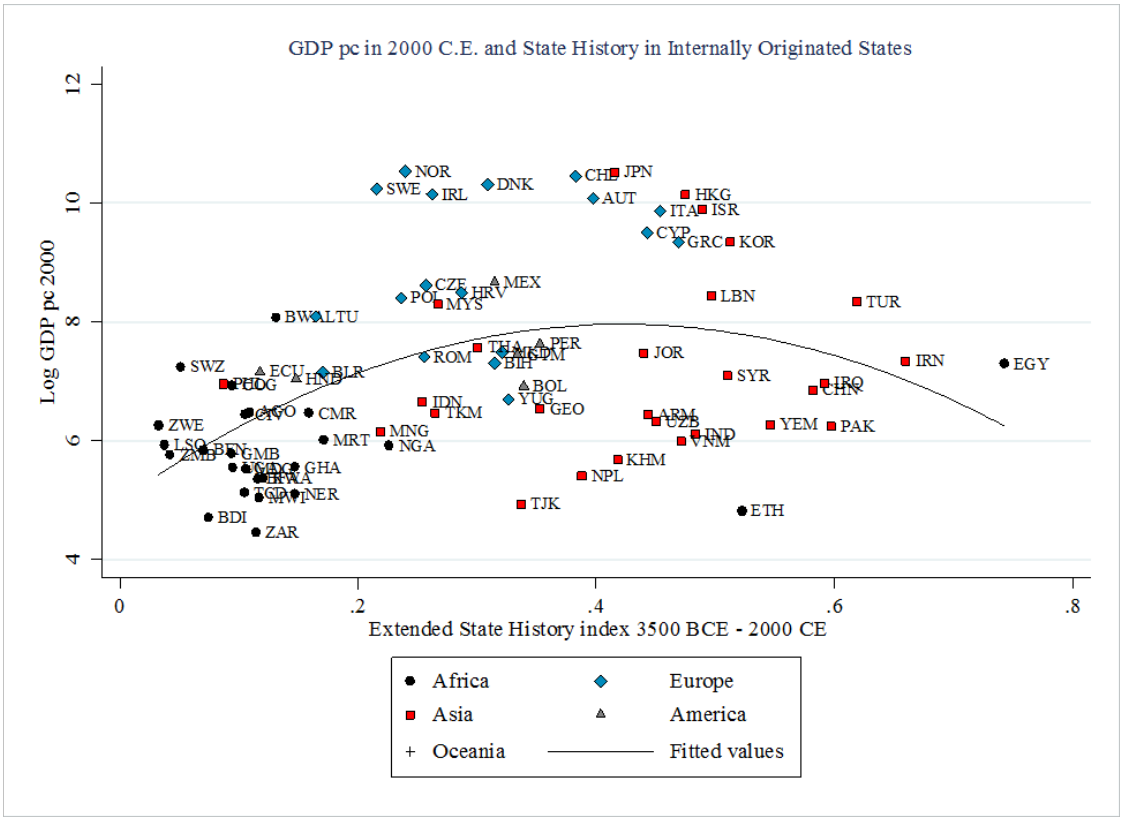
Note: The figure shows the development of the aggregated State index in Europe, Eastern Asia, West and Central Asia (including Turkey and India and the located countries in between), Latin America and the Caribbean (all countries in the Americas except Canada and USA), Africa (including North Africa), and the Western offshoots (USA, Canada, Australia, and New Zealand). Oceania is omitted. On the horizontal axis, negative values imply years BCE whereas positive values show the CE-period. Particular years with trend breaks are marked.

Figure D2: Non-linear relationship between Log GDP per capita in 2000 and Statehist 1-1950 CE



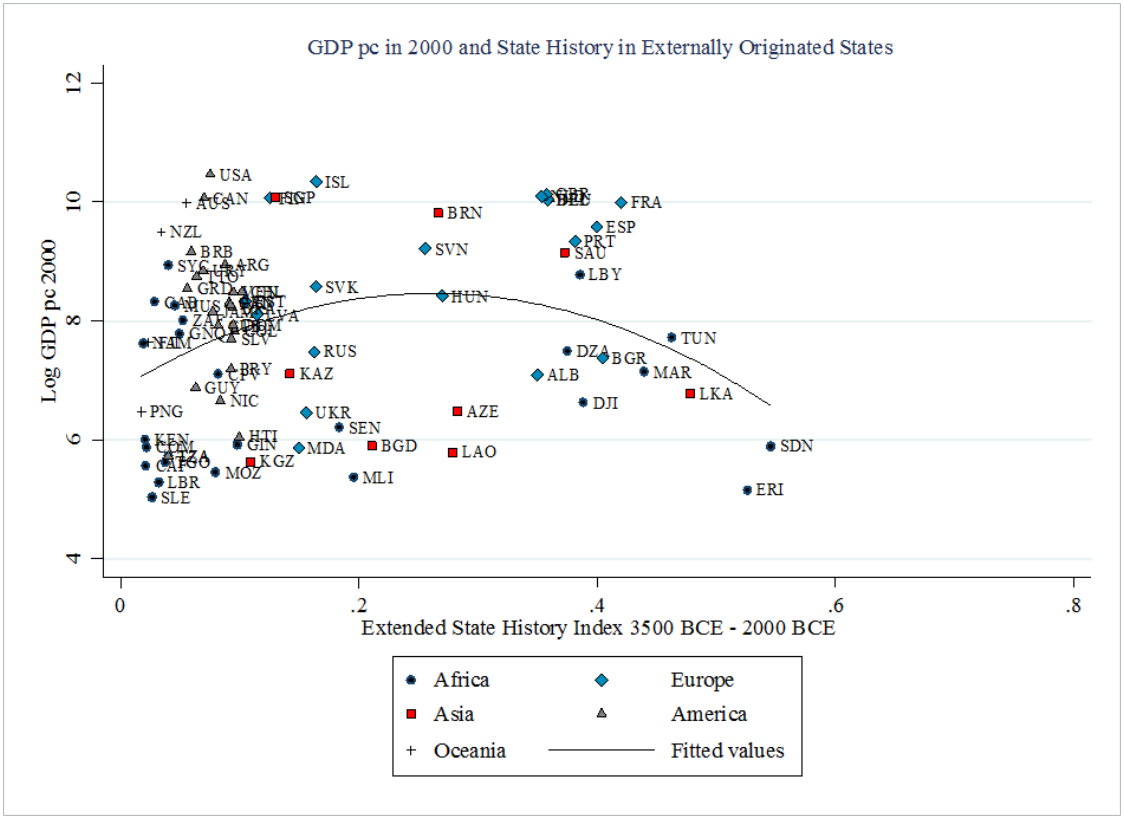
Note: The figure shows a fitted quadratic regression line corresponding to the estimates in Table 2, Panel B, column 2, with 154 country observations distinguished by 3-letter country isocodes.

Figure D3: GDP pc in 2000 C.E. and state history in internally-originated states



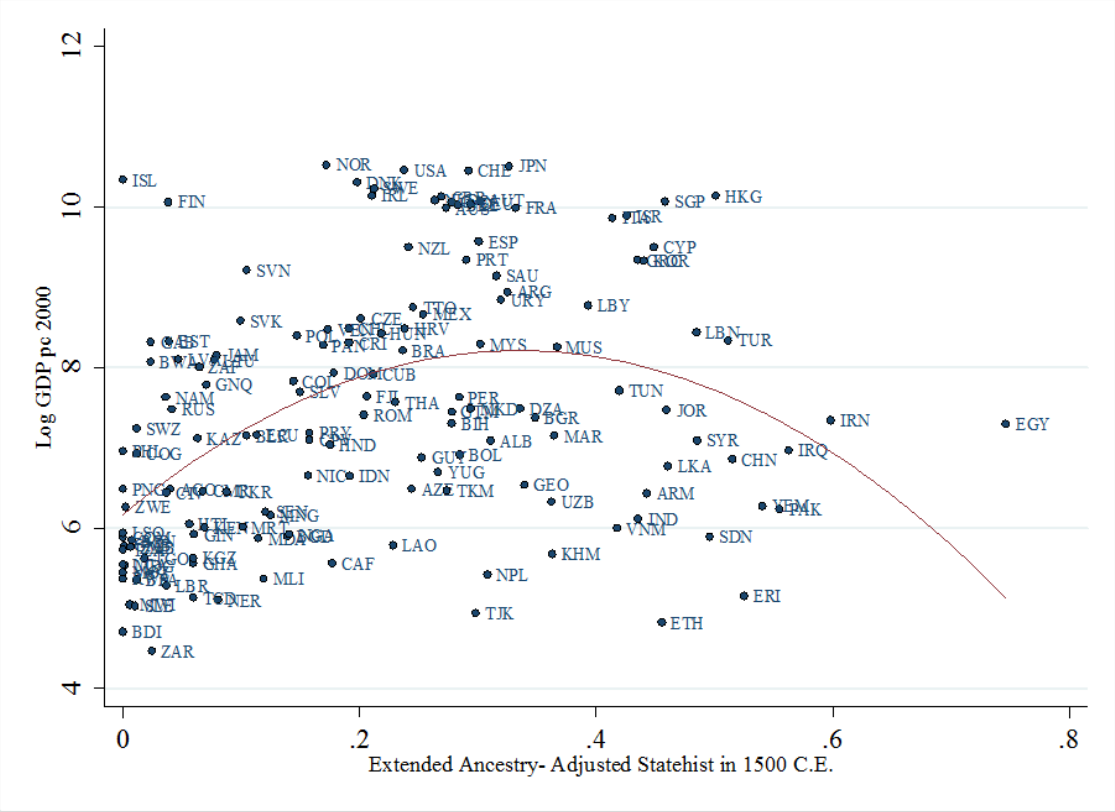
Note: The figure shows the relationship between Log GDP pc 2000 and Statehist, including a quadratic fit, for the subsample of countries where states were internally-originated.

Figure D4: GDP pc in 2000 C.E. and state history in externally-originated states



Note: The figure shows the relationship between Log GDP pc 2000 and Statehist, including a quadratic fit, for the subsample of countries where states were externally-originated.

Figure D5: GDP pc in 2000 C.E. and ancestry- adjusted state history



Note: The figure shows the relationship between Log GDP pc 2000 and ancestry-adjusted Statehist, including a quadratic fit, for all the countries in the sample.

Appendix E - Supplementary Tables

Table E1: Robustness checks-Log (GDP) per capita in 2000 and Statehist.

	Log (GDP) per capita in 2000				
	(1)	(2)	(3)	(4)	(5)
Statehist	9.565***	16.357***	5.985**	7.306**	7.306**
	(2.547)	(4.402)	(2.382)	(2.867)	(2.867)
Statehist squared	-6.609**	-15.975***	-3.803	-4.791	-4.791
	(2.746)	(5.219)	(2.926)	(3.083)	(3.083)
State Age	-0.663***	-1.892***	-0.434**	-0.557***	-0.557***
	(0.195)	(0.546)	(0.186)	(0.212)	(0.212)
Agyears	-0.480***	-0.040	0.009	-0.001	-0.001
	(0.171)	(0.078)	(0.083)	(0.091)	(0.091)
Agyears squared	0.047***				
	(0.016)				
State Age		0.242***			
squared		(0.087)			
Absolute centroid			-0.022		
latitude			(0.031)		
Absolute centroid			0.001		
latitude squared			(0.000)		
Distance from				-0.064	
Addis Ababa				(0.110)	
Distance from				-0.001	
Addis Ababa squared				(0.003)	
Predicted genetic					22.403
diversity					(73.470)
Predicted genetic					-9.011
Diversity squared					(55.974)
Constant	11.299***	11.312***	10.765***	12.264***	0.320
	(1.459)	(1.492)	(1.545)	(1.804)	(24.057)
Controls	yes	yes	yes	yes	yes
Continent FE	yes	yes	yes	yes	yes
Observations	125	125	125	125	125
R-squared	0.743	0.741	0.725	0.730	0.730

Note: The variables "Distance from Addis Ababa" and "Predicted Genetic Diversity" are those constructed by Ashraf and Galor (2013). Robust standard errors in parentheses.***p<0.01, ** p<0.05, * p<0.1

Table E2: Robustness checks-Log (GDP) per capita in 2000 and ancestry-adjusted statehist.

	Log (GDP) per capita in 2000				
	(1)	(2)	(3)	(4)	(5)
Ancestry – Adjusted	7.326***	7.875***	6.657***	6.919***	6.919***
Statehist	(1.724)	(1.911)	(1.890)	(1.918)	(1.918)
Ancestry – Adjusted	-8.103***	-9.780***	-7.532***	-7.476***	-7.476***
Statehist squared	(2.286)	(2.975)	(2.365)	(2.473)	(2.473)
State Age	-0.334**	-0.052	-0.017	-0.013	-0.013
	(0.154)	(0.079)	(0.083)	(0.088)	(0.088)
Agyears	-0.300**	-0.583**	-0.246*	-0.322**	-0.322**
	(0.138)	(0.247)	(0.147)	(0.154)	(0.154)
Agyears squared	0.029**				
	(0.014)				
State Age		0.071			
squared		(0.049)			
Absolute Centroid			-0.032		
Latitude			(0.032)		
Absolute Centroid			0.001		
Latitude squared			(0.001)		
Distance from				0.007	
Addis Ababa				(0.089)	
Distance from				-0.003	
Addis Ababa squared				(0.003)	
Predicted Genetic					70.494
diversity					(60.738)
Predicted Genetic					-46.137
Diversity squared					(46.088)
Constant	10.416***	10.380***	9.861***	11.160***	-15.763
	(1.527)	(1.578)	(1.573)	(1.789)	(20.284)
Controls	yes	yes	yes	yes	yes
Continent FE	yes	yes	yes	yes	yes
Observations	125	125	125	125	125
R-squared	0.737	0.732	0.733	0.734	0.734

Note: The variables "Distance from Addis Ababa" and "Predicted Genetic Diversity" are those constructed by Ashraf and Galor (2013). Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table E3: State history, Log Population Density and Urbanization in 1 CE.

Panel A	Log Population Density in 1 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist in 1 CE	4.350***	8.417***		-1.237	1.270	2.754	4.553
	(0.810)	(1.837)		(2.175)	(2.645)	(1.874)	(3.222)
Statehist in 1 CE squared		-8.254***		0.880	-1.063	-3.034	-4.265
		(2.838)		(3.251)	(4.258)	(2.512)	(3.322)
Agyears in 1 CE			0.455***	0.490***	0.456***	0.457***	0.457***
			(0.040)	(0.063)	(0.079)	(0.081)	(0.084)
Origtime in 1 CE					0.001	-0.016***	-0.010
					(0.004)	(0.006)	(0.010)
Origtime in 1 CE squared							-0.000
							(0.000)
State Age in 1 CE							-0.213
							(0.266)
Constant	-0.425***	-0.518***	-1.469***	-1.510***	-6.309***	-6.073***	-6.215***
	(0.140)	(0.144)	(0.159)	(0.168)	(1.267)	(1.296)	(1.270)
Observations	135	135	130	130	115	115	115
R-squared	0.154	0.182	0.455	0.458	0.717	0.800	0.803

Panel B	Urbanization in 1 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist in 1 CE	1.128***	2.778***		0.599	0.235	1.312	1.500
	(0.237)	(0.624)		(0.723)	(1.066)	(0.958)	(1.761)
Statehist in 1 CE squared		-3.260***		-1.013	-0.576	-1.312	-1.452
		(1.045)		(1.117)	(1.474)	(1.164)	(1.404)
Agyears in 1 CE			0.101***	0.093***	0.077**	-0.040	-0.039
			(0.020)	(0.027)	(0.034)	(0.039)	(0.039)
Origtime in 1 CE					0.004*	-0.000	-0.001
					(0.002)	(0.002)	(0.005)
Origtime in 1 CE squared							0.000
							(0.000)
State Age in 1 CE							-0.020
							(0.191)
Constant	2.566***	2.533***	2.348***	2.354***	1.592**	1.304	1.300
	(0.064)	(0.069)	(0.100)	(0.105)	(0.759)	(0.990)	(0.989)
Observations	128	128	128	128	125	125	125
R-squared	0.063	0.087	0.139	0.141	0.371	0.526	0.526
Controls	no	No	no	no	yes	yes	yes
Continent FE	no	No	no	no	no	yes	yes

Note: In panel A, the dependent variable is log population density in 1 CE and the main independent variables are the Statehist index between 3500 BCE and 1 CE, linear and squared. In panel B, the dependent variable is the urbanization rate in 1 CE and the main independent variables are the Statehist index between 3500 BCE and 1 CE, linear and squared. The data for historical population density is based on population data from McEvedy and Jones(1978) and land data from World Bank's World Development Indicators. The data for urbanization rate in 1 CE is taken from Comin, Easterly and Gong (2010) and is based on Peregrine (2003) and takes three values: 1 if the largest settlement is smaller than 100 persons; 2 if it is between 100 and 399 persons; and 3 if it is larger than 400 persons. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E4: State history and average technology adoption in 1 CE. Ancestry – Adjusted State history and technology adoption in 2000 CE

Panel A	Technology Adoption in 1 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statehist in 1 CE	0.762***	1.695***		0.415	0.016	0.534	0.788
	(0.122)	(0.279)		(0.390)	(0.519)	(0.351)	(0.601)
Statehist in 1 CE		-1.842***		-0.524	-0.150	-0.691*	-0.862*
squared		(0.479)		(0.572)	(0.746)	(0.412)	(0.508)
Agyears in 1 CE			0.064***	0.055***	0.064***	0.023	0.031*
			(0.007)	(0.012)	(0.016)	(0.015)	(0.016)
Origtime in 1 CE					0.001	-0.002***	-0.005***
					(0.001)	(0.001)	(0.002)
Origtime in 1 CE							0.000**
squared							(0.000)
State Age in 1 CE							-0.026
							(0.060)
Constant	0.688***	0.669***	0.554***	0.562***	-0.062	-0.123	-0.130
	(0.027)	(0.028)	(0.037)	(0.039)	(0.337)	(0.332)	(0.322)
Observations	128	128	128	128	124	124	124
R-squared	0.149	0.189	0.281	0.285	0.541	0.763	0.772
Panel B	Technology Adoption in 2000 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry – Adjusted	0.233**	1.332***		1.329***	0.794***	0.672***	0.787***
Statehist in 1500 CE	(0.092)	(0.275)		(0.309)	(0.215)	(0.219)	(0.223)
Ancestry – Adjusted		-2.088***		-2.085***	-1.125***	-0.938***	-0.887***
Statehist in 1500 CE sqr.		(0.513)		(0.514)	(0.287)	(0.317)	(0.303)
Agyears			0.011	-0.000	-0.010	-0.005	0.001
			(0.007)	(0.011)	(0.009)	(0.011)	(0.012)
Origtime					0.001*	0.001**	0.001
					(0.000)	(0.001)	(0.002)
Origtime squared							0.000
							(0.000)
State Age							-0.029
							(0.018)
Constant	0.399***	0.314***	0.405***	0.315***	0.708***	0.769***	0.711***
	(0.023)	(0.025)	(0.038)	(0.033)	(0.164)	(0.207)	(0.208)
Observations	130	130	129	129	125	125	125
R-squared	0.037	0.151	0.016	0.150	0.675	0.698	0.704
Controls	no	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	no	yes	yes

Note: In panel A, the dependent variable is the technology adoption index in 1 CE and the main independent variables are the Statehist index between 3500 BCE and 1 CE, linear and squared. In panel B , the dependent variable is the technology adoption index in 2000 CE and the main independent variables are the ancestry-adjusted Statehist index, between 3500 BCE and 1500 CE, linear and squared. The list of controls includes: absolute latitude, an indicator of whether the present-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E5: Robustness checks-Technology adoption (excluding agriculture) in 1 CE and 1500 CE and statehist.

Panel A	Technology adoption in 1 C.E. excluding agriculture					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1 CE	2.012***		0.526	0.113	0.600	1.025
	(0.328)		(0.464)	(0.564)	(0.391)	(0.691)
Statehist in 1 CE	-2.179***		-0.650	-0.271	-0.799*	-1.095*
squared	(0.567)		(0.682)	(0.821)	(0.462)	(0.587)
Agyears in 1 CE		0.075***	0.064***	0.071***	0.028	0.036*
		(0.008)	(0.014)	(0.017)	(0.017)	(0.019)
Origtime in 1 CE				0.001	-0.002***	-0.006***
				(0.001)	(0.001)	(0.002)
Origtime in 1 CE						0.000*
squared						(0.000)
State Age in 1 CE						-0.044
						(0.069)
Constant	0.604***	0.469***	0.480***	-0.152	-0.105	-0.115
	(0.032)	(0.041)	(0.043)	(0.391)	(0.403)	(0.397)
Observations	128	128	128	124	124	124
R-squared	0.201	0.293	0.298	0.577	0.759	0.768

Panel B	Technology adoption in 1500 C.E. excluding agriculture					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1500	2.969***		1.747***	1.783***	1.019***	1.549***
CE	(0.386)		(0.484)	(0.339)	(0.197)	(0.447)
Statehist in 1500	-3.524***		-2.684***	-2.015***	-0.728**	-0.857***
CE squared	(0.797)		(0.948)	(0.643)	(0.322)	(0.295)
Agyears in 1500		0.107***	0.076***	0.041***	-0.000	0.004
CE		(0.009)	(0.015)	(0.014)	(0.010)	(0.011)
Origtime in 1500				0.001**	-0.001*	0.001
CE				(0.001)	(0.001)	(0.001)
Origtime in 1500						-0.000*
CE squared						(0.000)
State Age in 1500						-0.071
CE						(0.055)
Constant	0.169***	-0.013	-0.007	-0.042	0.031	-0.053
	(0.027)	(0.039)	(0.039)	(0.266)	(0.149)	(0.161)
Observations	111	109	109	107	107	107
R-squared	0.498	0.510	0.589	0.813	0.915	0.922
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes

Note: The list of controls includes: absolute latitude, an indicator whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E6: State history and Log Population Density in 1 CE and 1500 CE-linear relationship.

Panel A	Log Population Density in 1 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1 CE	4.350***		-0.734	0.583	0.793	0.614
	(0.810)		(0.861)	(1.097)	(0.811)	(1.025)
Agyears in 1 CE		0.455***	0.484***	0.466***	0.481***	0.463***
		(0.040)	(0.057)	(0.068)	(0.078)	(0.083)
Origtime in 1 CE				0.001	-0.016***	-0.008
				(0.004)	(0.005)	(0.010)
Origtime in 1 CE squared						-0.000
						(0.000)
State Age in 1 CE						0.031
						(0.184)
Constant	-0.425***	-1.469***	-1.507***	-6.369***	-6.193***	-6.262***
	(0.140)	(0.159)	(0.167)	(1.281)	(1.297)	(1.282)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	135	130	130	115	115	115
R-squared	0.154	0.455	0.458	0.717	0.798	0.799
Panel B	Log Population Density in 1500 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1500 CE	3.883***		0.902	3.653***	3.605***	6.072***
	(0.670)		(0.957)	(0.943)	(0.915)	(1.803)
Agyears in 1500 CE		0.315***	0.269***	0.180***	0.115*	0.169**
		(0.042)	(0.064)	(0.064)	(0.069)	(0.069)
Origtime in 1500 CE				0.006**	-0.003	-0.017
				(0.003)	(0.004)	(0.012)
Origtime in 1500 CE squared						0.000
						(0.000)
State Age in 1500 CE						-0.334*
						(0.180)
Constant	0.359**	-0.411*	-0.345	-2.761*	-3.271**	-3.382***
	(0.144)	(0.212)	(0.223)	(1.487)	(1.343)	(1.258)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	154	147	147	128	128	128
R-squared	0.184	0.269	0.273	0.687	0.759	0.770

Note: The data for historical population density is based on population data from McEvedy and Jones(1978) and land data from World Bank's World Development Indicators. The list of controls includes: absolute latitude, an indicator whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E7: State history and urbanization in 1 CE and 1500 CE-linear relationship

Panel A	Urbanization in 1 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1 CE	1.128***		0.003	-0.135	0.475	0.170
	(0.237)		(0.232)	(0.360)	(0.313)	(0.728)
Agyears in 1 CE		0.101***	0.101***	0.082***	-0.030	-0.037
		(0.020)	(0.025)	(0.028)	(0.039)	(0.039)
Origtime in 1 CE				0.004*	-0.000	0.000
				(0.002)	(0.002)	(0.004)
Origtime in 1 CE squared						-0.000
						(0.000)
State Age in 1 CE						0.063
						(0.154)
Constant	2.566***	2.348***	2.348***	1.560**	1.260	1.285
	(0.064)	(0.100)	(0.104)	(0.755)	(0.995)	(0.984)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	128	128	128	125	125	125
R-squared	0.063	0.139	0.139	0.371	0.523	0.524

Panel B	Urbanization in 1500 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1500 CE	16.678***		18.666***	22.878***	21.837***	30.860*
	(2.384)		(3.452)	(6.078)	(7.446)	(16.414)
Agyears in 1500 CE		0.761***	-0.201	-0.135	-0.598	-0.411
		(0.177)	(0.239)	(0.356)	(0.474)	(0.472)
Origtime in 1500 CE				-0.052	-0.070*	-0.141
				(0.036)	(0.041)	(0.131)
Origtime in 1500 CE squared						0.001
						(0.001)
State Age in 1500 CE						-1.256
						(1.693)
Constant	4.487***	3.633***	5.119***	-7.853	-7.314	-8.246
	(0.569)	(1.019)	(1.015)	(8.277)	(8.019)	(8.291)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	83	83	83	76	76	76
R-squared	0.278	0.111	0.282	0.386	0.458	0.467

Note: The data for urbanization rate in 1 CE is taken from Comin, Easterly and Gong (2010) and is based on Peregrine (2003) and takes three values: 1 if the largest settlement is smaller than 100 persons; 2 if it is between 100 and 399 persons; and 3 if it is larger than 400 persons. The data for urbanization rate at 1500 CE is that reported by Acemoglu, Johnson and Robinson (2005), defined as the percentage of a country's total population residing in urban areas (each with a city population size of at least 5,000). The list of controls includes: absolute latitude, an indicator whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E8: State history and average technology adoption in 1 CE and 1500 CE-linear relationship

Panel A	Technology Adoption in 1 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1 CE	0.762***		0.106	-0.080	0.093	-0.001
	(0.122)		(0.122)	(0.196)	(0.134)	(0.246)
Agyears in 1 CE		0.064***	0.059***	0.065***	0.028**	0.032*
		(0.007)	(0.010)	(0.013)	(0.014)	(0.016)
Origtime in 1 CE				0.001	-0.002***	-0.005***
				(0.001)	(0.001)	(0.002)
Origtime in 1 CE squared						0.000*
						(0.000)
State Age in 1 CE						0.023
						(0.048)
Constant	0.688***	0.554***	0.559***	-0.071	-0.147	-0.139
	(0.027)	(0.037)	(0.039)	(0.332)	(0.329)	(0.320)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	128	128	128	124	124	124
R-squared	0.149	0.281	0.283	0.541	0.759	0.769

Panel B	Technology Adoption in 1500 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist in 1500 CE	1.227***		0.288	0.754***	0.684***	1.140***
	(0.157)		(0.193)	(0.172)	(0.135)	(0.353)
Agyears in 1500 CE		0.104***	0.088***	0.042***	-0.000	0.004
		(0.008)	(0.013)	(0.013)	(0.011)	(0.011)
Origtime in 1500 CE				0.001	-0.001**	0.001
				(0.001)	(0.001)	(0.001)
Origtime in 1500 CE squared						-0.000
						(0.000)
State Age in 1500 CE						-0.070
						(0.047)
Constant	0.315***	0.065*	0.090**	-0.143	0.026	-0.062
	(0.030)	(0.036)	(0.037)	(0.258)	(0.170)	(0.180)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	112	110	110	107	107	107
R-squared	0.389	0.532	0.541	0.776	0.897	0.904

Note: The list of controls includes: absolute latitude, an indicator whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E9: State history, ancestry-adjusted state history and technology adoption in 2000 CE-linear relationshi

Panel A	Technology Adoption in 2000 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Statehist	0.086		-0.074	0.063	0.128	0.289*
	(0.095)		(0.132)	(0.106)	(0.106)	(0.174)
Agyears		0.011	0.015	-0.008	-0.008	-0.000
		(0.007)	(0.010)	(0.008)	(0.011)	(0.012)
Origtime				0.000	0.001**	0.001
				(0.000)	(0.001)	(0.002)
Origtime squared						-0.000
						(0.000)
State Age						-0.032
						(0.021)
Constant	0.430***	0.405***	0.406***	0.761***	0.818***	0.791***
	(0.029)	(0.038)	(0.038)	(0.162)	(0.206)	(0.206)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	130	129	129	125	125	125
R-squared	0.006	0.016	0.018	0.639	0.679	0.684
Panel B	Technology Adoption in 2000 CE					
	(1)	(2)	(3)	(4)	(5)	(6)
Ancestry – Adjusted	0.233**		0.231	0.215*	0.194*	0.354**
Statehist in 1500 CE	(0.092)		(0.156)	(0.121)	(0.115)	(0.161)
Agyears		0.011	-0.000	-0.013	-0.011	-0.003
		(0.007)	(0.012)	(0.009)	(0.012)	(0.012)
Origtime				0.000	0.001**	0.001
				(0.000)	(0.001)	(0.002)
Origtime squared						0.000
						(0.000)
State Age						-0.034*
						(0.018)
Constant	0.399***	0.405***	0.401***	0.754***	0.790***	0.721***
	(0.023)	(0.038)	(0.035)	(0.162)	(0.203)	(0.204)
Controls	no	no	no	yes	yes	yes
Continent FE	no	no	no	no	yes	yes
Observations	130	129	129	125	125	125
R-squared	0.037	0.016	0.036	0.649	0.683	0.690

Note: The list of controls includes: absolute latitude, an indicator whether the modern-day country is landlocked, distance to coast and rivers, mean elevation, land suitability, percentage arable land, temperature, precipitation, percentage population at risk of contracting malaria. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

References

[1] Acemoglu, D., Johnson, S. and Robinson, J. (2005). "Institutions As the Fundamental Cause of Long Run Economic Growth" in Handbook of Economic Growth, Aghion, P. and Durlauf, S. eds., North-Holland, Amsterdam.

[2] Ahlerup, P. and Olsson, O. (2012). "The roots of ethnic diversity" Journal of Economic Growth, Springer, vol. 17(2), pp. 71 – 102, June.

[3] Ashraf, Q. and Galor, O. (2011). "Dynamics and Stagnation in the Malthusian Epoch" American Economic Review, 101(5), pp. 2003 – 2041.

[4] Ashraf, Q. and Galor, O. (2013). "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development" American Economic Review, 103(1), pp. 1 – 46.

[5] Bockstette, V., Chanda, A. and Putterman, L. (2002). "States and Markets: The Advantage of an Early Start" Journal of Economic Growth, Springer, vol. 7(4), pp. 347 – 69, December.

[6] Bradshaw Foundation (2007), available online at http://www.bradshawfoundation.com/stephenoppenheimer/.

[7] Comin, D., Easterly, W. and Gong, E. (2010). "Was the Wealth of Nations Determined in 1000 BC?" American Economic Journal: Macroeconomics, American Economic Association, vol. 2(3), pp. 65 – 97, July.

[8] Comin D., Hobijn, B. and Rovito, E., (2008). "A new approach to measuring technology with an application to the shape of the diffusion curves," The Journal of Technology Transfer, Springer, vol. 33(2), pages 187 – 207, April.

[9] Gallup, J. L., and Sachs., J. D. (2001). "The Economic Burden of Malaria," The American Journal of Tropical Medicine and Hygiene, 64(1-2): 85 – 96.

[10] McEvedy, C., and Jones, R.. (1978). "Atlas of World Population History," New York, NY:Penguin Books Ltd.

[11] Michalopoulos, S. (2012). "The Origins of Ethnolinguistic Diversity," American Economic Review, 102(4): 1508 – 39.

[12] Nordhaus, W. D. (2006). "Geography and Macroeconomics: New Data and New Findings," Proceedings of the National Academy of Sciences, 103(10): 3510 – 3517.

[13] Oppenheimer, S. (2003). "Out of Eden: The peopling of the World," London: Constable.

[14] Peregrine, P. N. (2003). "Atlas of Cultural Evolution," World Cultures: Journal of Comparative and Cross-Cultural Research, 14(1): 1 – 75.

[15] Putterman, L. with C.A Trainor (2006) "Agricultural Transition Year - Country Data Set," http://www.econ.brown.edu/fac/louis putterman/.

[16] Putterman, L. and Weil, D. N. (2010). "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," The Quarterly Journal of Economics, MIT Press, vol. 125(4), pp. 1627 – 1682, November.

[17] Ramankutty, N., Foley, A. J., Norman, J. and McSweeney, K. (2002). "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," Global Ecology and Biogeography, 11(5): 377 – 392.

Chapter II

The Impact of an Unexpected Wage Cut on Corruption: Evidence from a "Xeroxed" Exam*†

Oana Borcan
University of Gothenburg

Mikael Lindahl
University of Gothenburg

Andreea Mitrut‡
University of Gothenburg

Published in Journal of Public Economics (2014), 120: 32-47

Abstract

This paper aims to understand how corruption responds to an income loss. We exploit an unexpected 25% wage cut incurred in 2010 by all Romanian public sector employees, including the public education staff. We investigate a corruptible high-stake exam taking place shortly after the wage announcement. To measure corruption we compare changes in exam outcomes from 2007 to 2010 between public and private schools, as the latter were not affected by the policy. We find that the wage loss induced better exam outcomes in public than in private schools and we attribute this difference to increased corruption by public educators.

JEL Codes: H4, I2, J3

Keywords: austerity, bribes, public school principals, high-stakes exams

*All errors are our own. Andreea Mitrut gratefully acknowledges support from Jan Wallanders and Tom Hedelius Fond. Mikael Lindahl is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Torsten and Ragnar Sderberg Foundation, and also acknowledges financial support from the Scientific Council of Sweden [project 325-2009-7665] and the European Research Council [ERC starting grant 241161].

†We are grateful to our editor Brian Knight, two anonymous referees and to Per- Anders Edin, David Figlio, Peter Fredriksson, Rita Ginja, Per Johansson, Scott Imberman, Olof Johansson-Stenman, Edwin Leuven, Martin Ljunge, Eva Mrk, Ola Olsson, Sonja Opper, Amrish Patel, Maria Perrotta, as well as seminar participants at SOFI, Uppsala University, the 2013 DIAL conference, the 2012 NCDE conference, the 2012 CESifo Economics of Education conference in Munich, the Swedish 2012 National Conference in Stockholm, 11th EUDN Workshop in Toulouse and the HECER Economics of Education 2012 Summer Meeting in Helsinki for helpful discussions and suggestions. We also thank Diana Coman and SIVECO Romania for excellent help with the data.

‡Corresponding author: Andreea Mitrut, Dept Economics, University of Gothenburg, Box 630, 40530 Gothenburg, Sweden. andreea.mitrut@economics.gu.se.

1 Introduction

The last decades have witnessed fast growing political and academic efforts to break down the phenomenon of corruption into causes and effects. To date, many puzzles still remain regarding the key causes and determinants of corruption (see Olken and Pande, 2012 for a recent review of developments in this area). Among these, the degree to which corruption responds to a wage change is an underexplored topic of particular interest to policy makers. This paper attempts to shed light on the effects of wages on corruption in the public sector, exploring a quasi-natural experiment generated by an unexpected 25% wage cut incurred by the public sector employees in Romania in 2010. Understanding the consequences of a wage loss, especially for corruption, is particularly relevant in the context of the recent waves of austerity measures that have swept over most of the EU countries.[1] To our knowledge, this is the first paper that identifies a causal relationship between a wage cut in the public sector and corruption activities.

The idea that financial compensation is a crucial factor in the decision of whether to engage in fraudulent action was first formalized in 1974 with Becker and Stigler's seminal work. The key prediction from their model was that increasing the remuneration of public servants above the market-clearing wage can reduce bribery, and thus reduce the prevalence of corruption. Subsequently, this hypothesis has been empirically tested, initially using macro-level data. For example, exploring a cross-section of developing countries, Van Rijkenghem and Weder (2001) show a negative, but rather small, association between civil service compensation and corruption measured by the ICRG index, while Rauch and Evans (2000) find no significant relationship between bureaucrats' wages and corruption, but show that salaries correlate negatively with the bureaucratic delay. To date, few studies have used microlevel data to identify the deterrent effect that wages have on corruption. Di Tella and Schargrodsky (2003) exploit a crackdown on corruption in the procurement departments of Buenos Aires hospitals. They find that at higher levels of the staff's wages the crackdown is more effective in reducing the prices of hospital inputs when there is an intermediate level of monitoring. However, they also show that higher wages have no statistically significant effect when there is no monitoring or when monitoring is at a very high level. These results are consistent with the predictions of the Becker-Stigler model. Niehaus and Sukhtankar (2010) also find empirical support for the capacity of projected gains to reduce fraud. In this setting, however, the prospective rents are obtained from future opportunities to

[1] Similar measures regarding cuts in public sector wages have been proposed in other EU countries, e.g., Greece in 2011 and Spain in 2012.

collect bribes that rely strictly on keeping the job, which leads to an inter-temporal substitution of fraud today for rent-extraction in the future.[2]

While these studies are centered on the effect of an increase in remuneration on dishonesty, it is not obvious that a decrease in wages would have a symmetric impact on corruption.[3] Gorodnichenko and Sabirianova Peter (2007), to our knowledge, is the only study that has analyzed corruption in direct relation to low wages. Using micro data from Ukraine, these authors show that the wage differential between the private and (the much lower-paid) public sector does not translate into a difference in consumption, and they conclude that bribery must account for the observed wage gap. In doing so, they document the role of corruption in explaining the prevalence of low-paid public jobs, rather than the reverse. Thus, the impact of a decrease in wages on the prevalence of corruption, the object of our study, remains an open empirical question.

In the spirit of the shirking model proposed by Shapiro and Stiglitz (1984), lower wages could trigger a switchover to rents from corrupt activities, as the civil servant attempts to compensate for his lost income. At the same time, a different mechanism, working in the opposite direction, holds the prospect of unemployment as a deterrent for shirking or, as applied to our case, corruption (Shapiro and Stiglitz, 1984). Thus, particularly in a depressed economic time, as in 2010, an income loss may potentially prompt more risk-averse public employees to refrain from corruption because they fear losing their job and their only source of income when the market cannot accommodate them. The latter mechanism is also supported by an argument la Niehaus and Sukhtankar (2010) that the need to keep the public job with future bribe opportunities (relatively more lucrative than the diminished wage), may drive a temporary drop in corruption. Overall, these mechanisms convey an ambiguous

[2] Armantier and Boly (2011) carry out a controlled field experiment on the receptiveness of exam graders to bribe-offering. The effect of higher wages on corruption tested in their experiment is ambiguous. This paper belongs to a growing experimental literature on corruption using controlled field experiments (see Olken, 2007; Bertrand et al., 2006), as well as lab experiments (see Frank and Schulze, 2000; Abbink, 2002; Schulze and Frank, 2003; Barr et al., 2009; Barr and Serra, 2009). The latter category also yields mixed evidence on the impact of a wage increase on corruption.

[3] According to the prospect theory, agents perceive differently equivalent losses and gains. This is sustained by empirical evidence. For example, Armantier and Boly (2013) show in a field experiment that teachers performed better in a marking task when incentivized with a penalty, rather than with a bonus. If this applies to corruption, it is not clear whether reduced wages would increase corruption. Also, Niehaus and Sukhtankar (2010) argue that a significant wage decrease could increase the reliance on future bribe, and hence on keeping the public sector job with bribe opportunities, thus possibly discouraging an increase in fraud in the present. From the stand point of the wage-corruption relationship, our study is akin to the theoretical underpinnings of Becker and Stigler (1974). However, whereas the bribe in their model is exogenous, our analysis inquiries into how wages can alter corruption intensity. In this respect, our findings relate more closely to Shleifer and Vishny (1993) who take bribes to be endogenous and analyze how they respond to the market structure of corruption.

effect of lower wages on corruption, and identifying their impact is essentially an empirical exercise.

In this paper we show that a large reduction in the wages of civil servants in this case public school principals, together with teachers, and/ or the administration personnel can increase the incidence of corruption. Specifically, our study attempts to measure the effect of an exogenous 25% reduction in wages on corruption in the education sector in Romania. As part of an austerity plan, the Romanian public sector was hit by an unexpected wage cut announced on May 7th 2010, scheduled to take effect starting July 1st 2010. In June 2010, just between the announcement of the cut and its actual implementation, the annual national high school-leaving exam the Baccalaureate took place in the usual manner, testing approximately 200,000 students. The prevalence of corruption at the Baccalaureate exams was notorious and was attributed to the high-stakes character of the exam (it accounts for up to 100% of the university/college admission score) and the poor remuneration of teachers in general. As it happened, the 2010 exam signaled an unprecedentedly high number of allegations of fraud and bribery by school principals connected with the Baccalaureate. The 2010 spike in court investigations by the Romanian National Anticorruption Directorate (DNA), revealed how batches of identical answers had been distributed to students (by public educators), earning the 2010 exam a special title: "The Xeroxed exam".[4] Additional survey data on education corruption in Romania confirms that there was an increase in the incidence of bribery in public education in 2010 compared to 2006.[5]

Since we do not observe bribery and fraud directly, our strategy for understanding the impact of the wage cut on corruption is to compare the change in the Baccalaureate exam outcomes mainly the school level average grades and passing rates of the standardized Romanian language exam from 2007 to 2010 between public and private schools, as the latter category was not affected by the policy.[6] The argu-

[4]This title given by the media refers to the fact that many students were found to have identical test answers (including in essay type exams), which is unlikely to happen without special interventions, given the complexity of the subjects. We will return to the mechanisms of corruption later in the paper.

[5]We use Life in Transition Surveys I and II and rely on the question "In your opinion, how often do people like you have to make unofficial payments or gifts in these situations?"and we focus on public education. The answers range from 1 (never) to 5 (always). A t-test shows that the average score in 2010 is significantly larger than in 2006 (1.76 as opposed to 1.62) and the regression counterpart of this difference remains significant after we control for the usage of public education services.

[6]Because corruption is notoriously difficult to measure, many researchers resort to some indirect assessments, such as evaluating corruption through changes in the outcome of interest when moving into a treatment where corruption is more likely. A similar strategy has been, for example, employed in Olken (2007) or Bertrand et al. (2006).

ments in favor of interpreting the resulting change in exam scores as being due to changes in corruption are the following: 1) the timing between the announcement of the wage cut and the exam is far too short for other responses (for example, a change in the students' or in-class teachers' effort); and 2) using county specific variation in corruption we find that our effects are indeed driven by the most corrupted counties, whereas we find no impact of the wage cut in counties with little or no corruption. If we believe that exogenous shocks to private schools or responses in form of effort are likely to have a similar impact in the most and least corrupted counties, we can conclude that these confounders are unlikely to bias our baseline estimates. However, in Section 5.2 we discuss extensively alternative explanations and possible confounders to our interpretation of the main results.

Our results show a positive and significant change in the exam outcomes between public and private schools, which we attribute to an increase in incentives to engage in corrupt activities in 2010 relative to previous years. In particular, our results for the standardized Romanian written exam, a test which remained similar across years and is taken by all students, regardless of their track, indicate a wage cut-driven effect equivalent to a 0.26 S.D. increase in exam scores and an increase in school-level Romanian exam pass rates by 3.3 percentage points. The estimated effects are equivalent to a nearly 4% increase in both exam outcomes. We employ different falsification tests and sensitivity analysis to lend further credibility to our results.

While this study adds to the developing pool of knowledge about corruption in the education sector (see, for example, Ferraz et al., 2012; Duflo et al., 2014; Reinikka and Svensson, 2004, 2005; Muralidharan and Sundaraman, 2011; Glewwe et al., 2010), it also complements the findings in a related literature investigating incentives for teachers cheating and the dangers of high-stakes evaluation systems (Jacob and Levitt, 2003; Nichols and Berliner, 2007).

The paper is structured as follows: Section 2 presents an overview of the Romanian context, explaining the wage cut policy, the educational system and the implications for corruption. Section 3 provides the details of our data, while Section 4 outlines our empirical strategy and our main empirical findings. Section 5 provides some tests as to whether changes in exam scores following the wage cut can be interpreted as changes in corruption caused by the wage cut, while our conclusions are presented in Section 6.

2 Background

2.1 The 2010 unexpected public sector wage cut

The threat of recession posed by the unfolding international financial crisis in the fall of 2008 was largely overlooked by Romanian politicians, who confidently conveyed a disjunction between Romania and the world economy. The autumn 2008 Euro-barometer showed that more than 70% of Romanian respondents anticipated no change or even an improvement in the general economic situation of Romania.[7] Despite the IMF's prompting for moderation, upon preparing his 2009 electoral campaign and especially after winning the elections, the incumbent president promoted greatly optimistic prospects: "(...) we expect significant growth in the first part of 2010".[8]

May 7th, 2010 involving a 25% cut in wages for all public sector employees, the elimination of some of their financial and in-kind incentives (which were accounting for an additional up to 15% of the monthly remuneration), and a 15% reduction in pensions and unemployment benefits was unexpected, generating social instability and political divergence. The austerity measure was introduced in an attempt to reach the 6.8% budget deficit target agreed upon with the IMF (for more details about the unexpected announcement and the political situation in Romania in 2010, see also Bejenariu and Mitrut, 2012). Soon after, the Finance Minister publicly admitted that the governments' previous optimism had been deceptive.[9] Thus, following the May 7th announcement, on June 30th, the President promulgated the austerity law, which came in effect July 1st, with an initial duration of 6 months, until December 31st, 2010. To date, the public sector wages have not been restored to their initial level.

2.2 The structure of education and the high school exam in Romania

The standard design of the educational system in Romania is based on a division of three cycles, each containing four years: primary school (grades 1 to 4), middle school or gymnasium (grades 5 to 8), followed by a national exam which insures the admission into high schools on a: i) theoretical (or general) track, ii) techno-

[7] http://ec.europa.eu/public_opinion/cf/ "What are your expectations for the year to come with respect to the economic situation of your country (Romania)."

[8] http://www.evz.ro/detalii/stiri/basescu-romania-nu-va-fi-afectata-de-criza-837030.html (in Romanian).

[9] http://www.hotnews.ro/stiri-politic-7350294-sebastian-vladescu-era-foarte-usormintim-continuare-mai-imprumutam-vreo-sase-luni.htm (in Romanian)

logical track, and iii) vocational track (see NASFA Romanian Educational System, 2011). Upon completion of high school, students take the school-leaving exam - the Baccalaureate exam (akin to the French Baccalauréat) - which is a nationwide standardized test mandatory for obtaining the certificate of graduation from secondary school. Importantly, passing the Baccalaureate exam is a strict requirement for pursuing further professional training or for enrolling in tertiary education,[10] as the student's average grade on this exam accounts for up to 100% of the university admission score, and is the main criterion for being granted exemption from tuition fees (in public universities). Thus, passing this national examination (with high grades) is very important.

The Baccalaureate consists of several standardized tests taken in oral (testing knowledge of Romanian and a foreign language) and written form (containing multiple choice, elaborate answers and essays in different subjects, depending on track). These are graded on a scale from 1 to 10, and to pass the exam, a student should obtain a minimum score of 5 on each test and a minimum overall average score of 6, while scores of 7 and above are usually regarded as competitive for admission in higher education, The tests are held in examination centers, to which more high schools from the same locality are randomly assigned. The organization of the exam in every center is the responsibility of the exam committee, which consists of a chairman (typically a university professor), one or two deputy-chairmen (typically public high school principals), a person specialized in IT management (for technical support), and a number of public school teachers whose duty is to monitor the exam.[11]

The format of the Baccalaureate has been standard for the last ten years with two oral exams and four written tests, which take place over the course of two weeks toward the end of June every year. A few changes to the exam format in 2010 make the overall pass rate less comparable to earlier years.[12] The most important changes were the exclusion of oral tests from the overall score and the elimination of the fourth written test, all with abnormal score distributions highly concentrated at the top marks.[13] The tests are standardized for all students ascribed to each education

[10]At the very least, the degree obtained by passing this exam offers a basic qualification with the potential to earn the student a better placement in the labor market.

[11]These teachers are unrelated to the subject under evaluation or to the students, and are randomly assigned in pairs of two in each classroom by the exam committee.

[12]No other changes in the educational system took place in the period 2007-2010.

[13]The oral exams were pushed ahead of the written ones, to February, and they were rendered irrelevant to the overall exam grade. Also, a new examination of digital competencies was added to the oral section of the exam, and one track-specific written test was eliminated. The assessment became qualitative, categorizing the students into: experienced, advanced or average users. Also, in 2007, 2008 and 2009, in preparation for the exam, the students had access to 100-300 published

profile and track. The one test that is unique to all students regardless of profile and track is the written Romanian language exam. This, together with the fact that the conditions for this test have remained very similar across years makes it an ideal basis for comparison of student outcomes on the exam.[14]

As stated before, in 2010, the wage cut news arrived on May 7th, three weeks before the end of the school year, during which the graduation ceremonies take place. Since the exam is set in June, this close timing between the unexpected news and the exam reduces the possibility that the wage cut would have changed the test outcomes via increased effort by students, parents or teachers. Still, in Section 5, we will perform some sensitivity analyses to rule out this channel.

2.3 The corruption environment

The endemic post-communist corruption in the public sector has become proverbial among Romanians: a 2003 World Bank Report about corruption in Romania reveals that more than 67% of the respondents alleged that all or almost all public officials in Romania are corrupt, while more than 50% of the respondents believed that bribery is part of the everyday life in Romania.[15] This is particularly true in the education and health systems, where up to 66% of the respondents confirmed that they were paying the so-called atentie (unofficial payments or bribes).[16] More than a quarter of the students interviewed in the 2003 World Bank Diagnostic Survey of Corruption in Romania admitted to have provided some unofficial payments during the previous year.

written exam models with full answers for each discipline, some of which would have become the actual tests. In 2010 the test would resemble, but not perfectly match the models. All in all, we expect these changes, if anything, to decrease the test scores.

[14]We also claim that for the Romanian written exam it is more difficult to cheat in class (as one possible confounder to corruption), since students need to develop ideas and write essay-like questions as part of the examination.

[15]A 2010 study on corruption in Romania shows about 80% of the respondents to agree that the Government and Central Institutions are corrupted to a large and very large extent, a finding that is in linewith the idea that corruption has increased during the last years. www. agenda21.org.ro/download/%20Studiu%20perceptia%20cetatenilor%20asupra%20 coruptiei%20din%20institutiile%20publice.pdf (in Romanian).

[16]Paying the so-called atentie is very common. A World Bank Report on corruption in Romania confirms that up to 66% of the respondents have paid an atentie during a hospital stay, while 27% of the respondents have given atentii to vocational school (teachers), 25% to the primary school (teachers), 21% in the high-school system and 17% in the University Anderson et al., 2001. For education these are lower bounds: first, people do not like to admit they are bribing teachers, as may signal insufficient ability; second, these numbers are from survey questions to all households, regardless of the age of the household members and whether or not they have kids in school. A recent survey among university students reveals that about 72% of the students and 68% of the university teachers were involved in corrupted activities in relation to school (our calculations using the 2007 PEIS data, Gallup Romania).

Thus, one notable feature of the Romanian public schools that favors the propagation of corruption is the existence of a habitualized system of informal payments. These range from more innocuous forms such as the imposition of funds collected for covering school and classroom material expenses (fondul scolii/clasei) all the way to gifts demanded by teachers in exchange for favors such as not failing the students or inflating their grades.[17] Overall, the frequency of such exchanges over the entire course of school years sustains a dense clientelistic network. Among the most commonly invoked causes for dysfunctions in the public education system are: i) the poor remuneration of teachers in the public sector[18] and ii) the high-stakes of the high-school exit exam, particularly starting with the year 2002 when increasing numbers of universities included the Baccalaureate exam score as part of the admission process.

There is an overall consensus among the Romanian public that the Baccalaureate passing rates (anchored around 80%) and the underlying grades are artificially inflated through corruption. This "performance" is in complete opposition to international tests (PISA), where Romanian students earn among the lowest scores.[19] This inconsistency is shown in Fig. 1 where we show the 2009 upper secondary graduation rates and the PISA test scores for 16 European countries. Interestingly, Romania lies in the first part of the distribution of the upper secondary graduation rates (Fig. 1A), while, at the same time, is the European country with the lowest PISA scores (in Fig. 1B we show the reading test, but similar ranking is obtained for the mathematics and the science tests). Moreover, the introduction of video surveillance in 2011 coincided with a drop in average pass rates to a staggering 44%, further confirming that the exam had for years been corrupt.

The 2010 exam earned a special reputation and the suggestive title "The Xeroxed Baccalaureate" after a large number of cases of corruption at the exam (150 defendants compared to essentially none previously) caused a media storm.[20] Without precedent, many teachers and school principals were investigated by the Romanian National Anticorruption Directorate (DNA), in connection with the 2010 Baccalau-

[17]Center for Education (CEDU, 2006), Administration and practices lacking integrity in schools. http://www.cedu.ro/files/research/Administrare%20si%20practici% 20lipsite%20de%20integritate%20in%20scoala%20%20raport%20de%20cercetare.pdf (in Romanian).

[18]In Romania, similar to other transition countries, wages of the educational staff in the public sector are highly centralized and there is little variation across teachers. While there are no official statistics, it is the case that public teachers earn, on average, up to two times less than their private counterparts.

[19]See, for example, the 2009 PISA Executive Report: http://www.oecd.org/dataoecd/34/60/46619703.pdf and the 2009 OECD Report Education at a Glance http://www.oecd.org/dataoecd/41/25/43636332.pdf.

[20]http://www.pna.ro/faces/index.xhtml.

reate exam for having taken large amounts of money from students to help them pass or to raise their grades.[21] In particular, the school personnel was accused of arranging with committee members for selected papers of these students to be graded higher, partly changed or entirely replaced (Xeroxed) with correct answers. Some of these cases went to court and were finalized in 2011 and 2012 with prison sentences.[22] This evidence suggests that the exam in 2010 was characterized by an unusually high level of corrupt activity, which we explain through the additional incentives for fraud borne by the unexpected wage cut.

2.3.1 Possible mechanisms of corruption

As explained above, in Romania gift-giving and informal payments are very common, particularly in public institutions (see CEDU Report, 2006; Corruption in Public Institutions, 2010).[23][24] At the Baccalaureate, the unofficial payments resulting in grade inflation can be, broadly, summarized as follows:

a) Collective bribes the so-called "protocols" are informal but commonly accepted funds (money) collected on various occasions, among which is graduation.[25] The graduating students, shortly before the end of the school year, collect these contributions to "organize" the Baccalaureate exam, which are in fact used to "grease the wheels " such that the invigilators and other committee members turn a blind eye to cheating in the exam rooms (copy aids, talk among students, etc.). However, in-class cheating and thus, implicitly the protocol, is feasible for both public and private students, who are randomly and anonymously mixed in exam rooms, under the same surveillance. We will rule out differential in-class cheating in Section 5.

b) Individual bribes some students (individually or in small groups) may give extra bribes for extra favors. These favors come in many forms: distributing of correct solutions during the exam for the contributing students, bribing the evaluators

[21] http://www.ziare.com/stiri/arestare/directori-de-liceu-arestati-pentru-fraude-labacalaureat-1029179;
http://www.adevarul.ro/scoala_educatie/liceu/150-000_de_leifraudarecord_la_Bacalaureat_0_292771226.

[22] www.desteptarea.ro/zeci-de-condamnari-in-dosarul-spaga-la-bac.html (in Romanian).

[23] Center for Education (CEDU, 2006), "Administration and practices lacking integrity in schools ", see footnote 17. The National Agency for Public Workers (2010), "Perceptions about corruption in public institutions", http://www.anfp.gov.ro/DocumenteEditor/ Upload/proiecte%20in%20derulare/Studiu%20perceptie%20coruptie%20sept%202010.pdf.

[24] Hallak and Poisson (2007) provide a comprehensive taxonomy of corruption in education. The forms of fraud tackled in this paper are not restricted to the Romanian educational system. In Russia, Ukraine and Uzbekistan (Silova and Bray, 2006) the sale of grades is common, while in India the high school exam annual pass rates dropped from 61% to 17% in 1992, when police were stationed at the examinations centers (Kingdon and Muzammil, 2009). For more such illustrations see Lewis and Pettersson (2009: 45).

[25] See also Center for Education CEDU, 2006 (footnote 17).

to score selected papers higher, cooperating with the exam committee to single out the marked papers and improve them or completely replace (Xerox) them with correct ones before sending them to the evaluation centers.[26] In particular, using the already developed informal network at the high-school level, students use the teachers/school principals' channel to send their bribes to the exam committee members and/ or the evaluators for higher grades. Although the composition of the exam committees is made public only 48 h before the exam, the chairman and the IT staff are known months in advance. Note that the school principals typically have a very dense web of connections, having been randomly allocated to be part of the exam committees formed around the Baccalaureate in different years.

The individual bribes are somewhat more relaxed for the public students given the well-established informal networks in public schools.[27] However, the existence of corruption in private high schools cannot be ruled out but, as private school principals are not in exam committees, the chain of events necessary for a bribe from a student to result in higher exam scores is less likely to be fulfilled for private school students. Thus, we ground our identification strategy in the conjunction of this form of corruption with this differentiation between public and private schools' access to a corrupted network.[28]

3 Data and descriptive statistics

3.1 The data set

In our empirical exercise we use three main sources of data. Firstly, we use administrative data from 2007 to 2010, essentially covering the universe of students

[26]It was actually this form of bribe that led to the court cases in 2010 mentioned above. The 2010 Report of Activity of the National Anticorruption Court enumerates the investigated crimes at the 2010 Baccalaureate: bribe giving and taking; influence peddling; stealing, destruction and falsification of official documents, all involving large amounts of money. Individual bribes amounted to 350 Euro for passing one written test and 500 Euro for passing the overall exam. The total prejudice was at least 150,000 Euro. We do not have information about the number of high school students involved in individual bribing, but in the PEISGallup 2006 data, 55% of the university students admitted to have been paid "gifts " to get higher exam grades (admittedly, these are low stake-exams).

[27]Note that there is a cost associated with engaging in corrupt activities for educators the risk of getting caught and losing future earnings. Although no official sources detail on the monitoring and detection process, the 2010 Report of Activity of the National Anticorruption Court reveals that most cases of corruption at the exam have been detected as a consequence of reporting of the crime by some party involved in the corrupt deal (usually students). This gives a good indication that the larger the portfolio of clients a public educator serves, not only the larger is the private benefit, but also the higher is his risk of getting caught.

[28]In our sensitivity analysis we attempt to isolate the collective bribe channel from the individual bribes by controlling for exam center.

enrolled in the Baccalaureate exam, with individual information about their gender, school, their personal specialization track (theoretical/general, technological or vocational), whether the student passed the exam and the scores on each exam. From these scores we will construct our outcomes of interest. We also know whether the student was present at the exam or expelled from the exam room due to in-class cheating.[29] Secondly, we complement the data above with a measure on the students' poverty status, using individual information on the students eligible for the Money for High School (MHS) program of financial assistance for high school students with a monthly income per family member below 180 RON (about 53 USD). The 2007-2010 data provided by the Ministry of Education covers information on all the eligible students' school in every year of application.[30]

Finally, our third source of data is the 2010 Study Performance in High School (SPHS) data, collected by Statistics Romania. The SPHS records information on a broad set of high school characteristics for all high schools in the country: the high school name and a unique identification code; the address of the school (locality and county); the type of school (whether private or public); and detailed information about the number of students by gender and ethnicity, the number of teachers and school principals by gender, type of employment contract, and their age structure. We can thus match these data with the administrative students' records at the final exam by the school's unique identification code to construct our working sample. The key information for our empirical strategy is whether the student comes from a private or a public school. We only consider counties that have both private and public schools (19 out of a total of 42 counties). Thus, for the main analysis we rely on an unbalanced panel of between 824 and 850 schools for each academic year (127,500 students on average per academic year); among these approximately 6% are private schools (up to 5000 students per academic year).[31]

3.2 Descriptive statistics

Summary statistics for our main variables of interest, separately for 2007 through 2010 are found in Table 1. For our working sample, about 26.5% are theoretical or

[29] With our data, we only observe students that have been registered for the Baccalaureate.

[30] In particular, an applicant was eligible if he had a gross monthly income per family member not larger than 180 RON in the previous three months before applying. For the years 2007-2010 all students that were eligible and applied have received the scholarship. For more information about this program see Borcan et al. (2014).

[31] Our main results when using the entire sample are overall similar to those in the main analysis but less precisely estimated. Additionally, we will show some results at the examination center for all centers with at least one private school and where the share of private students is about 25%.

general schools, around 8% are vocational schools, and the rest of around 66% are technological or mixed schools.

We show descriptive statistics for exam scores and pass rates for the Romanian written exam at the school level, where we have weighted each school by the number of students taking the tests in the exam. Table 1 shows an increase in the average grade at the written Romanian test in 2010 relative to previous years, particularly 2009 and 2007. This test is directly comparable across years as its format has remain similar relative to previous years and all students, regardless of their profile, track or ethnicity, need to pass this standardized exam. This makes it an ideal basis for comparison of student outcomes across years. Thus the school-level average grades for the written Romanian exam and the share of students (at the school level) passing the written Romanian exam are our main outcomes of interest.

Finally, it is important to note that private and public schools differ in the levels of our key outcomes. Private schools consistently exhibit average passing rates and average Romanian grades below those of public schools. This indicates an overall lower performance of private schools, related to the selection of lower achieving students into private high schools in the 9th grade, a common occurrence in Romania.[32] This is why later in the paper we: 1) estimate the impact on exam scores between public and private school students in 2010, relative to previous years, controlling for pre-treatment differences in exam scores for previous years, county fixed effects and county-specific time trends, and school fixed effects, and 2) conduct estimation on a matched sample of public and private schools, with similar levels and trends in exam scores, and on type of track (and on other characteristics), prior to the wage cut in 2010.

4 Estimation strategy and baseline results

4.1 Identification strategy

We attempt to understand whether an income loss led to changes in corruption behavior, measured through a change in exam outcomes. Specifically, the policy we evaluate is the May 7th, 2010 unexpected wage cut for all public sector employees, affecting more than 90% of the Romanian education staff. The intuition is as follows. Before the 2010 exam, we assume exam outcomes to be inflated, for both public

[32]This is true on average, as a small number of private high schools select and train top students. For a description of the selection of Romanian students into the 9th grade see Pop-Eleches and Urquiola (2011).

and private schools.[33] Additionally, it is probably reasonable to assume that the incentives and level of corruption intensity for private schools should stay constant.[34] As we have argued before, a substantial wage loss for the public school staff has, ex-ante, unclear implications for corruption: on the one hand, teachers may attempt to compensate for their forgone income by increasing the prevalence of bribing and corruption; at the same time, an income loss may prompt teachers to refrain from corruption because the need to keep their job along with future bribe opportunities becomes more salient.

Our main empirical strategy to assess the impact of a change in corruption incentives caused by an unexpected wage cut is a simple difference-in-difference (DD) specification. In particular, we will compare school-level exam outcomes for the public and private schools in 2010 relative to earlier years. Because private and public students are alphabetically mixed in exams rooms and subject to the same examinations, the private school students constitute a natural control group. If the wage cut has caused an increase in corrupt behavior of the educators in the public schools (through bribes, as discussed in Section 2), we expect to see an increase in exam scores in public school, relative to private schools.

Our baseline specification is the following equation:

$$y_{sct} = \alpha + \beta \cdot Public_s \cdot yr2010_t + \delta' \cdot X_{sct} + \varphi_t + \theta_s + \theta_c \cdot t + \epsilon_{sct} \qquad (1)$$

where s indexes a school in county c at year t. y_{sct} is one of our two main outcomes of interest: 1) the school-level average grade for the written Romanian language exam and 2) the school-level share of students passing the written Romanian language exam; $Public_s$ is an indicator that equals 1 if school s is public and 0 if it is private; $yr2010_t$ is an indicator that equals 1 if it is for the 2010 final exam and 0 if it is for any other year; X_{sct} includes the share of poor students and the share of male students in school s in year t; φ_t represent 3 year indicators; θ_s includes schools indicators and $\theta_s \cdot t$ are county-specific yearly trends. Our main coefficient of interest

[33] A natural test of the validity of this assumption is actually the Baccalaureate exam in 2011. Following different anti-cheating initiatives and threats (for example, installing video cameras in schools during the exam, threatening the staff with dismissal), over half of the students taking the exam failed (see Borcan, Lindahl and Mitrut, 2014).

[34] While we assume that corruption in private schools did not change after the 2010 wage cut announcement, one may argue that this policy impacted indirectly the private teachers' labor market, making them potentially less inclined to take bribes for fear of getting fired. Thus, this could have generated lower exam scores in private schools, due to less corruptible private school teachers. We hereby work under the assumption that corruption (if any) in private schools stays constant between 2010 and previous years, or that the alternative labor market situations equally affected for private and public school teachers. We will also run several sensitivity analyses in Section 5.

is β, the DD-estimand, which measures the change in outcomes in 2010, after the abrupt wage cut, relative to previous years, for public relative to private schools. We weight all regressions with the number of (per school) students taking the exam.[35] In the regressions we cluster the standard errors at the municipality level, since an important part of schools' financing is decided by the municipal administrations (resulting in 254 clusters).

By including school fixed effects, we are able to control for unobservable time-invariant school characteristics. In alternative specifications, we replace θ_s with θ_c, which includes 19 county indicators.[36] In this case we expand the list of controls to include a separate $Public_s$ indicator and an additional vector X_{sc} containing two indicators for the track of the school: theoretical and technological (the base is vocational).[37]

We account for possible changes in the composition of students at the school level by including controls for the students' gender and poverty status, which, if correlated with the events in 2010, may otherwise alter estimates of the β coefficient.

A necessary condition for an estimate of to capture the effect of a sizable wage cut on corrupted exam scores is that the interaction term $Public_s \cdot yr2010_t$ is uncorrelated with the error term in Eq. (1). Our key assumption in order to get consistent estimates of β in Eq. (1) is therefore that, in the absence of the wage cut, we would not observe any difference in the change in the exam scores between public and private schools in 2010 relative to earlier years (the parallel trend assumption). To investigate the plausibility of this assumption we will estimate a less restrictive

[35]The estimates are very similar if we estimate un-weighted regressions.

[36]The difficulty in estimating correct standard errors in DD models where a policy changes only for a small number of groups is discussed in Conley and Taber (2011). Their argument is that unless the number of treated groups is large, standard methods for inference are inappropriate. In this study we have treated and control units (public and private schools) represented in all the 19 counties. Hence, if we see geographical clusters (for instance counties) as units of treatment, their critique is not relevant for this study. Of course, one can also think of their critique as being relevant for non-geographical dimensions (such as all public schools being one unit of treatment and all private schools being one unit of control). However, although we discuss this issue more in detail in Section 5.2, we think that it is unlikely that there are important specific shocks (unrelated to the wage cut) that affect public schools but not private schools. This assertion gets additional support from the facts that a) we get similar sized standard errors whether or not we cluster the standard errors at the school, the locality or at the county level, something which can be reconciled with the Conley & Taber argument being valid here only in the unlikely case of shocks hitting public and private schools differently between but not within counties, and b) we do not find that exam scores evolve differently in public and private schools prior to the wage cut, hence supporting the claim that observed differences in outcomes between public and private schools are not due to group-specific shocks.

[37]We do not include other school related characteristics since we only have this information for the year 2010. We will perform some tests using this information and show these results in Appendix B.

version of Eq. (1) and add two interaction terms, the public and yearly indicators for 2008 and 2009, to the baseline model.

We also try to address concerns related to other changes that may have affected private and public schools differently in 2010 relative to previous years and that could confound the estimated effect β. Firstly, because private schools are, on average, different than public schools along other dimensions, we conduct additional estimations where, for a subsample of schools, we are able to control for student performance measured prior to high-school admittance and we also use matched samples of private and public schools to check our main results. Secondly, because a differential change in exam scores between public and private school students might occur for reasons unrelated to corruption we, in Section 5.2, discuss and investigate a number of additional potential threats to the interpretation of our results.

4.2 Results from baseline estimations

In this section, we present the basic findings from estimating Eq. (1). Table 2 displays the DD estimation results from our chosen baseline specification featuring the average grade (Panel A) and the pass rate on the written Romanian exam(Panel B) as our main outcomes of interest. Columns (1) and (2) present the DD estimates unconditional on pre-treatment dynamics, while columns (3)-(4) display the estimated coefficients from the fully-interacted model. Columns (1)-(3) include school indicators, whereas column (4) presents the estimates from the model with county fixed effects.[38] All columns include year-indicators and county indicators interacted with a time trend.

We note already in column (1) that for both outcomes, the DD estimate of the wage cut is positive and statistically significant. When we add controls for school student composition, the DD-estimate increases slightly. Focusing on the DD-estimates reported in column (2), we find that the average grade score has increased with 0.27 points and the average pass rate has increased with 3.3 percentage points (a 3.7% increase) for students in public schools relative to private schools, in 2010 compared to previous years. Interpreting the estimate for the average grade score in terms of effect sizes, the size of the estimated effect is equivalent to a 0.26 S.D. increase in scores on the Romanian exam (amounting to a 3.9% increase).[39]

[38] In all columns we use the same unbalanced panel. Estimates for the sample of schools with data in all years (balanced panel) are available upon request. They are similar in magnitude to the results from the specifications with school fixed effects, but slightly less precisely estimated.

[39] The calculation of the effect size is based on the school-level distribution in exam outcomes reported in Table 1. If we instead use the student-level distribution for the Romanian written

Identifying a causal effect of the wage cut on corruption through the DD estimate hinges crucially on the parallel trend assumption. If exam scores would have increased more in public schools than in private schools, even in the absence of the wage cut, our DD estimates would be too high. Column (3) in Table 2 presents estimates from regressions which allow for a flexible form of pre-treatment dynamics by including the publicyear interactions for 2008 and 2009 (the omitted year is 2007). For neither outcome are the estimates for the 2009 and 2008 year-specific public indicators significantly different from zero.[40] This suggests that public and private schools do not differ significantly in their evolution of exam scores during the pre-treatment years, validating the parallel trend assumption.[41] These results therefore lend support to our hypothesis that the change in grades in public schools relative to private schools in 2010 relative to previous years not driven by different trends in the performance of the two types of schools, but rather plausibly related to the wage cut through the increased incidence of corruption.[42]

Lastly, we note that the estimates in column (4) where we have replaced the school indicators with county indicators generate larger estimates and similar standard errors relative to the first three columns.

5 Sensitivity analysis and alternative explanations

Because our identification strategy is based on observational data, it deviates from the ideal setting of a randomized experiment. To consolidate the credibility of our

exam (where the standard deviation is 1.674 in 2010s) we get the estimated effect to be equivalent to a 0.16 S.D. increase in scores on the Romanian exam.

[40]When we add more structure to the pre-treatment dynamics and replace $Public_s \cdot yr2009_t$ and $Public_s \cdot yr2008_t$ by the interaction of $Public_s$ with a linear time trend, the estimate for $Public_s \cdot yr2010_t$ decreases somewhat (to 0.251) and is statistically insignificant, but still shows a large 2010 jump from what would be expected from the estimated trend which indicates an increase by 0.008 (for public relative to private schools) for each year.

[41]Note that the estimates for the $Public_s \cdot yr2008_t$ interaction are large relative to 2007 and 2009 for both outcomes. However: (i) the 2010 DD estimate is significant and is the largest in magnitude, whereas the estimates for the pre-treatment interaction terms are always insignificant; (ii) the estimates for the $Public_s \cdot yr2010_t$ interaction term, are similar in models with and without pre-treatment dynamics.

[42]The results shown in Table 2 are based on students in all high-school tracks. The theoretical tracks are generally the first choice for skilled students in the admission to secondary education. In order to investigate the potentially differential impact across school tracks, we also performed estimations separately, for theoretical and non-theoretical schools and we find a similar-sized contribution to the wage cut effect, even though the effects for theoretical schools are imprecisely estimated. Finally, as already explained, we only focus on the written Romanian exam because this is a standard exam for all children, regardless of the track and sub-track. Other exams, more specific for each track and sub-track (e.g., some theoretical track students would take Mathematics difficulty 1 while others Mathematics difficulty 2; some would choose between Physics and Chemistry), are more difficult to analyze.

findings, we perform some additional analyses where we attempt to gauge the sensitivity of our results to using private schools as the control group, and to eliminate some confounding factors and to build a compelling case against alternative behavioral responses to the wage-cut news.

5.1 Are the treatment and control groups similar enough?

One could object that private schools are not an ideal control group to public schools and there is always a possibility that the controls included in the specifications underlying the results above are insufficient to adjust for such differences. Most importantly, the average exam scores and pass rates differ significantly between public and private schools. Additionally, although probably of less importance, the control group (6% of the sample) is notably smaller than the treated group. To check if these issues are likely to bias our baseline estimates we perform different sensitivity checks.

5.1.1 Estimations controlling for student performance prior to high-school admission

Our first exercise attempts to rule out the possibility that the DD estimate is driven by differential student intake in the public and private schools in the 2010 cohort, and to tease out the effect of student composition from the general public-private score gap. To do this, we make use of additional data available from the Ministry of Education covering the high-school students' gymnasium (5th-8th grade) average graduation grade (i.e., the average of all scores from grades 5 to 8), which we refer to as student "ability " below.[43] Unfortunately, this information is only available for the students that completed gymnasium in 2004-2006 and were admitted to high schools, with standard admission procedures, in 2008-2010.[44] Hence we rely on a smaller (and potentially slightly different) sample than for the baseline estimates.

[43]This proxy should capture students' true ability reasonably well. Firstly, because this measure captures all grades in all subjects during the four years of middle school (gymnasium). Secondly, there are less incentives to inflate this grade through corruption as all students in Romania are admitted into high-school, so this is not a high-stake grade as compared to e.g., the Baccalaureate (for more details about the centralized transition between middle and high school, eighth to ninth grade, in Romania see Pop-Eleches and Urquiola, 2011). Furthermore, as shown before, following the 2011 anti-cheating initiatives and threats (installing video cameras in schools during the exam, threatening the staff with dismissal), the passing rate for the Baccalaureate failed with more than 45% in 2011 relative to before (see Borcan et al., 2014), whereas the drop was much smaller (about 17%) for the 8th grade standardized evaluation.

[44]Moreover, we do not have this information for around 60 schools, because the gymnasium performance is only made public by high schools that organize a standard admission process, whereas some vocational and private schools have independent admission procedures.

In Table 3 we show results for the average grade (Panel A) and the pass rate on the written Romanian exam (Panel B) from estimating Eq. (1) for the years 2008-2010. We start in column (1) by replicating the baseline estimates from Table 2 (the second specification), in column (2) we include controls for the average initial ability of the students in each high-school and for the share of students per school for whom we have information on ability, while in column (3) we add interactions with the 2010 year indicator.

First we note that, despite some potential change in the composition of schools for the years 2008-2010, our results in column (1) are comparable with those in our main Table 2. Next, we learn that controlling for student ability has little effect on the size of the DD estimate using the pass rate as the outcome, but that the DD estimate using the average grade as an outcome which now becomes smaller and insignificant. However, the specification underlying the estimates in column (2) is quite restrictive as it assumes that student ability has the same impact on Baccalaureate outcomes in all years. This is especially problematic since there are reasons to expect that the importance of ability for later outcomes differs depending on how corrupt these outcomes are. Therefore, in column (3) we also interact student ability with the 2010 indicator. The result is then that the DD estimate is statistically significant and similar in size as in column (1). This reassures us that the wage cut effect is independent of the initial ability and of the interaction of the ability level with the exam structural changes in 2010 (i.e., a potentially more favorable response of higher-ability students to the increase in exam difficulty).[45] This means that we are able to pin down what typically distinguishes public and private schools and to ensure that the DD estimate is not driven by any difference in student composition. Finally, we note that the coefficient of the interaction between average school ability and the 2010 indicator is negative and significant, while ability itself has a large positive coefficient. This means that in 2010, ability has a lower impact on exam outcomes, while being in a public school in the same year, conditional on ability, has a larger impact on exam outcomes than in previous years. Results in Panel B for the pass rate outcome have a similar pattern as those shown above for the average written Romanian. Overall, these findings seem to support our hypothesis that the DD estimate captures an increase in corruption in public relative to private schools.

[45] Additionally, from county fixed effect estimations not reported here we see that controlling for average ability seems to reduce the public-private gap before 2010.

5.1.2 Evidence from matched public and private schools

With our next exercise we address the potential concerns that the public schools included in the treatment group might not have comparable private schools. Because we do not have enough pre-treatment school level information we attempt to match public to private schools using exam scores in 2007-2009 (to capture both the levels and the trend), student composition in terms of share of poor students and the gender split, track, and county.[46] As we match on pre-treatment outcomes, our strategy here is to simply compare the matched public and private school outcomes for the year 2010. Results are reported in Table 4. In column (1) we show the resulting matching estimates without any controls, while in columns (2) and (3) we add the student composition controls and the theoretical track indicator. The estimates in the first two columns are somewhat bigger than our baseline DD estimates reported in Table 2. When we add controls for the exam scores prior to 2010 (in column 3) and also county fixed effects (in column 4), we learn that the matching estimates decrease quite a lot. However, since the precision also increases, we still obtain statistically significant positive estimates for both outcomes when including the full set of controls. The magnitude of the estimates is also quite similar to our baseline DD–estimates, thus matching techniques are reassuring in what concerns our baseline parametric estimates.

5.1.3 Examination centers with both private and public schools

Finally, we also limit the sample to schools in examination centers where there was at least one private school and estimate regressions similar to our baseline.[47] These results are reported in Appendix (Table A1) and are in line with our main results in Table 2. We also include examination center indicators to control for unobservables at the center level (location, size related to the number of schools and, implicitly, to the collective bribe). This could potentially rule out collective bribe for schools, some of which are assigned to the same exam center. That estimates do not change with the inclusion of examination fixed effects, suggests that individual bribes are the main mechanism for why we find the wage cut to increase the corrupted exam scores.

[46] We use the nearest neighborhood and 1-to-1 matching (without replacement) to match a public to each private school. Our matching is done using the psmatch2 command in STATA (Leuven and Sianesi, 2003).

[47] For this exercise, we have identified on a case–by–case basis the school composition of centers to which at least one private school was assigned each year. The percentage of private school students in this sample is about 25%.

5.2 Alternative explanations

Because, like most of the previous work, we do not have a direct measure of bribes, in this section we discuss some potential confounding explanations that could bias our main estimates. In particular, students, parents, teachers, proctors and/or exam committee members may respond to the wage cut announcement in ways that are actually unrelated to corruption, but that can nevertheless impact scores on the exam taken in June. Another concern is related to other possible exogenous macro-level shocks incurred in 2010 (or before) that may differentially affect public and private school students or teachers, and that, in turn, would impact differently the exam scores. In addition, the Baccalaureate exam changed in 2010. Although we cannot provide fully conclusive evidence, in what follows we attempt to discuss all these alternative explanations that may bias/confound our main results.

5.2.1 Exploring the regional variation in corruption

One ideal setting to test these concerns would be to estimate Eq. (1), for the same time period, in a setting where there is no corruption in education, but where circumstances are otherwise identical. While the nature of the policy we analyze precludes us from finding and using such a setting, we can still use the variation in corruption at the county level in Romania. In Fig. 2 we show the county-level variation in corruption as proxied by the frequency of payment of bribes and gifts in the public education system. In particular, we use the Life in Transition Survey (2010) and aggregate the scores assigned to responses to the question "In your opinion, how often do people like you have to make unofficial payments or gifts in these situations?", considering only the situations regarding the receipt of public education.[48] Using these aggregate scores, we divide counties into more and less corrupt if they situate above/below the median corruption. Next, we estimate our model separately for most and the least corrupted counties in an attempt to check whether the wage cut impact is differential across counties. If other exogenous shocks (e.g., macro-level shocks) or other responses (e.g., change of effort or cheating not related to corruption) had a similar impact across all counties, significant estimates exclusively in the more corrupted counties would support the corruption channel.

[48]The "Life in Transition Survey, After the Crisis" (LiTS II, 2010), was the second public attitudes survey conducted jointly by the European Bank for Reconstruction and Development and the World Bank. It surveyed 39,000 households in 34 countries, including Romania. The goal was to assess "public attitudes, well-being and the impacts of economic and political change " (www.ebrd.com), particularly those brought by the financial crisis. The survey is nationally representative, conducted face-to-face on samples of randomly chosen 1000 households from each country.

In Table 5, Panel A for the average grade at the written Romanian exam and Panel B for the share of students passing the written Romanian exam, we find that our positive interaction effects are driven by effects in the most corrupted counties, while the estimates in the least corrupted counties are much smaller and never statistically significant.[49] The challenge with this exercise is that corruption may be correlated with factors that may have also affected the performance of the students differently. Indeed, investigating other county level characteristics reveals not only that richer counties (higher GDP, less poverty, less unemployment) tend to be more corrupt, but also that less trust in justice and people is associated with more corruption. These other factors could lie behind the difference in performance across counties, so this split by corruption level cannot fully dismiss alternative explanations. We discuss these confounding stories in more detail below.

5.2.2 Some alternative explanations

We have discussed the overall economic context in 2010 (see Section 2) and particularly the fact that the international financial crisis was taken lightly in Romania. Indeed, the autumn 2008 and 2009 Eurobarometer showed that more than half of the Romanian respondents anticipated no change and some even expected an improvement in the general economic situation of the country.[50] As such, we believe that it is reasonable to assume that the austerity measures were not anticipated, neither in their unprecedented scope and magnitude, nor their timing. However, below we discuss how the overall macroeconomic situation or other mechanisms, like the changes in the exam structure, can affect proctors, evaluators, educators and/or students efforts' and which, in turn, can confound our main results.

i) *Proctors' effort* is a potential confounding story, particularly if the proctors decreased their effort following the wage cut or the overall economic situation, resulting in more students cheating during the 2010 exam compared to previous years. This may have a stronger effect, on average, on the public students, if they are more predisposed to cheating. To shed light on this issue, we employ our main strategy on a measure of the share of students caught cheating (in class) and expelled from the exam, from the total number of students taking the exam (at the school level).

[49] These results are robust when using alternative measures of county-level corruption. In particular, we constructed a proxy based on the share of people having an informal network, at the county level, based on a question from the 2007 Romanian Barometer of Public Opinion: "Is there anyone (i.e., informal network) that could "help" you solve (i.e., informally): issues in court/trials, medical problems, city hall, police, or issues related to the local authorities.". The results, available in online Appendix B (Table B2) are similar to those in Table 5.

[50] http://ec.europa.eu/public_opinion/cf/: "What are your expectations for the year to come with respect to the economic situation of your country (Romania)".

The interaction term between the public and the year indicators is never significant in Table 6, which seems to support that, indeed, what we measure is not a change in in-class cheating.

ii) *Evaluators' effort* may have also changed (as a result of the wage shock and/or the macro conditions), in that they may have potentially decreased effort when evaluating the exams. If this was the case, we expect this to be particularly relevant for the students who were on the verge of passing.[51] The minimum requirements for passing each test and the overall exam are 5 and 6, respectively. Therefore, if there were proportionally more public than private students with scores 5-6, a less stringent assessment in 2010 could favor the public students, conducing to the observed average difference in outcomes. Then, in 2010, we would expect, on average, more public than private students passing the written Romanian exam with scores 5-6. To check this channel we consider in Table 7 a new outcome the share of students, at the school level, that passed the written Romanian exam with scores above 5 and below 7. Indeed, the interaction term between the public and the year indicators is only significant in the first two columns, and if anything, it is negative (Panel A), dismissing the story about marginal improvement of public students' scores due to a change in evaluators' assessment effort. Interestingly, in Panel B of the same table we show that the only positive significant increase for the public students relative to their private peers, in 2010 relative to the previous years, is found at the upper tail of the score distribution, for scores of 7 and above (within this range of scores, students would be competitive enough for admission into higher education, so there are higher stakes from achieving these scores). In Fig. 3 we show, separately by year, the average school shares of students attaining scores in each one of six categories: below 5, 5-5.99, 6-6.99, 7-7.99, 8-8.99 and 9-10, for public and private schools, respectively. Public and private schools differ at competitive scores (7 and above) in 2010 relative to before (particularly 2009 and 2007): while for private schools there is a slight decrease in the average student share in this range, for public schools the average shares in segments 7-7.99 and 8-8.99 show an increase. For the scores 5-6.99, the average shares in private schools are quite stable across years, while they decrease somewhat for public schools in 2010, relative to before. Overall based on the trends shown in Fig. 3, both a decrease in the scores for the private schools and an increase in the scores for the public schools appear to have contributed to the differences observed in 2010. The figure therefore corroborates the results in Table

[51]The implicit assumption here is that students who fail to pass a test (with score 5) or the overall Baccalaureate (with score 6) are more likely to appeal and/or re-take the exam in August, implying more effort.

7. This is also partly supported by anecdotal evidence that the Xeroxed exams helped students to achieve competitive scores of 7 and above.[52]

iii) *Educators' effort* in teaching activities could be affected by the substantial wage cut and/or by the overall economic context. We dismiss the former channel because the courses were already finished at the time of the wage cut announcement. Moreover, if anything, a lower teacher wage would likely lead to lower student achievement, which would mean that we would underestimate our main effect estimates. However, the overall economic context may have changed the educators' effort. Particularly worrisome for our interpretation is whether the educators have increased effort during in-class teaching in the months before the wage cut, differentially in public and private schools.[53] Moreover, because our main effect seems to come from the upper part of the grade distribution, this would mean that only competitive students were affected by the possible change in educators' effort. Overall, this remains a channel that we cannot completely dismiss.

iv) *Students' effort* may have also been affected by the changes in 2010 not directly related to the wage cut policy. For instance, the marginal benefits of going to college may have changed as a result of the 2010 macroeconomic context which, in turn, may have changed the incentives of students to study for the Baccalaureate. If the returns of going to college increased relatively more for the public than for the private students, the former may have put additional effort in passing the exam.[54] Below we outline some conceivable "symptoms " of the change in students' incentives, which may confound our interpretation of the results:

[52]In the wake of corruption trials, student testimonies confirm that bribes were paid to ensure a score of 7 and above. Source: http://adevarul.ro/news/eveniment/dimitriebolintineanu-1_51d31f61c7b855ff56f42753/index.html (in Romanian).

[53]Alternatively, teachers may have reacted to the unstable economic situation by increasing the supply of private tutoring which would result in better outcomes for students. We have looked into the 2008-2010 Romanian Household Budget Survey and, albeit a very small sample, we find no change in the share of students taking private tutoring in 2010 vs. 2009 and 2008.

[54]Another reason for the students' effort to evolve differently between the public and private school students is if their parents are affected differently by the wage cut (if e.g., the public school students are more likely to have parents employed in the public sector). Even if this is the case, it is not obvious in what direction this would affect our estimates: parents affected by the wage cut might be more willing to pay bribes in order to avoid future university fees for their children or, lower incomes mean that there are less available resources to be spent on bribes. Because we are lacking data on the occupations of the parents, we are not able to investigate this issue. In addition to the issue about student effort, if, for example, students fear that the evaluators will be more demanding in 2010 as a behavioral reaction to the wage cut because both public and private students are graded by public teachers, their level of awareness should be the same. Thus, their incentives to invest in marginally more preparation, either individual or through potential private tutoring, should not differ. We have looked into the 2008-2010 Romanian Household Budget Survey and, albeit a very small sample, we find no change in the share of students taking private tuition in 2010 vs. 2009 and 2008.

a) Changes in *student selection* (either with respect to the share of exam takers or the students' background) related to changes in students' incentives: One concern is the differential evolution in the share of public and private students taking the Baccalaureate. Lack of data about the number of graduates prevents us from constructing an accurate measure of the dropout rates over time. Using a rather restrictive proxy, we see that the share of 12th graders enrolled in the final exam sustained a larger increase in private than in public schools.[55] Even if this were accurate, this would be unlikely to have happened on grounds of the wage cut (announced on May 7th 2010), since the exam registration period was December 2009. However, the dropout rates may be affected by the overall economic conditions. This would be a problem for our estimates if marginal students were of lower ability: we might suspect that exam scores could decrease more in private than in public schools in 2010 relative to before, partly because of changed composition of students. For lack of accurate dropout rates before 2010, we cannot control for the share of exam takers in the regressions, which would account for a variety of unobserved factors to do with motivation. Still, we can at least include in our regressions a proxy for family income as an additional control, as a way to partially control for students' selection. For instance parent's income may determine a change in the motivation on whether to invest in education and exert effort, particularly during an economic downturn. If, for instance, fewer low-income students take the exam, discouraged by the economic turmoil, then this would affect only the public schools, potentially generating the results we see. And similar arguments in the opposite direction could be made. Our strategy to deal with this issue is to control in all our regressions for the share of poor students among the students who take the exam in each school, a variable which we have yearly data for. Thus, we ensure that our results are not explained by the income composition.

[55] Our preferred proxy suggests that the share of exam takers has increased from 2007 to 2010: the shares are 0.85, 0.90, 0.89 and 0.91 in public schools and 0.68, 0.60, 0.68 and 0.81 in private schools, for the years 2007, 2008, 2009 and 2010, respectively. These numbers are calculated in the following way: We know how many students took the Baccalaureate exam, but do not know how many students are enrolled in high school and decide not to take the exam. Also, we know how many students in each school graduated from high school (graduation is based on course work during the four years of high school and is decided about one month before the Bac exam) in 2010, but not for earlier years (since we only have HS graduation data for 2010). The best we can do is to assume that high school graduation is constant over time, and use the number of graduates per school in 2010 as the denominator. This makes it possible to approximate the fraction taking the Baccalaureate (out of the total number of graduates) in each year for each school. Note that we are getting that the share of exam takers is above 100% for about 10% of the schools (we have then restricted these schools to have a share equal to one). Also note that, by construction, this preferred proxy is of lower quality in the years before 2010.

b) Changes in *students' effort* resulting from macroeconomic conditions. As described above, survey evidence indicates that in December 2009 Romanians anticipated no change or even expected an improvement in the general economic situation of the country. Despite this, we cannot exclude that the worsening of the macroeconomic climate may have affected the students' marginal benefits of going to college. If students in public and private schools differ along characteristics like ability and income, the marginal cost of effort may differ across schools. Students with a lower marginal cost of effort are likely to respond more to changes in the returns to education. The variation in the marginal cost of effort that comes from variation in income can, at least partially, be dealt with by controlling for the school share of poor students, our proxy for income. Another part of this variation can be explained by students' ability. Using the students' ability proxy, in Table 3 we control for differences in ability and their interaction with the year 2010 indicator (containing all general changes) which should capture partially the interaction between ability and the 2010 marginal costs of effort. The effect of being a public student in 2010 is very similar in regressions with and without ability controls (columns 1 and 3). However, this test cannot fully dismiss the interaction between the macroeconomic changes and other unobserved characteristics in public and private school students.

c) Changes in students' effort (and teachers' reactions) resulting from the 2010 changes in the Baccalaureate exam. As some subject tests were no longer included in the Baccalaureate exam in 2010, the Romanian written test became relatively more important. As a result, students may have put additional effort in studying for the Romanian written test. While we do control for students' ability and a family income proxy, this may still be a valid argument that may potentially confound our results.[56] However, without more detailed individual level data it remains difficult to show conclusive evidence to dismiss this channel.

v) Finally, we acknowledge that the deterioration of the country's economic situation may have generated an increase in corruption directly, not necessarily via the wage cut. This possibility cannot be dismissed, despite survey evidence indicating that most Romanians were optimistic about the overall economic situation of Romania prior to the austerity measures in May 2010. In particular, if public teachers perceived and resented the general economic deterioration more acutely than the private teachers, they may have been more tempted to resort to illicit in-

[56]In an attempt to dismiss the wage cut anticipation effects on students' parents' or teachers' efforts due to changes in the exam structure, we consider the no-stake oral Romanian exam held only in February 2010. For this year we compare public and private students' scores and we find that the significant positive gap between them disappears when we control for previous performance (scores 5th-8th grade). Despite this, the possibility remains that students simply reacted differently to the overall changes in exam format in 2010.

comes regardless of the pay cut. In this case, the estimates here still reflect a story of economic adversity and corruption, but causality runs from the general economic depression rather than reduced bureaucrats' compensation per se.

6 Conclusion

This study responds to the imperative call for diagnosing the causes of corruption, particularly those stemming from the financial incentives of civil servants. We exploit an unexpected wage cut of 25% incurred by the entire public sector in 2010, to investigate the causal relationship between wage loss and the intensity of corruption. We base our analysis in the educational system, which was largely affected by the reduction in wages. Using data from the national Romanian Baccalaureate exam, we employ difference-in-differences strategies and estimate the effect of the wage cut on exam outcomes in the public schools, in comparison with private schools which did not experience any wage shock. Our estimates show that the wage cut caused a disproportionate change in average grades and passing rates in public high schools relative to private ones between 2010 and previous years. We attribute the estimated increased difference in exam outcomes between public and private schools to an intensification of corrupt activity by public school staff that is related to the wage loss. Our conclusion is also supported by the fact that we find no significant effects of the public school indicator for the pre-treatment years, and a series of tests which rule out some confounding stories. However, we need to be cautious when interpreting the main results because, as emphasized in the previous section, there are several channels that may confound the interpretation of our main mechanism.

Our results provide a snapshot of the undesired impact the policies of budget contraction had on the illicit behavior of affected agents, which is of particular relevance in the context of the recent adoption of austerity measures by post-crisis financially distressed EU members. Such drastic types of reductions in public spending are particularly dangerous in vulnerable environments that are already predisposed to corruption.

References

[1] Abbink, K., 2002. Fair salaries and the moral costs of corruption. Working Paper 2002-05. University of Bonn, Germany.

[2] Anderson, J., Cosmaciuc, B., Dininio, P., Spector, B., Zoido-Lobaton, P., 2001. Diagnostic surveys of corruption in Romania. World Bank, retrieved from http://www1.worldbank. org/publicsector/anticorrupt/romenglish.pdf.

[3] Armantier, O., Boly, A., 2011. A controlled field experiment on corruption. Eur. Econ. Rev. 55, 10721082.

[4] Armantier, O., Boly, A., 2013. Framing of incentives and effort provision. Working Paper. Federal Reserve Bank of New York.

[5] Barr, A., Serra, D., 2009. The effects of externalities and framing on bribery in a petty corruption experiment. Exp. Econ. 12 (4), 488-503.

[6] Barr, A., Lindelow, M., Serneels, P., 2009. Corruption in public service delivery: an experimental analysis. J. Econ. Behav. Organ. 72 (1), 225-239.

[7] Becker, G., Stigler, G., 1974. Law enforcement, malfeasance, and the compensation of enforcers. J. Legal Stud. 3, 1-19.

[8] Bejenariu, S., Mitrut, A., 2012. Austerity measures and infant health. Lessons from an unexpected wage cut policy. Working Paper 2012:4. Uppsala University (February).

[9] Bertrand, M., Djankov, S., Hanna, R., Mullainathan, S., 2006. Obtaining a driving license in India: an experimental approach to studying corruption. Q. J. Econ. 122 (4), 1639-1676.

[10] Borcan, O., Lindahl, M., Mitrut, A. 2014. Fighting Corruption: Monitoring and Punishment. Effects of a National Campaign on High-Stake Exams, mimeo, Uppsala University.

[11] Conley, T.G., Taber, C.R., 2011. Inference with difference in differences with a small number of policy changes. Rev. Econ. Stat. 93 (1), 113-125.

[12] Di Tella, R., Schargrodsky, E., 2003. The role of wages and auditing during a crackdown on corruption in the city of Buenos Aires. J. Law Econ. 46 (1), 269-292.

[13] Duflo, E., Hanna, R., Ryan, S., 2014. Incentives works: getting teachers to come to school. Am. Econ. Rev. 102 (4), 1241-1278.

[14] Ferraz, C., Finan, F., Moreira, D.B., 2012. Corrupting learning: evidence from missing federal education funds in Brazil. J. Public Econ. 96.

[15] Frank, B., Schulze, G.G., 2000. Does economics make citizens corrupt? J. Econ. Behav. Organ. 43 (1), 101-113.

[16] Glewwe, P., Nauman, I., Kremer, M., 2010. Teacher incentives. Am. Econ. J. Appl. Econ. 2 (3), 205-227.

[17] Gorodnichenko, Y., Sabirianova Peter, K., 2007. Public sector pay and corruption:measuring bribery from micro data. J. Public Econ. 91 (56), 963-991.

[18] Hallak, J., Poisson, M., 2007. Corrupt Schools, Corrupt Universities: What Can Be Done? UNESCO, Institute for Educational Planning, Paris, France.

[19] Jacob, B., Levitt, S.D., 2003. Rotten apples: an investigation of the prevalence and predictors of teacher cheating. Q. J. Econ. 118 (3), 843-877.

[20] Kingdon, G., Muzammil, M., 2009. A political economy of education in India: the case of Uttar Pradesh. Oxf. Dev. Stud. Taylor and Francis J. 37 (2), 123-144.

[21] Leuven, E., Sianesi, B., 2003. psmatch2: Stata module to perform full Mahalanobis and propensity scorematching, common support graphing, and covariate imbalance testing. http://ideas.repec.org/c/boc/bocode/s432001.html (version April 2012).

[22] Lewis, M., Pettersson, G., 2009. Governance in education: raising performance. World Bank Human Development Network Working Paper, December.

[23] LiTS II Report, 2010. Life in Transition, After the Crisis. European Bank for Reconstruction and Development (EBRD) and the World Bank (WB).

[24] Muralidharan, K., Sundararaman, V., 2011. Teacher performance pay: experimental evidence from India. J. Polit. Econ. 119 (1), 39-77.

[25] Nichols, S.L., Berliner, D.C., 2007. Collateral Damage: How High Stakes Testing Corrupts America's Schools. Harvard Education Press, Cambridge, MA.

[26] Niehaus, P., Sukhtankar, S., 2010. Corruption dynamics: the golden goose effect. BREAD Working Paper 223.

[27] Olken, B., 2007. Monitoring corruption: evidence from a field experiment in Indonesia. J. Polit. Econ. 115 (2), 200-249 (April).

[28] Olken, B., Pande, R., 2012. Corruption in developing countries. NBER Working Paper Series. vol. w17398.

[29] Pop-Eleches, C., Urquiola, M., 2011. Going to a better school: effects and behavioral responses. NBER Working Papers 16886. National Bureau of Economic Research, Inc.

[30] Rauch, J.E., Evans, P.B., 2000. Bureaucratic structure and bureaucratic performance in less developed countries. J. Public Econ. 75 (1), 49-71.

[31] Reinikka, R., Svensson, J., 2004. Local capture: evidence from a central government transfer program in Uganda. Q. J. Econ. 119 (2), 678-688.

[32] Reinikka, R., Svensson, J., 2005. Fighting corruption to improve schooling: evidence froma newspaper campaign in Uganda. J. Eur. Econ. Assoc. 3, 259-267.

[33] Schulze, G.G., Frank, B., 2003. Deterrence versus intrinsic motivation: experimental evidence on the determinants of corruptibility. Econ. Gov. 4 (2), 143-160.

[34] Shapiro, C., Stiglitz, J.E., 1984. Equilibrium unemployment as a worker discipline device. Am. Econ. Rev. 74 (3), 433444 (June).

[35] Shleifer, A., Vishny, R., 1993. Corruption. Q. J. Econ. 108 (3), 599-617.

[36] Silova, I., Bray, M. (Eds.), 2006. Education in the Hidden Market Place: Monitoring of Private Tutoring. Open Society Institute, New York.

[37] Van Rijckeghem, C., Weder, B., 2001. Bureaucratic corruption and the rate of temptation: dowages in the civil service affect corruption, and by how much? J. Dev. Econ. 65 (2), 307-331.

Figures and Tables

Figure 1: Upper secondary graduation and PISA tests, country ranking

Figure 1A

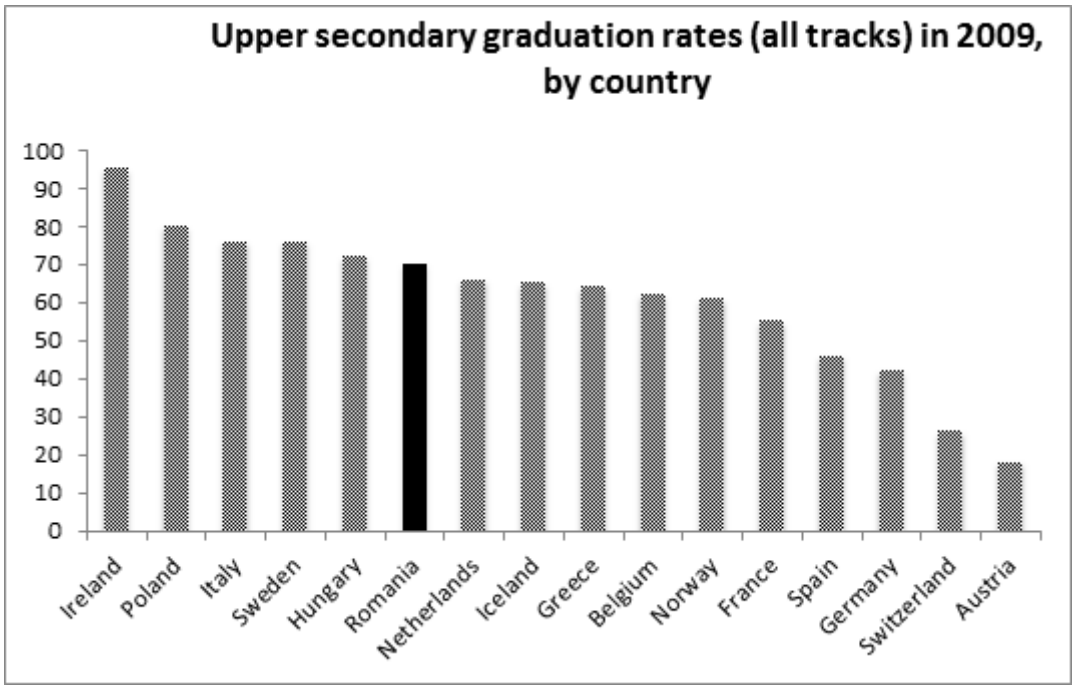
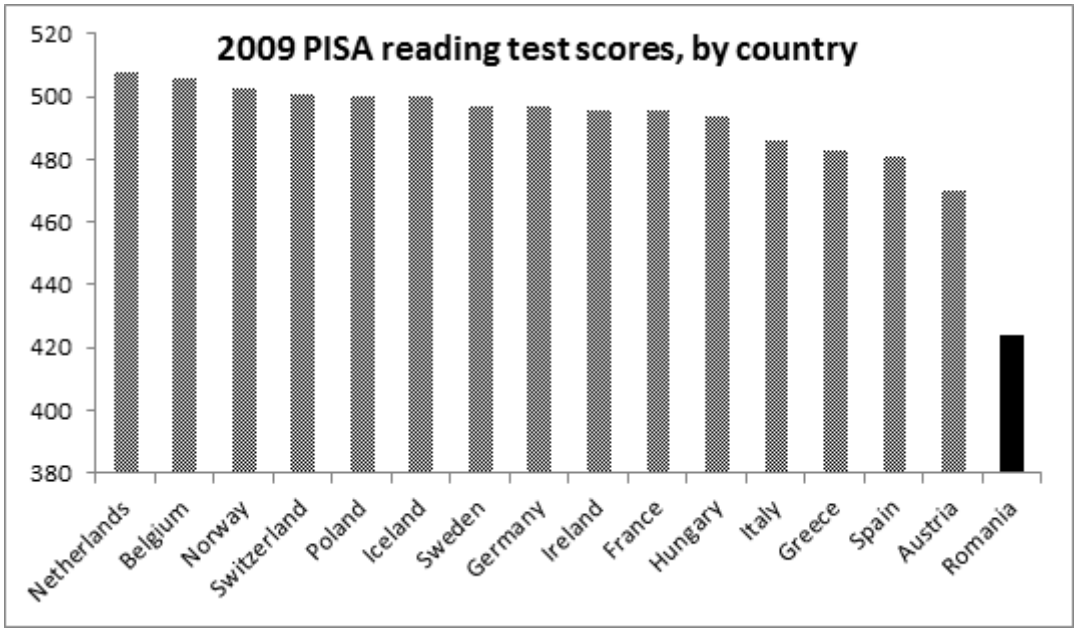
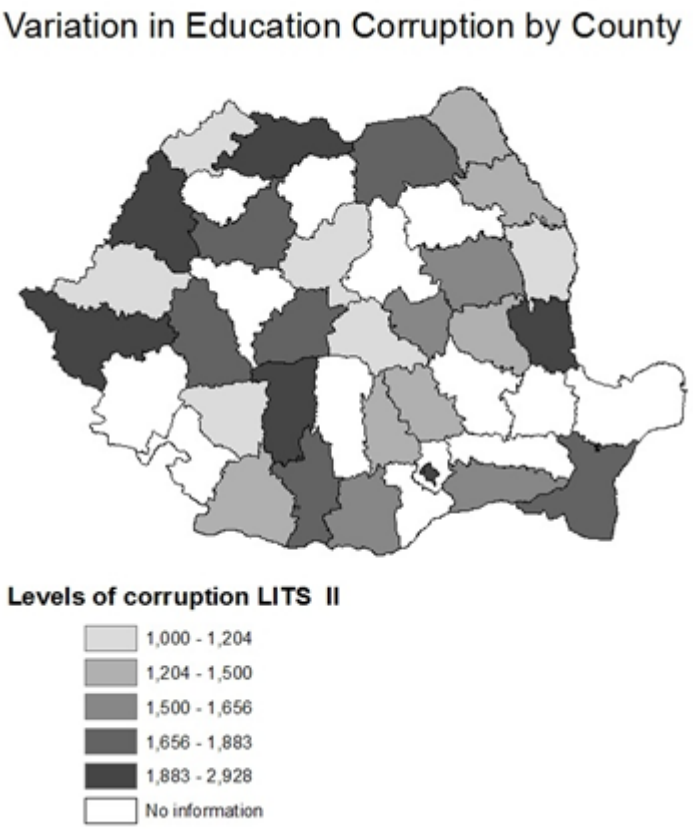


Figure 1B



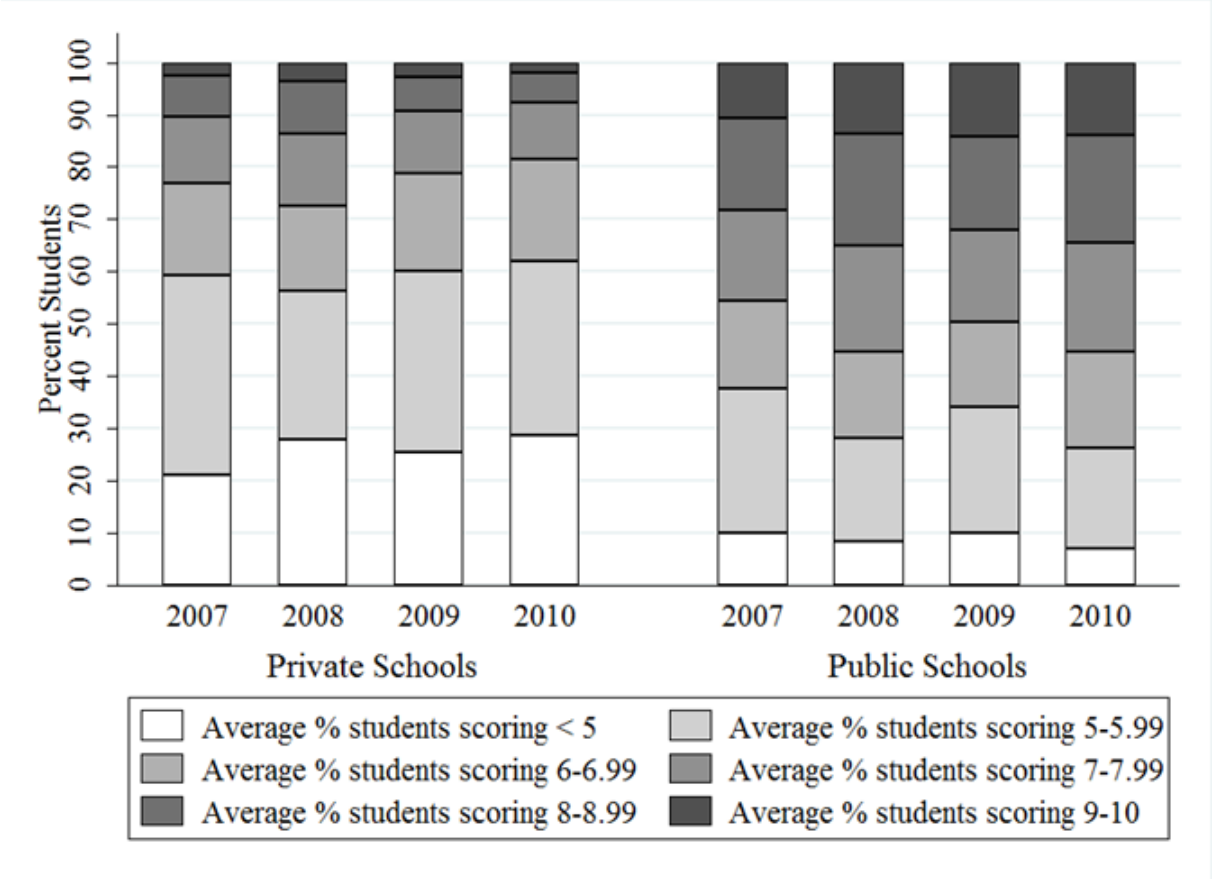
Notes: Our calculations using the UNESCO Institute for Statistics data from 2009 (available at http://www.uis.unesco.org/Education) and the 2009 PISA reading test scores (available at http://www.oecd.org/pisa/pisaproducts/pisa2009keyfindings.htm). Please note that: 1) we have used all European countries for which we have both information on upper secondary graduation rates and PISA 2009 tests; 2) Romania scores last and similar figure (1B) would have been obtained if using the mathematics or science tests scores.

Figure 2



Notes: Source: our calculations using the 2010 Life in Transition Surveys. We use the question: "In your opinion, how often do people like you have to make unofficial payments or gifts in these situations?" focusing only on the receipt of public education. The answers are scored 1-5 where 1 corresponds to "never" and 5 to "always" . For each county we display the average over all respondents scores within the respective county.

Figure 3: The Romanian written exam average shares of students, by scores, by private and public schools, and by year.



Notes: Each bar segment is the average across all private schools (and public schools, respectively) of the share of students (in each school) who attain written Romanian scores in one of the 6 categories: below 5, 5-5.99, 6-6.99, 7-7.99, 8-8.99 and 9-10. All shares are weighted with the number of (per school) students taking the exam.

Table 1: Descriptive Statistics 2010-2007

		2010 (N=850)		**2009 (N=841)**	
		Mean	S.D.	Mean	S.D.
All schools	Public schools	0.937	0.241	0.942	0.232
	Share poor students	0.184	0.18	0.184	0.187
	Share male students	0.498	0.164	0.494	0.17
	Theoretic track	0.264	0.441	0.265	0.441
	Vocational track	0.08	0.271	0.079	0.27
	Technologic and mixed tracks	0.655	0.475	0.655	0.475
	Average Grade Romanian written exam	7.000	1.060	6.755	1.175
	Average Pass Rate Romanian written exam	0.939	0.086	0.912	0.111
Private	Average Grade Romanian written exam	5.618	0.813	5.746	0.783
	Average Pass Rate Romanian written exam	0.804	0.116	0.839	0.113
Public	Average Grade Romanian written exam	7.090	1.036	6.845	1.185
	Average Pass Rate Romanian written exam	0.945	0.074	0.917	0.102

		2008 (N=824)		**2007(N=837)**	
		Mean	S.D.	Mean	S.D.
All schools	Public schools	0.947	0.222	0.947	0.223
	Share poor students	0.144	0.162	0.106	0.133
	Share male students	0.483	0.167	0.477	0.163
	Theoretic track	0.266	0.442	0.26	0.439
	Vocational track	0.081	0.273	0.078	0.269
	Technologic and mixed tracks	0.651	0.476	0.661	0.473
	Average Grade Romanian written exam	7.007	1.091	6.686	1.109
	Average Pass Rate Romanian written exam	0.930	0.088	0.918	0.104
Private	Average Grade Romanian written exam	5.834	1.078	5.846	0.849
	Average Pass Rate Romanian written exam	0.816	0.143	0.855	0.134
Public	Average Grade Romanian written exam	7.032	1.077	6.712	1.106
	Average Pass Rate Romanian written exam	0.933	0.085	0.92	0.102

Notes: Average Grade Romanian written exam - the average grade in the Romanian written exam at school level; Average Pass Rate Romanian written exam the share of students per school who passed the Romanian written exam.

Table 2: Main effects, 2007-2010 academic years

Panel A: Average grade score on the standardised written Romanian exam				
	(1)	(2)	(3)	(4)
Public*Yr10	0.232**	0.266**	0.314**	0.455***
	(0.110)	(0.105)	(0.138)	(0.116)
Public*Yr09			0.023	0.123
			(0.137)	(0.106)
Public*Yr08			0.128	0.175
			(0.195)	(0.174)
Share Poor		0.143	0.142	-0.731***
		(0.224)	(0.225)	(0.188)
Share Males		-1.079***	-1.074***	-2.517***
		(0.189)	(0.189)	(0.211)
Theoretic				0.889***
				(0.050)
Technologic				-0.443***
				(0.059)
Public				0.753***
				(0.196)
Observations	3,324	3,324	3,324	3,324
R-squared	0.921	0.923	0.923	0.603
Panel B: Share of students passing the standardised written Romanian exam				
	(1)	(2)	(3)	(4)
Public*Yr10	0.030*	0.033**	0.031*	0.054***
	(0.016)	(0.015)	(0.019)	(0.017)
Public*Yr09			-0.019	-0.006
			(0.024)	(0.019)
Public*Yr08			0.022	0.031
			(0.034)	(0.028)
Share Poor		0.041	0.043	-0.002
		(0.026)	(0.026)	(0.018)
Share Males		-0.095***	-0.094***	-0.178***
		(0.029)	(0.029)	(0.022)
Theoretic				0.040***
				(0.005)
Technologic				-0.035***
				(0.007)
Public				0.059**
				(0.027)
Observations	3,324	3,324	3,324	3,324
R-squared	0.790	0.792	0.792	0.409
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Notes: All regressions are weighted with the number of (per school) students taking the exam. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Main effects, controlling for student ability, 2008-2010 academic years

Panel A: Average grade score on the standardised written Romanian exam			
	(1)	(2)	(3)
Public x yr10	0.258**	0.204	0.274**
	(0.124)	(0.124)	(0.135)
Average 5-8 grade score		0.392***	0.462***
		(0.091)	(0.109)
Share non-missing 5-8 score		0.419**	0.456***
		(0.168)	(0.172)
Share poor	-0.147	-0.321	-0.278
	(0.303)	(0.298)	(0.322)
Share males	-0.901***	-0.688***	-0.598**
	(0.271)	(0.261)	(0.267)
Average 5-8 grade score x yr10			-0.097**
			(0.039)
Share poor x yr10			-0.277
			(0.242)
Share males x yr10			-0.178
			(0.127)
Observations	2,297	2,297	2,297
R-squared	0.939	0.941	0.942
Panel B: Share of students passing the written Romanian exam			
	(1)	(2)	(3)
Public x yr10	0.040**	0.036**	0.051***
	(0.018)	(0.017)	(0.018)
Average 5-8 grade score		0.025**	0.046***
		(0.012)	(0.014)
Share non-missing 5-8 score		0.030	0.039
		(0.023)	(0.024)
Share poor	0.017	0.002	-0.009
	(0.034)	(0.039)	(0.042)
Share males	-0.060*	-0.047	-0.028
	(0.033)	(0.033)	(0.031)
Average 5-8 grade score x yr10			-0.027***
			(0.006)
Share poor x yr10			-0.028
			(0.028)
Share males x yr10			-0.022
			(0.020)
Observations	2,297	2,297	2,297
R-squared	0.826	0.827	0.833
Year FE	YES	YES	YES
School FE	YES	YES	YES
County trends	YES	YES	YES

Notes: All regressions are weighted with the number of (per school) students taking the exam. All regressions use the sample of schools from 2008-2010. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Matching private and public schools

Panel A: Average grade score on the standardised written Romanian exam				
	(1)	(2)	(3)	(4)
Public	0.676**	0.368*	0.256	0.305**
	(0.287)	(0.192)	(0.160)	(0.122)
Controls	NO	YES	YES	YES
Pre-reform outcome	NO	NO	YES	YES
County FE	NO	NO	NO	YES
Observations	78	78	78	78
R-squared	0.100	0.477	0.805	0.901
Panel B: Share of students passing the standardised written Romanian exam				
	(1)	(2)	(3)	(4)
Public	0.060**	0.043	0.043**	0.033*
	(0.029)	(0.026)	(0.020)	(0.018)
Controls	NO	YES	YES	YES
Pre-reform outcome	NO	NO	YES	YES
County FE	NO	NO	NO	YES
Observations	78	78	78	78
R-squared	0.065	0.225	0.627	0.768

Notes: All regressions are weighted with the number of (per school) students taking the exam. Controls include: theoretic track, share poor students, share males. The standard errors, shown in parentheses, are clustered at the locality level. Pre-reform outcome is the lag outcome from 2007, 2008 and 2009. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Main effects by county level of corruption, 2007-2010 academic years. Corruption proxy = unofficial payments in (public) education

Panel A: Average grade score on the standardised written Romanian exam				
I. Most corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	0.320***	0.359***	0.428***	0.555***
	(0.069)	(0.075)	(0.141)	(0.129)
Public*Yr09			0.015	0.072
			(0.176)	(0.130)
Public*Yr08			0.206	0.284
			(0.244)	(0.217)
Share Poor		-0.441	-0.438	-0.127
		(0.358)	(0.360)	(0.268)
Share Males		-0.968***	-0.961***	-2.560***
		(0.271)	(0.272)	(0.247)
Theoretic				0.926***
				(0.069)
Technologic				-0.349***
				(0.077)
Public				0.626**
				(0.253)
Observations	1,645	1,645	1,645	1,645
R-squared	0.920	0.922	0.922	0.594
II. Least corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	-0.020	0.004	-0.028	0.169
	(0.317)	(0.289)	(0.279)	(0.195)
Public*Yr09			-0.004	0.139
			(0.169)	(0.202)
Public*Yr08			-0.106	-0.122
			(0.190)	(0.189)
Share Poor		0.416	0.417*	-1.278***
		(0.252)	(0.251)	(0.268)
Share Males		-1.136***	-1.141***	-2.448***
		(0.288)	(0.286)	(0.395)
Theoretic				0.837***
				(0.081)
Technologic				-0.528***
				(0.084)
Public				0.965***
				(0.236)
Observations	1,303	1,303	1,303	1,303
R-squared	0.930	0.933	0.933	0.611
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Panel B: Share of students that passed the written Romanian exam				
I. Most corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	0.043***	0.046***	0.052***	0.074***
	(0.014)	(0.015)	(0.015)	(0.017)
Public*Yr09			-0.018	-0.008
			(0.030)	(0.023)
Public*Yr08			0.042	0.052
			(0.043)	(0.036)
Share Poor		-0.014	-0.010	0.050**
		(0.038)	(0.038)	(0.023)
Share Males		-0.083*	-0.081*	-0.190***
		(0.045)	(0.045)	(0.026)
Theoretic				0.051***
				(0.006)
Technologic				-0.026***
				(0.010)
Public				0.040
				(0.032)
Observations	1,645	1,645	1,645	1,645
R-squared	0.782	0.783	0.785	0.417
II. Least corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	-0.011	-0.010	-0.041	0.004
	(0.037)	(0.035)	(0.043)	(0.043)
Public*Yr09			-0.039	-0.002
			(0.037)	(0.043)
Public*Yr08			-0.047	-0.016
			(0.036)	(0.032)
Share Poor		0.062*	0.065*	-0.046*
		(0.036)	(0.037)	(0.027)
Share Males		-0.102**	-0.103**	-0.168***
		(0.042)	(0.041)	(0.044)
Theoretic				0.027***
				(0.006)
Technologic				-0.044***
				(0.011)
Public				0.105*
				(0.055)
Observations	1,303	1,303	1,303	1,303
R-squared	0.815	0.817	0.818	0.414
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Notes: In particular, we use the question: "In your opinion, how often do people like you have to make unofficial payments or gifts in these situations?", and we focus only on public education. All regressions are weighted with the number of (per school) students taking the exam. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Share of expelled students (caught cheating) from the exam, 2007-2010 academic years

Share of expelled students from the exam				
	(1)	(2)	(3)	(4)
Public*Yr10	-0.005	-0.005	-0.009	-0.007
	(0.007)	(0.007)	(0.007)	(0.006)
Public*Yr09			-0.005**	-0.004**
			(0.002)	(0.002)
Public*Yr08			-0.004	-0.005
			(0.003)	(0.003)
Share Poor		-0.002	-0.001	0.001
		(0.002)	(0.002)	(0.001)
Share Males		0.000	0.000	0.002***
		(0.002)	(0.002)	(0.001)
Theoretic				-0.000**
				(0.000)
Technologic				0.000
				(0.000)
Public				0.001
				(0.001)
Observations	3,324	3,324	3,324	3,324
R-squared	0.330	0.331	0.333	0.066
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Notes: All regressions are weighted with the number of (per school) students taking the exam. Columns (1)–(4) unbalanced panel; column (5) balanced panel. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Main effects by scores, 2007-2010 academic years

Panel A: Share students that passed the Romanian written exam with scores 5-6				
	(1)	(2)	(3)	(4)
Public*Yr10	-0.038**	-0.041**	-0.039	-0.026
	(0.019)	(0.019)	(0.024)	(0.023)
Public*Yr09			-0.013	-0.011
			(0.022)	(0.019)
Public*Yr08			0.026	0.043*
			(0.028)	(0.026)
Share Poor		0.002	0.004	0.159***
		(0.045)	(0.046)	(0.032)
Share Males		0.108***	0.109***	0.418***
		(0.040)	(0.041)	(0.030)
Theoretic				-0.183***
				(0.010)
Technologic				0.072***
				(0.013)
Public				-0.089***
				(0.030)
Observations	3,324	3,324	3,324	3,324
R-squared	0.883	0.884	0.884	0.556
Panel B: Share students that passed the Romanian written exam with scores 7-10				
	(1)	(2)	(3)	(4)
Public*Yr10	0.068***	0.074***	0.070**	0.080***
	(0.022)	(0.022)	(0.031)	(0.025)
Public*Yr09			-0.006	0.005
			(0.030)	(0.023)
Public*Yr08			-0.004	-0.012
			(0.042)	(0.038)
Share Poor		0.039	0.040	-0.161***
		(0.055)	(0.056)	(0.043)
Share Males		-0.203***	-0.203***	-0.595***
		(0.046)	(0.046)	(0.046)
Theoretic				0.223***
				(0.013)
Technologic				-0.106***
				(0.015)
Public				0.148***
				(0.042)
Observations	3,324	3,324	3,324	3,324
R-squared	0.913	0.914	0.914	0.598
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Notes: All regressions are weighted with the number of (per school) students taking the exam. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

A Appendix – Supplementary Tables

Table A1: Main effects using the sample of exam centers: mixed public with private schools

Panel A: Average grade score on the standardised written Romanian exam						
	(1)	(2)	(3)	(4)	(5)	(6)
Public*Yr10	0.272*	0.322**	0.420**	0.583***	0.523***	0.591***
	(0.144)	(0.140)	(0.164)	(0.158)	(0.167)	(0.211)
Public*Yr09			0.107	0.220	0.270*	0.172
			(0.154)	(0.133)	(0.144)	(0.172)
Public*Yr08			0.174	0.227	0.213	0.163
			(0.168)	(0.151)	(0.160)	(0.203)
Share Poor		0.188	0.134	-0.958**	-1.027***	-1.298***
		(0.546)	(0.548)	(0.413)	(0.380)	(0.487)
Share Males		-1.347***	-1.335***	-2.198***	-1.855***	-2.557***
		(0.426)	(0.425)	(0.308)	(0.293)	(0.376)
Theoretic				0.777***	0.558***	0.289*
				(0.116)	(0.116)	(0.155)
Technologic				-0.269**	-0.339***	-0.413***
				(0.113)	(0.094)	(0.112)
Public				0.850***	0.903***	0.896***
				(0.163)	(0.130)	(0.166)
Observations	738	738	738	738	738	417
R-squared	0.945	0.947	0.947	0.706	0.829	0.842
Panel B: Share of students passing the standardised written Romanian exam						
	(1)	(2)	(3)	(4)	(5)	(6)
Public*Yr10	0.031	0.035*	0.038	0.060**	0.056**	0.085**
	(0.021)	(0.020)	(0.024)	(0.024)	(0.026)	(0.036)
Public*Yr09			-0.011	0.004	-0.000	0.013
			(0.023)	(0.021)	(0.019)	(0.026)
Public*Yr08			0.027	0.038	0.032	0.032
			(0.031)	(0.025)	(0.025)	(0.034)
Share Poor		0.059	0.069	-0.002	-0.037	-0.080
		(0.059)	(0.059)	(0.041)	(0.041)	(0.057)
Share Males		-0.142**	-0.133**	-0.137***	-0.132***	-0.185***
		(0.066)	(0.065)	(0.028)	(0.027)	(0.040)
Theoretic				0.033***	0.016	-0.012
				(0.010)	(0.012)	(0.017)
Technologic				-0.018*	-0.046***	-0.048***
				(0.011)	(0.011)	(0.015)
Public				0.067***	0.070***	0.058*
				(0.024)	(0.024)	(0.033)
Observations	738	738	738	738	738	417
R-squared	0.849	0.852	0.854	0.516	0.661	0.667
Year FE	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	NO	NO	NO
County FE	NO	NO	NO	YES	NO	NO
Center FE	NO	NO	NO	NO	YES	YES
County trends	YES	YES	YES	YES	YES	YES

Notes: All regressions are weighted with the number of (per school) students taking the exam. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

B Appendix – Supplementary Analysis

B1. Heterogeneous effects

Finally, in this section, we explore whether corruption responds to the wage cut in distinct ways across high schools with different characteristics. In particular, in Table B1, we look at DD estimates in schools with different proportions of female students (Panel A of Table B1), different ethnic compositions (Panel B), varying shares of teachers paid by the hour (Panel C) and, different age of the school principal (Panel D).

The most interesting findings are the following:

a) The DD estimates are significant only for high schools with a minority population of female students, suggesting that male dominated schools are more prone to appeal to corruption especially when the financial incentives are accentuated. While this does not exclude milder forms of fraud, such as increased male to female student cheating in the exam rooms, this finding is also consistent with an outward shift in demand for illegal grades meeting the increased supply by didactic staff, where male students are dominant.

b) The impact of the wage cut is significant in ethnically mixed high schools (defined as having the share of Romanians less than 1), which is true both for the average pass and for the average grade in the Romanian written exam.

c) The findings are mixed for schools with a different share of teachers working part time. Effects are larger in magnitude for those with higher prevalence (i.e., the share of teachers paid by the hour is larger than the mean=11%), suggesting they might be more responsive to monetary incentives. This might indicate that less organised schools or teachers who have loose ties to the teacher labor market (by being hired on a temporary contract), are more easily influenced by principals to be involved in corrupt behavior. However, it should be noted that very few schools have a high proportion of part-time teachers. If we exclude the few schools with more than 50% of teachers paid by the hour, we get positive and statistically DD-estimates that are in line with our baseline estimates.

d) Schools with a younger school principal (i.e., smaller than the mean age=48) are more responsive to monetary incentives. This might be in line with the increase in corruption in schools over time in Romania, so that older principals were used to working in a system of less corruption.

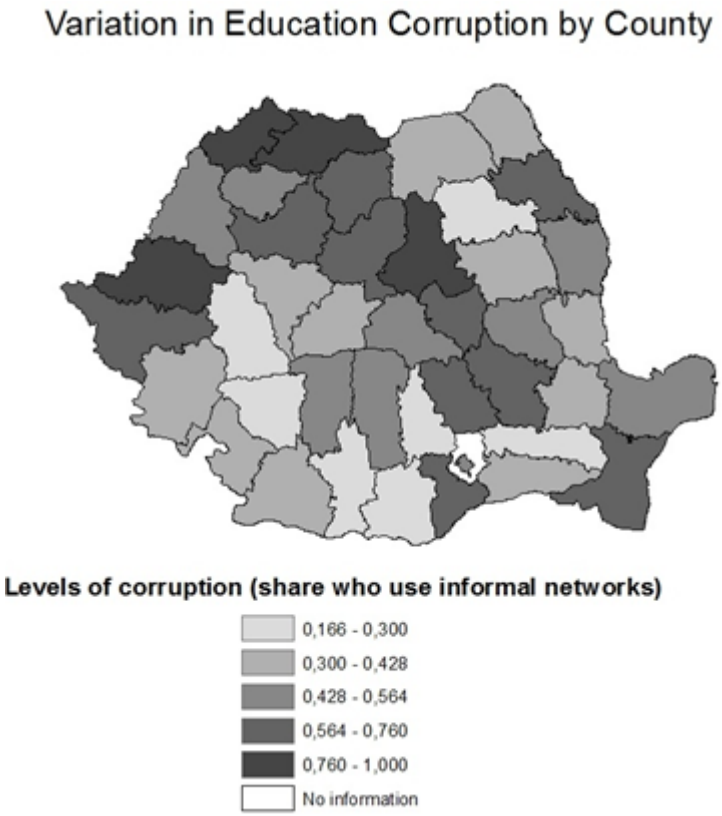
Table B1: Heterogeneous effects: gender, ethnic composition, teacher and management composition, all outcomes, 2007-2010 academic years

	Average Grade Romanian Exam				**Share Students Passing Romanian Exam**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Gender	**I. Female share<0.5**		**II. Female share>0.5**		**I. Female share<0.5**		**II. Female share>0.5**	
pub10	0.266*	0.262	0.000	-0.238	0.048**	0.043*	0.019	-0.006
	(0.156)	(0.173)	(0.193)	(0.186)	(0.022)	(0.023)	(0.027)	(0.026)
Observations	1,488	1,488	1,836	1,836	1,488	1,488	1,836	1,836
R-squared	0.875	0.875	0.935	0.935	0.759	0.760	0.834	0.834
Panel B: Ethnicity	**I. Share Romanians=1**		**II. Share Romanians<1**		**I. Share Romanians=1**		**II. Share Romanians<1**	
pub10	0.149	0.312	0.328**	0.209*	0.027	0.036	0.055***	0.046**
	(0.174)	(0.216)	(0.138)	(0.120)	(0.023)	(0.041)	(0.021)	(0.022)
Observations	830	830	2,492	2,492	830	830	2,492	2,492
R-squared	0.923	0.924	0.922	0.922	0.811	0.812	0.782	0.783
Panel C: Part-time teachers	**I. Share part-time<=0.11**		**I. Share part-time>0.11**		**I. Share part-time<=0.11**		**I. Share part-time>0.11**	
pub10	0.355*	0.668***	0.260**	0.111	0.027	0.059***	0.038**	0.021
	(0.181)	(0.169)	(0.130)	(0.159)	(0.029)	(0.008)	(0.019)	(0.022)
Observations	2,070	2,070	1,226	1,226	2,070	2,070	1,226	1,226
R-squared	0.926	0.927	0.917	0.917	0.801	0.801	0.780	0.781
Panel D: Principals age	**I. Principals age<48**		**II. Principalsage>=48**		**I. Principals age<48**		I. Principals age<48	
pub10	0.357***	0.405***	0.073	0.187	0.055***	0.062***	-0.015	-0.012
	(0.126)	(0.133)	(0.149)	(0.235)	(0.017)	(0.013)	(0.026)	(0.026)
Observations	1,602	1,602	1,624	1,624	1,602	1,602	1,624	1,624
R-squared	0.917	0.917	0.931	0.931	0.801	0.801	0.790	0.790
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
School FE	YES	YES	YES	YES	YES	YES	YES	YES
County trends	YES	YES	YES	YES	YES	YES	YES	YES
Public x year interactions	NO	YES	NO	YES	NO	YES	NO	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: All regressions are weighted with the number of students (per school) taking the exam and the standard errors, shown in parentheses, are clustered at the municipality level. We show: in Panel A schools with different shares of female students; in Panel B ethnically homogenous and non-homogenous schools; Panel C shares of teachers paid by hour contract (0.11=mean); Panel D average age of schools management (48 years=mean). The benchmark year in columns (2), (4), (6) and (8) is 2007. ***p<0.01, **p<0.05, *p<0.1

B2. Results using an alternative measure of corruption

Figure B1



Notes: Our calculations using the 2007 Public Opinion Barometer, Soros. We use the question: "There is anyone (i.e., informal network) that could "help" you solve (i.e., informally): issues in court/trials, medical problems, city hall, police, or issues related to the local authorities"

Table B2: Main effects by county level of corruption, 2007-2010 academic years. Corruption proxy = share that use informal network

Panel A: Average grade score on the standardised written Romanian exam				
I. Most corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	0.350***	0.376***	0.445***	0.532***
	(0.093)	(0.096)	(0.168)	(0.139)
Public*Yr09			0.025	0.082
			(0.199)	(0.149)
Public*Yr08			0.189	0.220
			(0.269)	(0.234)
Share Poor		-0.176	-0.175	-0.750**
		(0.371)	(0.370)	(0.328)
Share Males		-1.038***	-1.028***	-2.437***
		(0.244)	(0.244)	(0.319)
Theoretic				0.936***
				(0.063)
Technologic				-0.395***
				(0.077)
Public				0.650***
				(0.219)
Observations	1,941	1,941	1,941	1,941
R-squared	0.923	0.924	0.924	0.592
II. Least corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	0.035	0.093	0.100	0.357*
	(0.263)	(0.243)	(0.213)	(0.212)
Public*Yr09			0.011	0.155
			(0.130)	(0.100)
Public*Yr08			0.010	0.106
			(0.196)	(0.204)
Share Poor		0.263	0.262	-0.770***
		(0.263)	(0.268)	(0.227)
Share Males		-1.194***	-1.194***	-2.607***
		(0.276)	(0.276)	(0.201)
Theoretic				0.834***
				(0.087)
Technologic				-0.511***
				(0.096)
Public				0.942***
				(0.354)
Observations	1,383	1,383	1,383	1,383
R-squared	0.919	0.921	0.921	0.614
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Panel B: Share of students that passed the written Romanian exam				
I. Most corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	0.037**	0.038**	0.036	0.060***
	(0.015)	(0.015)	(0.022)	(0.020)
Public*Yr09			-0.023	-0.006
			(0.035)	(0.027)
Public*Yr08			0.025	0.035
			(0.051)	(0.041)
Share Poor		0.039	0.042	0.028
		(0.038)	(0.038)	(0.030)
Share Males		-0.082**	-0.079**	-0.181***
		(0.036)	(0.035)	(0.037)
Theoretic				0.044***
				(0.006)
Technologic				-0.035***
				(0.010)
Public				0.057*
				(0.033)
Observations	1,941	1,941	1,941	1,941
R-squared	0.790	0.791	0.792	0.394
II. Least corrupted counties	(1)	(2)	(3)	(4)
Public*Yr10	0.020	0.025	0.025	0.042
	(0.035)	(0.033)	(0.036)	(0.033)
Public*Yr09			-0.012	-0.008
			(0.029)	(0.025)
Public*Yr08			0.019	0.023
			(0.026)	(0.028)
Share Poor		0.032	0.033	-0.022
		(0.036)	(0.037)	(0.020)
Share Males		-0.117**	-0.116**	-0.173***
		(0.045)	(0.046)	(0.019)
Theoretic				0.034***
				(0.007)
Technologic				-0.034***
				(0.010)
Public				0.063
				(0.046)
Observations	1,383	1,383	1,383	1,383
R-squared	0.791	0.795	0.795	0.434
Year FE	YES	YES	YES	YES
School FE	YES	YES	YES	NO
County FE	NO	NO	NO	YES
County trends	YES	YES	YES	YES

Notes: In particular, we use the question: "There is anyone (i.e., informal network) that could "help " you solve (i.e., informally): issues in court/trials, medical problems, city hall, police, or issues related to the local authorities". All regressions are weighted with the number of (per school) students taking the exam. The standard errors, shown in parentheses, are clustered at the locality level. *** p<0.01, ** p<0.05, * p<0.1

Chapter III

Fighting Corruption in Education: What Works and Who Benefits?*

Oana Borcan
University of Gothenburg

Mikael Lindahl
University of Gothenburg

Andreea Mitrut
University of Gothenburg

Abstract

We investigate the efficiency and distributional consequences of a corruption-fighting initiative in Romania targeting the endemic fraud in a high-stakes high school exit exam, which introduced CCTV monitoring of the exam and credible punishment threats. We find that punishment coupled with monitoring was effective in reducing corruption. Estimating the heterogeneous impact for students of different ability, poverty status, and gender, we show that fighting corruption led to efficiency gains (ability predicts exam outcomes better) but also to a worrisome score gap increase between poor and non-poor students. Consequently, the poor students have reduced chances to enter an elite university.

JEL Codes: I21, I24, K42

Keywords: corruption, high-stakes exam, bribes, monitoring and punishment

E-mail: Oana.Borcan@economics.gu.se, Mikael.Lindahl@economics.gu.se and Andreea.Mitrut@economics.gu.se (corresponding author), respectively. All errors are our own. Andreea Mitrut gratefully acknowledges support from the Jan Wallanders and Tom Hedelius Foundation. Mikael Lindahl is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Torsten and Ragnar Söderberg Foundation, and also acknowledges financial support from the Scientific Council of Sweden and the European Research Council [ERC starting grant 241161]. We thank Matthew Lindquist, Randi Hjalmarsson, Brian Knight, Marco Manacorda, Erich Battistin, Pedro Dal Bo, and David Weil for useful comments and discussions. Seminar participants at Brown University, the "Economics of Education and Education Policy" workshop, IFAU, Uppsala, UCLS, Stockholm School of Economics (SITE), SSB, Oslo, and Gothenburg University provided useful comments. We also thank the Ministries of Education in Romania and Moldova for providing data, as well as Diana Coman and SIVECO Romania for excellent help with the data.

1 Introduction

Equality of educational opportunity for individuals of similar ability is a key ingredient in a society that wants to promote growth and increase social mobility. A meritocratic education system increases the efficiency of how talented individuals are allocated by rewarding ability and not family income. However, in many countries, hurdles such as tuition fees and school and neighborhood segregation may reinforce inequality of opportunities across generations and increase inequality by limiting skill acquisition and access to higher education for poor individuals of high ability. An additional barrier to higher education, mostly prevalent in developing countries, is corruption in education, including bribes taken by teachers to facilitate admission to education or to inflate grades and scores on high-stakes exams. Moreover, corruption in education may act as an added tax, putting the poor students at a disadvantage and reducing, once more, equal access to human capital (see the 2013 Global Corruption Report, GCR).

This paper analyzes the implications of the fight against corruption in a setting of endemic fraud, cheating, and grade selling in the public education system in Romania.[1] Particularly, we investigate the efficiency and distributional consequences of a national anti-corruption campaign targeting the Romanian high school exit exam the Baccalaureate.[2] The campaign was initiated in 2011 in response to the 2010 Baccalaureate, which marked a peak in corruption for exam grades and generated a media storm after Romanian National Anticorruption Directorate revealed how batches of identical answers had been distributed to students by public teachers (see Borcan, Lindahl and Mitrut, 2014).[3] The campaign consisted of two distinct components: 1) increasing the threat of punishment for teachers and students caught taking/giving bribes and 2) closed-circuit TV (CCTV) monitoring of the exam centers in an effort to eradicate mass cheating and bribes during the examination.[4]

Our aim in this paper is to first evaluate the efficiency of the national anti-corruption campaign and subsequently to understand who the winners (and losers) are, especially in terms of students' poverty status, ability, and gender characteris-

[1] The prevalence of corruption in the Romanian, as well as many other countries public education is acknowledged in the World Bank Report Gobal Corruption Report, 2013.

[2] Corruption in this setting refers to the giving of bribes for permission to cheat or for higher scores than deserved.

[3] This exam became known in the media as the "Xeroxed exam," referring to the fact that many students were found to have identical test answers including in essay type exams.

[4] While similar policies are currently discussed in other countries, Moldova and Cambodia have already implemented a similar policy targeting the endemic corruption in connection with the high school exit exam, resulting in 56% and 26% of students passing the exam compared with over 94% and 87%, respectively, in the past.

tics. To accomplish our first objective, we evaluate the punishment and monitoring components of the campaign. For teachers, the punishment side of the campaign comprised threats of dismissals and imprisonment, while corrupt students risked being suspended from any retakes for over a year. The commitment to punish teachers and students caught red-handed was demonstrated by the high number of trials related to exam fraud immediately after the 2010 Baccalaureate. The installation of CCTV cameras in exam centers, the second component of the campaign, was an effort to eradicate mass cheating and fraud. This measure was not announced until May 2011, i.e., one month prior to the high-stake Baccalaureate exam. Just over half of the counties had video surveillance in 2011, while the rest installed cameras in 2012 when CCTV surveillance became mandatory. Hence, for the monitoring part of the campaign we have access to quasi experimental variation in camera installation, which we utilize in a difference-in-differences (DD) framework, comparing counties treated with the camera monitoring (some in 2011 and all in 2012) with those not treated (all in 2009-2010 and some in 2011). This yields an estimate of the effect of increased monitoring on high-stakes Baccalaureate scores. The punishment component was implemented across the country at the same time, but because of its strict implementation and since we can use a placebo test as control, we are able to say something about the impact of the monitoring and punishment combined.

Having established that the anti-corruption campaign did have an overall effect in lowering Baccalaureate scores and pass rates, we next investigate who the winners and losers from the campaign are. We analyze the heterogeneous effects of the anti-corruption campaign for the students: high vs. low ability, high vs. low income (poor), and males vs. females. This will give us an idea of how different groups fare in a more or less corrupt education system. Given that bribing requires economic resources and is an opportunity to circumvent effort and ability in producing high scores, we hypothesize that eliminating or decreasing corruption in relation to the Baccalaureate benefits poor students and makes ability a more important predictor of the Baccalaureate score. As the Baccalaureate score is the only or major admission criteria for higher education in Romania, we expect our results on Baccalaureate outcomes to carry over to the admission to higher education. To corroborate this finding, we have collected additional data to directly investigate the consequences of the (monitoring component of the) anti-corruption policy for admission to higher education at an elite university.

We provide a number of interesting findings. We find that exam outcomes dropped sharply already in 2011 and that the drop came from both the monitored and non-monitored counties, yet it was larger in the monitored ones. By 2012, the

average pass rate had almost halved. In the DD analysis we find that the presence of CCTV cameras reduced the Romanian written exam score by 0.12 SD and the probability of passing the Baccalaureate by 8.3 percentage points. We interpret these estimates as the additional effect of introducing monitoring. The analogous analysis of a no-stakes exam, with no scope for corruption (the oral Romanian exam), reveals neither a general drop in scores in 2011 or 2012, nor a decrease in response to monitoring. We interpret this as suggestive evidence that punishment works well, in particular when complemented with monitoring. Moreover, we corroborate this finding with very similar pattern for pass rates at the baccalaureate in Moldova, a country with a very similar educational structure as Romania which introduced harsher punishments in 2012 and CCTV cameras in 2013.

As expected, the campaign increases the importance of ability for exam outcomes, implying efficiency gains. More surprisingly, our findings contradict our original expectation that fighting corruption should close the score gap between poor and non-poor students. The results indicate that the anti-corruption measures made the already underperforming poor students relatively worse off than non-poor students. The campaign induced an increase in achievement gaps, in that groups performing relatively worse prior to the campaign (low ability, poor, males) became even more worse off relative to the groups performing better. This was likely the result of the structure of corruption and the pre-existing inequalities hidden behind it, as we discuss in section 6. Importantly, we are also able to investigate the consequences of the anti-corruption policy (the monitoring component) for admission to higher education. Using data from an elite university, we show that, while the ability and gender composition of students at this top school was unaffected by the introduction of cameras, the monitoring significantly reduced the chances of admission for poor students, hence confirming most of the results found for the Baccalaureate.

Our paper makes several contributions to the literature on fighting corruption and on the economic consequences of corruption. Economic theory argues that the right combination of increasing the probability of detection (through monitoring) and the threat of punishment may reduce corruption by increasing its costs (Becker and Stigler, 1974). However, evaluation of policies that combine punishment and monitoring has proven to be a challenging task (Hanna et al., 2011; Svensson, 2005). The setting we have for the year 2011 is one where, akin to a Becker-Stigler model of crime, we have a combination of incentives and varying detection probabilities. Counties that installed cameras faced both a stronger incentive (credible punishment threat) and increased monitoring, whereas counties that did not install cameras faced the new punishment threats but no increase in actual monitoring. This allows us

to bring additional evidence on the interplay between punishment and monitoring and their effects on exam outcomes. Our research therefore complements the literature on anti-corruption policies, which has so far explored monitoring through official audits (Ferraz and Finan, 2008, 2011; Di Tella and Schargrodsky, 2003) and community-based monitoring interventions (Duflo et al., 2012; Reinikka and Svensson, 2004, 2005; Olken, 2007), and has also analyzed changes in incentives (Banerjee et al., 2008; Duflo et al., 2012). Some of these studies shed light on the interplay and relative effectiveness of monitoring and incentives in discouraging dishonest practices.[5] Our paper offers evidence that monitoring is effective insofar as it enables incentive schemes to operate better, even in the high-stakes setting of a high school exit exams of crucial importance for future education and success in the labor market. The paper also contributes additional evidence of the effectiveness of monitoring to an emerging literature on the role of CCTV cameras in combating crime (Priks, 2014, 2015; King et al. 2008, Welsh and Farringdon, 2009, 2003).[6]

One important contribution of our paper is the estimated impact of fighting corruption on equality of educational opportunity. While social scientists have argued that (income) inequality is positively correlated with the level of corruption (see, e.g., You and Khagram, 2005; Rothstein and Uslaner, 2005), little is known about the distributional consequences of the various means to fight corruption and particularly how curbing corruption influences inequality of opportunity in a society. This is problematic as corruption might adapt and transform to circumvent new constraints, generating a redistribution of resources and opportunities that could increase inequality. Importantly, empirical evidence on the welfare consequences of

[5]Nagin et al. (2002) report on a field experiment which showed that that decreasing the rate of monitoring observable by employees led them to shirk more, independently of how good their alternatives in the labor market were relative to their job. Di Tella and Schargrodsky (2003) examine the effects of wages and audits during a crackdown on corruption in Buenos Aires hospitals. They find that the wages played no role in reducing corruption (inferred from the drop in previously inflated hospital input prices) when the probability of detection was close to 100%, but only when auditing was less frequent. Duflo et al. (2012) show that monitoring with tamper-proof cameras worked in reducing teacher absenteeism insofar as it was instrumental in implementing an incentivizing attendance-based wage scheme. Their model predicts that at the very least, punishment prospects (fear of dismissal) should put a bound on dishonest behavior. Banerjee et al. (2008) follow the punishment approach of incentives and show that credible threats of punishment (through pay cuts and dismissal) were indispensable in getting government nurses in India to come to work, even when camera monitoring was in place. The impact of changing monitoring or incentives of corruption and shirking linked with the education process is also illustrated in Glewwe et al. (2010).

[6]The effectiveness of CCTV cameras in reducing crime is a current topic, with million dollars being spent in this public safety infrastructure. Priks (2014) and Priks (2015) documents the causal effects of CCTV cameras on unruly behaviour and some types of crime, using temporal variation in CCTV installation in Swedish stadiums and underground. King et al (2008) showed that property crime was reduced as a result of CCTV monitoring on the streets of San Francisco.

corruption remains very scarce.[7] By separating the effects of corruption elimination between low- and high-income students and between low- and high-ability students, we will also infer the consequences of corruption on educational opportunity for students from different backgrounds a perspective neglected in previous studies. Allocative inefficiencies, for instance in the selection into higher education, can have great consequences for longer-run economic development and economic inequality (Banerjee et al., 2012). Our paper also related to the large literature on how credit constraints (in this paper in the form of bribes) affect continuation to higher education, especially with regard to selection across the ability and family income distribution (see Lochner and Monge-Naranjo, 2012, for a survey).

The paper is structured as follows. Section 2 presents the setting and the anti-corruption initiatives. Section 3 provides the details of our data. Section 4 provides a graphical analysis of the data. Section 5 outlines our empirical strategy. Section 6 presents our main empirical findings. Section 7 presents the effects on admission to university. Our conclusions are presented in Section 8.

2 Background

2.1 The Romanian education system

The Romanian pre-university education starts with elementary school, which is divided into primary school (1st to 4th grade) and secondary school, or gymnasium (5th to 8th grade). Upon graduation from secondary school, i.e., at the end of 8th grade, the students need to pass a national standardized exam. The score from this exam and the student's graduation grade point average (5th to 8th grade) contribute with equal weights to the student's tertiary or high school admission grade. Based on this score and a comprehensive list of ranked high schools, the student is systematically allocated by the Ministry of Education (through a computerized, transparent allocation procedure) to a high school and a specific track at that school: i) a theoretical track, which includes humanities and sciences,[8] ii) a technological track, which includes technical training, services, and natural resource- and environment protection-oriented education, or iii) a vocational track, which includes

[7]Exceptions include Ferraz et al. (2012), who explore variation in corruption in education across Brazilian municipalities, showing how more corruption translates into lower scores for the students, thereby assessing the efficiency costs of corruption, and Choe et al. (2013), who show survey evidence from Bangladesh that corruption in education is most taxing for the poor and less educated. Similarly, Hunt (2007) shows evidence from Peru that the victims of misfortune (crime) are also more likely to be victims of bribery.

[8]The theoretical track is typically the most popular among high-ability students.

arts, military, theology, sports and teaching (for more details on the allocation, see Pop-Eleches and Urquiola, 2013).

Upon completion of high school, students take the Baccalaureate exam. This high-stakes nationwide standardized test is mandatory in order to obtain a tertiary education degree. Admission to university or further training as well as access to the labor market are almost exclusively based on this test.[9] The exam takes place every year in June and consists of a few oral and written standardized tests, with slight alterations across years.[10] The tests within each subject and may have different degrees of difficulties across tracks, but they are standard within one track. The only exception is the written exam in Romanian language and literature, which is the one test that is identical for all students regardless of track, and its format has remained unchanged over the years.

2.2 The Baccalaureate and the Anti-corruption Campaign

The pressure of passing the Baccalaureate exam (with high scores) has been constantly rising since about 2002. It was around then that the increase in the number of private universities and the introduction of tuition fees in public higher education began. This made the university admission exams less relevant as the Baccalaureate scores attained increasing shares in the admission criteria (up to 100%), raising the stakes of the high school exit exam. The combination of the high stakes and poor remuneration of public school teachers created an endemic corruption environment surrounding the Baccalaureate exam, as also documented by Borcan, Lindahl, and Mitrut (2014).[11]

The unofficial payments behind the Baccalaureate exam can be summarized as follows:[12] *i) Collective bribes* which are funds collected from the students before, or just before the exam. These are voluntary but very common, usually perceived as a

[9] All tests and school grades in Romania are scored on a scale from 1 to 10, and to pass a student must obtain a minimum score of 5 on each test. However, to pass the Baccalaureate a student needs at least 5 on each exam and a minimum overall average score of 6.

[10] The most important changes were the exclusion of oral tests from the overall score starting in 2010 and the elimination of the fourth written test. All these tests displayed abnormal score distributions highly concentrated at the top marks.

[11] A 2003 World Bank Report on corruption in Romania reveals that more than 67% of the respondents alleged that all or almost all public officials in Romania are corrupt, while more than 50% of the respondents believed that bribery is part of the everyday life in Romania. The figure was particularly high for the education and health systems, as up to 66% of the respondents confirmed that they were paying the so-called atentie (unofficial payments or bribes). According to the Global Corruption Barometer from Transparency International, in Romania in 2010/2011, 37% of respondents believed the education system was corrupt or extremely corrupt, which was above the world average.

[12] This distinction is based on examples of bribes documented in the court cases and official press releases of the National Anticorruption Directorate (retrieved from www.pna.ro - in Romanian).

norm by all students and are used to "grease the wheels"("protocol" meals or, small gifts for the exam committee) or directly given to the exam committee and proctors to turn a blind eye or even help in-class cheating. Because these bribes affect what happens during the exam, it is this type of corruption ("pay to cheat") that the CCTV monitoring can reduce. *ii) Individual bribes*, which are large sums (hundreds of euro) transferred privately by the more affluent students to members of the exam committee to increase the student's score, or to replace the exam paper with a correct version.[13] This is usually done with the help of a student's teacher or school principal who act as intermediaries for the bribe transfers. The corruption trials following the 2010 exam illustrate this form of bribing: "*The defendant [school principal, name] claimed and received from the defendants [names] the total amount of 7.000 RON, which she then transferred to the defendant [name]. This money was received in order for the latter, as examiner in Romanian language, to give higher scores for the (contributing) candidates*" (National Anticorruption Directorate Press release No. 473/VIII/3, 2010). Thus, while punishment threats may affect the incidence of individual bribes, CCTV monitoring cannot capture these private deals. The fact that the pass rates of 80-90% until 2009 did not reflect ability but rather mass fraud was common knowledge among teachers, principals, parents, and students.[14]

Following the 2010 Baccalaureate, which was marked by a surge in grade-inflating corruption generated by the 25% public sector wage cut in May 2010, a high number of teachers were brought to trial on allegations of selling grades.

In response to this scandal, the Ministry of Education started a Baccalaureate "cleaning" campaign in 2011. In a first step, the Ministry publicly appealed to all schools and teachers involved in the exam to better enforce the examination rules and threatened to punish teachers caught receiving bribes with a pay cut and/or jail,[15] while also promoting a zero tolerance policy against collective bribes. Additionally, a new rule stipulated that parents and NGOs had the right to enroll as

[13] *"Around the time of the Baccalaureate exam, June 2010, in the exam center [name], the defendants [name] - principal, [name] - deputy principal, [name] - secretary and [name] - teacher, [...] have [...] planned and organized a fraudulent exam, in which students who paid various amounts of money passed the tests. [...] On June 28, 2010, after the written Romanian exam, upon a police search of the high school premises, 56 envelopes containing money and the names of the students [who have contributed] have been identified. In total 91.850 RON (equivalent to 21,360 EUR) and 7,750 EUR have been found. In addition, [the principals] have received 19,000 RON, 1,850 EUR and 8 envelopes containing unspecified amounts from students interested in passing the exam."* Press release No. 633/VIII/3, National Anticorruption Directorate, November 29, 2010.

[14] For a more detailed treatment of the state of corruption in Romania, particularly in the education system, see Borcan, Lindahl, and Mitrut (2014). Based on PISA test scores, the authors also document the strong contrast between national exam scores and true ability compared with other European countries.

[15] Threats ranged in severity from being excluded from the examination for a few years to going to jail (following the 2010 example).

exam proctors in order to increase transparency. In terms of harsher punishments, the new rules also stipulated that students caught cheating would be banned from re-takes for at least one year. On top of these measures, there was a recommendation to organize the exam, when possible, in centers equipped with surveillance cameras. The introduction of CCTV cameras was reinforced in May 2011 through public appeals by the Ministry of Education to the county inspectorates. However, because the request was not binding, each of the 42 county inspectorates decided independently whether or not to install CCTV cameras in the examination centers by the end of May.[16] As a result, twenty-five counties had cameras installed in the examination centers and 17 did not, blaming lack of funds. Where installed, the cameras were placed in the front of the room, or on the hallways, and the camera footages were collected and screened by the county inspectorates. Descriptive statistics (Appendix C, Table C1) confirm that the counties that did not install CCTV cameras in 2011 were poorer than the others. We discuss the county self-selection later on.

Thus, in 2011, counties that installed cameras faced both a credible punishment threat and increased monitoring, while non-implementers faced a credible punishment threat but no additional actual monitoring.[17] Consistent with this, the national average pass rates plummeted to a staggering 44.5% (from around 70% in 2010). Both implementers and non-implementers of the camera policy sustained a drop in the pass rates, but the drop was much larger in the monitored (pass rates decreased to 41%) compared to the non-monitored counties (where pass rates dropped to 51%). In 2012, the Baccalaureate methodology was further modified and CCTV cameras became mandatory in *all* counties, which reduced the pass rates even further to 41.5%.

The gradual introduction of monitoring allows us to compare education outcomes in a corrupt (in 2011 in non-monitored counties and in 2010 and before) and a non-(or less) corrupt system (in 2011 in monitored counties and in 2012 in all counties). This variation sets the foundation for our empirical strategy, as described in Section 4.

[16] *Metodologia de organizare si desfasurare a examenului de bacalaureat*, 2011, Annex 2 of the Ministry of Education's Decision no. 4799/31.08.2010, concerning the organization of the Baccalaureate exam.

[17] However, the latter counties may have expected a higher rate of monitoring due to the increasing pressure from counties that complied. For instance, counties that decided against the Ministry's recommendation may have feared being targeted with more frequent inspections. Since agents' behavior responds to perceived monitoring, which does not necessarily coincide with objective monitoring (Nagin et al., 2002), we can plausibly assume that the expected detection probability increased also in non-implementing counties, but to a lower extent than in implementing counties.

3 Data

For the purpose of our empirical investigation we employ several datasets:

i) Administrative data provided by the Ministry of Education and covering the universe of students enrolled at the Baccalaureate exam (typically 200,000 students every year) from 2009 to 2012. From this source we retrieve each student's exam outcome (scores and whether the student passed or not), track (theoretical, technological or vocational), date of birth, gender, and the county, locality, and school of enrollment.[18]

ii) Administrative data covering the universe of students admitted to high schools. This data contains information on each student's high school and secondary school, the average scores in 5th through 8th grade, and the average scores on the 8th grade national standardized exam. We employ data from 2005-2008 of the same students who took the Baccalaureate in 2009-2012 (about 600,000 students, as some postpone high school education). In what follows we will use the average scores of the four years of lower secondary school (gymnasium) as a proxy for ability, as it captures all scores in all subjects during these four years (see also Pop-Eleches, 2009).[19]

iii) Because the administrative data under i) and ii) does not cover student poverty status, we construct this measure from individual information on the students eligible for the Money for High School (MHS) public program of financial assistance for high school students from poor households, for the cohorts 2009-2012. This data contains information on all eligible students' school and family income for each year when they submitted an application. The MHS (administrated by the Ministry of Education) disbursed a monthly allowance of 180 RON (∼53USD) per student. An applicant was eligible if the gross monthly income per family member was not higher than 150 RON in the 3 months prior to applying.[20] This warrants

[18]We opted to use data from 2009-2012 because joining the 2008 Baccalaureate data and corresponding high school admission data (admission in 2004 was somewhat different than in the following years) yielded a slightly lower matching rate, causing a risk of having a selective sample of students in 2008.

[19]Moreover, the 8th grade exit exam is not a high stake test compared with the Baccalaureate, as all students in Romania are admitted to high school, diminishing the incentives to inflate this grade through corruption (for more details on the centralized transition from 8th grade to high school in Romania, see Pop-Eleches and Urquiola, 2013). Finally, following the 2011 anti-cheating initiatives and threats (installing video cameras in schools during the exam, threatening staff with dismissal), the passing rate for the Baccalaureate dropped by more than 45% in 2011 relative to before (see Borcan, Lindahl, and Mitrut, 2014), whereas the drop was much smaller (about 17%) for the 8th grade standardized exam.

[20]Students could reapply at the beginning of every school year. The MHS funds have been disbursed every year since 2004 with no limit on how many times a student can apply as long as they were eligible. However, because of the rising number of requirements, from 2009-2010 a new criterion was introduced demanding that the student have a very good school attendance rate. A little over 100 students were denied the allowance because of low attendance in 2010-2011.

the use of MHS beneficiary status as a proxy for the economic status of the students' families (which we code using an indicator for poor students). For further discussion, please see Appendix B.

When we merge the datasets i), ii), and iii), we obtain our working sample of 553,903 students for whom we have historical school data from grade 5 (beginning of lower secondary) to grade 12 (the end of high school). Additionally, in an attempt to understand the allocation of students to university studies following the anti-corruption campaign, we will merge this data with individual data from the admission to a top Romanian university from 2009 to 2012, generating a sample of 15,395 students. We discuss this data when we address this issue later in the paper.

Table 1 outlines some key statistics for our main variables, separately by year. We note that the Romanian language written exam scores (the test most amenable to comparison, as it is identical for all students and similar across years) declined from an average of 7.07 in 2009 and 7.32 in 2010 to 6.51 and 6.37 in 2011 and 2012, respectively.[21] The overall Baccalaureate pass rate declined from 85.2% in 2009 and 75.3% in 2010 to 54.9% and 51.9% in 2011 and 2012, respectively.[22] It is important to note the drop in the 2010 pass rates, in spite of the increase in corruption (see also Figure 1 below). The main explanation behind this fall, as also supported by the 2010 official report from the Ministry of Education and Borcan, Lindahl, and Mitrut (2014), lies in a few changes in the exam structure (see Appendix C, Figure C2): *a)* The oral Romanian exam, compulsory for all students, was rendered irrelevant to the calculation of the overall Baccalaureate grade (and passing). Before 2010, 99% of the students passed this exam (a minimum grade of 5), with 50% of the students receiving an implausible score between 9 and 10 (out of 10). *b)* One elective exam was removed in 2010. Before this year around 75% of the students chose physical education for this elective test (of whom more than 90% scored a maximum score of 10).

Table 1 also shows that the share of poor students (as proxied by the MHS recipient status) is relatively stable across years (about 22%), while the number of males taking the exam decreases slightly (from about 50% in 2009 to 45% in 2011). Furthermore, higher ability students, as proxied by students with an above median

[21]The increase in 2010 is discussed in Borcan, Lindahl, and Mitrut (2014) to be a direct consequence of the 2010 public sector austerity measures and the sudden increase in corruption related to the Baccalaureate.

[22]Note that the higher pass rates in our descriptive statistics tables compared with the national averages are due to the fact that we do not include exam re-takes (i.e., instances where a student who has failed the exam in previous years re-takes the exam) in these numbers (or in the estimations). However, when we repeat our analyses including the exam repeats, the results are essentially the same, just slightly larger in magnitude.

5th-8th grade score, seem to be proportionally more numerous in 2011 and 2012. This apparent (positive) change in the composition of test takers indicates that our results could actually be a lower bound of the true effects of the anti-corruption campaign.

4 Graphical Evidence

We start with an illustration of the evolution of exam outcomes over time in Figures 1 and 2. These figures summarize the essence of our findings. Figure 1 shows the 2004-2012 pass rates and written Romanian averages, separately for early and late installers. The notable patterns are: *1)* in both early and late implementers, the Romanian written scores and the overall pass rates dropped quite sharply in 2011,[23] suggesting that the anti-corruption campaign as a whole was effective in both types of counties and that the part of this campaign that included threats of punishment played the largest role; *2)* the drop in performance in early implementation counties is larger in 2011 than in late implementation counties, suggesting that monitoring per se was effective; *3)* while early implementation counties display constant performance levels in 2012 relative to 2011, the score in late implementation counties continue to drop in 2012, reaching levels below the early implementers. This suggests that monitoring had an effect not only if introduced in combination with punishment (as was the case in 2011 for the installing counties), but also in situations where punishment for corruptive behavior had been in place for a year; and *4)*, the score for the late implementers continued to drop in 2012, when the objective monitoring was introduced and were reaching levels below the early implementers, even though the late implementers had higher scores in 2010. This may indicate that the late and early implementers may differ along some characteristics, suggesting the need to account for self-selection into treatment.

These patterns are perfectly preserved in Figure 2, which displays the evolution of scores from 2009 to 2012, separated by ability, gender and poverty status (Figures 2A, 2B and 2C, respectively). Figure 2A reveals very similar score evolutions in early and late implementers for high-ability students (above the median 5th-8th grade graduation score). Low-ability students start off at much lower scores and sustain a much sharper drop when corruption-fighting strategies are in place. The same applies for the performance of male relative to female students. The most striking contrast is perhaps between poor and non-poor students. The score dip

[23] As we explain in the data section, there is a drop already in 2010 for pass rates because of changes in the exam structure.

associated with camera monitoring in 2011 is larger for the already worse-off poor students. Overall, the graphs indicate that the camera monitoring was effective in reducing the cheating and fraud opportunities, particularly for groups prone to engage in corrupt behavior. However, the pattern observed is also that the dispersion in exam outcomes between groups increased, and those who performed poorly on the Baccalaureate prior to the corruption-fighting measures did even worse after their introduction. Note also that the patterns across groups are very similar regardless of whether or not we look at the overall change from 2010 to 2012, or if we investigate the pattern related to the installation of CCTV cameras. This suggests that we can potentially use the well identified estimates from the effect of camera installation to draw inference about how groups fare before and after the introduction of corruption-fighting measures.

In what follows we test the camera impact for the entire sample and by groups more formally. It should be noted that later installers were, on average, better off before 2010 and that we observe parallel trends before 2011 in early and late implementers for average pass rates but that the trends converge somewhat for the written Romanian score. We discuss issues of selection into camera treatment in the identification section below. In the estimations we will also add student controls in order to control for possible compositional sample changes over time across treatment and controls. In addition, we will present results from placebo regressions, using outcomes from a no-stakes exam. We argue that if we do not find an effect (of monitoring or of the threat of punishment) on such a low-stakes outcome, it will be a strong indication that our main results are unlikely to be driven by unobservable factors that could have potentially affected achievement even in the absence of anti-corruption measures. Additionally, we will show that possible differential pre-reform trends do not affect our main results.

5 Estimation strategy

To assess more formally the impact of corruption-fighting measures on exam outcomes, we employ a difference-in-differences (DD) strategy. In particular, we use the variation between counties and over time in the installation of CCTV cameras to separate out the effect of actual monitoring from the effect of harsher punishment captured by the 2011 and 2012 year indicators. The general specification is:

$$y_{ict} = \alpha + \beta T_{ct} + \gamma' \cdot X_{ict} + \varphi_t + \theta_c + \epsilon_{ict}, \tag{1}$$

where i indexes a student attending a school in county c in year t. y_{ict} is one of our two main outcomes of interest, i.e., 1) the score on the standardized written Romanian language exam and 2) an indicator equal to 1 if the student passed the Baccalaureate exam and 0 otherwise; T_{ct} is an indicator equal to 1 if the student is CCTV monitored (for all counties in 2012 and for 25 counties in 2011) and 0 otherwise; X_{ict} includes indicators for gender, for whether the student comes from a poor family, for the graduation score prior to entering high school as a proxy for student ability (as described in Section 3), for high school track and for rural area; φ_t includes year indicators; and θ_c includes 41 county indicators. In some of the estimations we replace the county indicators with a full set of school or family indicators. In all regressions we cluster the standard errors at county level, since the treatment implementation is county-wide (resulting in 42 clusters).

The DD estimate, $\hat{\beta}$, will capture the impact of CCTV installation on exam scores, based on the variation in exam outcomes within counties over time (after vs. before camera installation). Since no county had cameras installed in 2009-2010 and some counties installed them in 2011 and the rest in 2012, this estimate will be a weighted average of the exam score effects for those installing cameras in 2011 and 2012, respectively.

The 2011 and 2012 year coefficients are of interest since they capture the shift in exam outcomes relative to earlier years, net of the impact of cameras. However, these indicators can be interpreted causally only under the very strong assumption that the sole source of variation in exam outcomes 2011-2012 relative to before is the corruption-fighting campaign. This is obviously a restrictive assumption as a number of other factors might have changed across years, e.g., different changes as a result of the overall economic situation. To investigate the plausibility of this assumption, we estimate equation (1) using as outcome the scores from the low-stakes oral Romanian exam. This exam is also part of the Baccalaureate and covers the same topics as our main outcome, i.e., the written exam, but does not count towards the overall grade and there is consequently no scope for corruption. Hence, in our model using performance in this exam as the dependent variable, the year indicators' coefficients can be read as pure year effects. If the estimates for the year indicators and the DD indicator are zero, we believe we can make a reasonably strong argument for an interpretation of the year indicators as saying something about the overall impact of punishment threats. This is especially likely since the changes in exam scores (as shown in Figures 1 and 2) are so large it would be unlikely to find other factors that could explain this whole shift. Yet, we need to be cautious when interpreting the year effects as effects of the anti-corruption policy

(see Section 6.1.2 below). Similarly, when we estimate equation (1) separately by sub-groups, we focus on comparing the resulting estimates across these groups. The identifying assumption is then that there are no other factors that could explain, e.g., a diverging pattern.

Finally, the question of self-selection of counties into the CCTV monitoring treatment warrants some discussion. Since the CCTV surveillance was not enforced in 2011, county inspectorates had the final decision on the matter. The choice not to install cameras was typically motivated by lack of funds. Thus, any claim of random assignment into camera treatment would be untenable in this context. To learn more about the selection into exam monitoring, we look at the mean differences in outcomes and controls between early and late installers in the pre-reform years 2009-2010 (Appendix C, Table C1). We learn that student ability or performance does not differ across counties, and neither do our survey-based proxies for corruption norms.[24] Yet, on average, early installment counties seem to have significantly fewer poor students and be slightly larger. This supports the official justifications and also reassures us that the factors affecting the monitoring decision are accounted for in our baseline regressions. Under the assumption that county fixed effects or specific time trends account for any unobserved county-level characteristics related to the camera decision, poverty, and the observed exam outcomes, the DD estimator yields the causal impact of the CCTV monitoring on exam outcomes.

6 Results

Our aim is two-fold. First, we aim to assess the impact of the corruption-fighting campaign and particularly to understand the separate impact of the campaign mechanisms, i.e., monitoring and increased threat of punishment (Section 6.1). Second, we inquire about who benefits and who loses from curbing corruption by looking at the heterogeneous effects of the campaign on the high-stake Baccalaureate exam scores for poor vs. non-poor, males vs. females, and most vs. least able students (Section 6.2). Additionally, we attempt to understand whether the transition from a system with unhindered corruption (2010 and before) to one where corruption opportunities should be drastically reduced (2011 and 2012) by the anti-corruption

[24] We compute a proxy based on the share of people having an informal network, at the county level, using the answers to a question from the 2007 Romanian Barometer of Public Opinion: *"Is there anyone (i.e., informal network) that could "help" you solve (i.e., informally): issues in court/trials, medical problems, city hall, police, or issues related to the local authorities."* We also compute of a proxy for the level of confidence in justice, based on perceived trustworthiness of the justice courts, elicited in the same survey.

campaign significantly changed the composition of students admitted at an elite university (Section 7).

6.1 The Overall Impact of the Campaign

6.1.1 The effect of installing CCTV cameras

Table 2 presents results from estimating equation (1) for the scores on the written Romanian exam, a standardized test that has the same structure across years and tracks (columns 1-2), and for the probability of passing the Baccalaureate exam (columns 3-4). In columns (1) and (3) we only include the CCTV monitor indicator, year indicators (base is 2010), and county indicators, while in columns (2) and (4) we add the controls described previously.

We note in column (1) that the written Romanian score decreases by about 0.22 points due to camera monitoring, which is equivalent to a 0.12 SD decrease in scores on the Romanian exam relative to the sample mean. The CCTV camera effect remains very similar in column (2) when we include the rest of our control variables. For the probability of passing the Baccalaureate exam, the main results show a similar pattern as for the written Romanian exam. In particular, the impact of CCTV camera monitoring lowered the probability of passing the Baccalaureate by around 8.3 percentage points. We also note that, relative to 2010, the 2011 and 2012 year indicators clearly exhibit much lower values.[25] Yet, at this point it is difficult to assess whether these negative coefficients indicate a response to punishment threat or some other changes. We provide details on the effect of the punishment threat on exam outcomes in the next subsection.[26]

There are several concerns related to whether the CCTV monitoring in Table 2 above can indeed be interpreted as the effect of the campaign exclusively due to increasing monitoring. In particular, the negative impact of monitoring on test scores may reflect not only corruption fighting per se but (also) test anxiety from the newly introduced CCTV cameras. While we cannot fully dismiss this possibility, we believe that anxiety from monitoring would not account for such a large drop in scores. In the same line, Bertoni et al. (2013) show that the negative impact of the

[25]One apparent surprising result is the negative coefficient of the 2009 indicator for the written Romanian exam score. The reason for this pattern is the escalading corruption in relation to the Baccalaureate grades, which, especially for the written Romanian exam, peaked with the 2010 exam following a 25% wage cut for all public school educators as shown in Borcan, Lindahl, and Mitrut (2012).

[26]We also note that the 2009 year indicator is positive when we look at the probability of passing the Romanian exam. This is because the probability of passing drops already in 2010 due to additional changes in the exam structure/passing requirements as discussed in Section 2.

presence of an external examiner on test scores is due to reduced cheating rather than to anxiety. Moreover, the evidence from the psychology literature (Chapell et al., 2005) indicates that females display higher levels of anxiety during tests than males, while we will show in the heterogeneity analysis that males perform worse compared with females following the campaign.

We also address some additional concerns in alternative specifications which introduce tighter controls (county specific trends, school and family fixed effects), all of which leave the results unchanged. To save space, we report and discuss these robustness tests in Appendix 1.

Our results seem to indicate that monitoring lowered the exam scores as a result of reduced ability to engage in petty and mass in-class cheating, which had been possible in the past, as discussed in Section 2.2, subsequent to *collective bribes* being paid to the exam committee members. Yet, we cannot fully exclude that, even in the presence of CCTV monitors, some students would resort to *individual bribes* (before/after the exam takes place). We will return to this point in our heterogeneity analysis.

6.1.2 Can we separate out the punishment from the overall effect of the campaign?

We have shown that the CCTV monitoring part of the campaign indeed had an impact on curbing corruption as it resulted in statistically significant lower test scores and pass rates. Interpreting the year effects in Table 2 as showing the effects of the threat of punishment, net of increased monitoring, is much more problematic. To convincingly establish that the threats of prosecution for teachers and re-take restrictions for students were credible enough to reduce corruption, we would ideally like to contrast the written exam with a no(low)-stakes exam with no scope for fraud and thus no impact of the anti-corruption campaign on the scores. This test would be more compelling if this exam's intrinsic features were comparable to the high-stakes exam that it is compared against. Conveniently, the Romanian language is tested both via an oral and a written exam during the Baccalaureate, both covering the same topics.[27] However, since 2010, the oral exam has been rendered irrelevant for the calculation of the overall Baccalaureate score and converted to an objective aptitude test, which students cannot fail, but in which they can qualify as an "excellent," "good," or "sufficient" language user (performance levels are marked by a score of 3, 2, and 1, respectively). As the same skills are required for the two

[27]The Romanian language exam covers the same topics from all four high school years; the oral exam takes place a couple of weeks before the written.

exams but the written one is a high-stakes while the oral is a no-stakes, the oral exam is the ideal placebo test described above. To make the Romanian written and oral exams comparable we start by translating the latter exam scores, available only on a non-cardinal scale, into percentile ranks using the data from 2010-2012.[28] Next, we standardize both the percentile rank oral Romanian scores and the written Romanian scores (mean zero, standard deviation one) for the 2010-2012 cohorts.

We report the results from this exercise in Table 3. Note that the structure of Table 3 is somewhat different from Table 2. In columns 1-4 we show results for the written Romanian exam and in columns 5-8 for the oral Romanian exam, both standardized.[29] Columns 1 and 5 show results from a simple specification with only (except for county fixed effects) an indicator variable equal to 1 in 2011 and 2012 (*After11*), when the corruption-fighting campaign was in effect, and zero in 2010, when it was not.[30] In column 1, we find that the scores on the written Romanian exam decreased sharply in 2011-2012 relative to 2010, which is line with the graphical evidence in Section 4. The drop is equivalent to about one-third of a standard deviation. When we look at the oral exam (column 5) we find instead a small *increase* in scores. This suggests that the impact of the overall campaign in curbing corruption is real.

Next, in columns 2 and 6 we also add the camera indicator in an attempt to tease out the effect of increased monitoring from the overall effect of the campaign. In these specifications, we argue that the 2011-2012 indicator captures the impact of the threats of punishment for the written Romanian exam. In column 2 we see that the monitoring did make up a non-trivial part, about one-third, of the overall campaign effect. The DD estimate is statistically insignificant and very small when we look at the oral exam. We also estimate the model allowing for separate year effects (as in Table 2) and see again that the 2011 and 2012 year indicators for the oral exam have the opposite sign and are much smaller in magnitude than those for the written exam, confirming that performance was not negatively affected by

[28] We use percentile ranks since the oral exam is expressed on an ordinal scale. This is a useful transformation because if, for instance, the distribution of scores is such that there are relatively few students with a level 3 score, then these students get a higher rank score. Note also that, since we also want to compare the estimates for the year indicators, we rank the scores using all three years combined.

[29] Because the oral Romanian exam is a no-stake exam, the use of CCTV cameras was optional (even in 2012), and actually very few schools monitored this exam. We do not know which schools had CCTV cameras during the oral Romanian exam.

[30] We experimented with an regression discontinuity design using birth months as running variable, hence just adding month of birth to the specification in columns 1 and 5. This generated similar results for the 2011-2012 treatment dummy. However, the sensitivity analysis revealed problems with endogenous location around the cutoff due to a very non-strict rule of when during the year a pupil could start school.

a general year trend; if anything, scores may have actually increased, in which case the 2011 effect for the written exam may be underestimated. By association with the baseline findings, these results tell us that the year effects do not seem to explain the negative 2011 change in written exam scores.

Overall, these results lend support to our hypothesis that the increased threat of punishment brought by the campaign has curbed corruption, as seen in the drop in scores. Importantly, the drop in scores in non-monitored counties supports this hypothesis, but this does not imply that the incentive effect is independent from that of monitoring. Given that these counties would have plausibly also perceived a higher detection risk, even in the absence of cameras, the punishment threat came into effect because it was enabled by enhanced monitoring. The campaign was even more effective when the probability of detection was even higher, in the presence of CCTV monitoring.

6.1.3 Additional evidence on the effectiveness of monitoring combined with punishment

The setup we have does not include a situation where monitoring increases in the absence of punishment. However, as theory and a few empirical studies suggest, increasing the probability of detection is unlikely to work without increasing the costs to being detected, and equally, punishment is ineffective if the chances it will be applied are very low. To offer additional evidence that monitoring and punishment are most effective if combined, we take advantage of a similar policy in Moldova, a neighbor country with a very similar education system as Romania,[31] facing similar corruption problems in connection with the high-stake Baccalaureate exam. Inspired by the Romanian anti-corruption policy, a crackdown on Baccalaureate corruption in Moldova started in 2012, when the Ministry of Education obliged students to sign a special document just before the exam confirming that they are free of any additional source of cheating (mobile phones, books) during the exam and if caught with any source of cheating they would be banned from the exam for at least a year, regardless of whether they used the source or not. If caught taking bribes or letting the students cheat, teachers would also be punished.[32] In addition to these

[31]In Moldova more than 76% of the population speak Romanian as their native language, and the Baccalaureate, which is very similar to the one in Romania, includes also a Romanian language written test.

[32]There was no clear rule but the methodology stipulated that the punishment would be according to the Moldavian Labor Code. In addition, in 2012 the methodology introduced a recommendation to install CCTV cameras, but this recommendation was not followed ("The video cameras *may* be introduced in exam centers"). (Source: The Baccalaureate Methodology for the organization of the 2012 Baccalaureate exam, section IV, article 50.)

punishment threats, a new methodology prescribed mandatory installation of CCTV cameras in all exam centers in 2013.[33] This roll-out is similar to the Romanian anti-corruption campaign, but with a one-year lag; yet the threat of punishment in 2012 in Moldova was not as credible as it was in the 2011 Romanian case (due to Romania's unprecedentedly high number of trials related to the 2010 exam fraud). Therefore, we expect a less significant drop in pass rates in 2012 in Moldova relative to the large drop in 2011 in Romania.

Figure 3 shows a comparison of the evolution of pass rates in the two countries from 2007 to 2013. The Moldavian pass rates were still anchored at above 90% until 2012, while in Romania, where the campaign was well underway in 2011, the pass rates were drastically reduced to 44%. Moldavian pass rates sustained a mild drop in 2011 and 2012, reaching 88.3% in 2012. However, when the CCTV cameras were introduced in 2013, we note a 20% drop in pass rates in Moldova (reaching 68.3%). The figure suggests that the intended effects of the anti-corruption campaign were felt in both countries when a high level of monitoring coupled with punishment was reached (in 2011 and 2012 in Romania and in 2013 in Moldova).

We conclude that monitoring and punishment interact with each other and, more specifically, that monitoring enables and enhances the effectiveness of punishment.

6.2 Heterogeneous Effects of the Anti-Corruption Campaign

After having established that the campaign had a drastic effect on the test scores and probability of passing the Baccalaureate high-stake exam for the average student, in this section we focus on the efficiency and distributional side effects of curbing corruption and, in particular, look at the heterogeneous impact for students who differ in ability, poverty and gender. We already saw from the evolution of scores in Figures 2a-2c that the groups particularly affected by the camera policy are the same groups for which the exam outcomes dropped the most from 2010 to 2012: poor, low-ability, and male students.

To lend additional credibility to these findings we now turn, in Table 4, to a regression analysis using the DD approach as specified in equation (1), but now applied to sub-samples of students that differ in background characteristics: ability (columns 1 and 4), poverty status (columns 5 and 8), and gender (columns 9 and 12).[34] We estimate separate regressions for low- and high-ability,[35] poor and non-

[33]The Baccalaureate Methodology for the organization of the 2013 Baccalaureate exam, section IV, article 48.

[34]Results with school indicators are very similar.

[35]We divide students into high and low ability, according to an average graduation score above/below the median 8.81.

poor,[36] and male and female students, respectively. To save space, we only show results for the full specification, including the camera and year indicators, as well as additional controls. The estimates for the camera and year indicators from specifications that do not include controls for background variables are almost identical. This is reassuring as it means that mean-reversion is not driving the differences in results across groups.

Focusing on the camera effect, we find larger negative effects for low-ability students than for their high-ability peers. For the written Romanian exam (columns 1-2), the exam monitoring resulted in 0.354 unit lower test scores (about one-sixth of a standard deviation), which is three times as large of a drop as seen for high-ability students; for the pass rates we also see difference between high and low-ability students, but it is smaller and not statistically significantly different. Next, the results for poor and non-poor students confirm again the graphical analysis from Section 4: as a result of the camera policy, poor students'test scores decreased twice as much and pass rates about 50% more compared with non-poor students. For males and females we see a smaller difference, although the directions of the effects reconfirm that those doing worse pre-campaign lose more.

Next, we attempt to draw some inferences about changes over time, from a fully corrupt (in 2010) to a less corrupt system (in 2012 when the anti-corruption policy was fully implemented). We use a similar placebo test as in Table 3 for the years 2010-2012. Table 5 shows results separately by ability (Panel A), poverty (Panel B) and gender (Panel C) for the low-stakes oral and the high-stakes written Romanian exam. As in Table 3, for comparability, both outcomes are standardized results. The magnitudes of the differences in the estimates across groups are qualitatively similar as in Table 4, although here expressed in standard deviation units. The pattern of estimates for the oral exam is quite striking, as the estimates for the camera indicator and the year indicators are very similar across groups regardless of whether we compare high and low ability, poor and non-poor, or males and females. The only exception is for the oral exam in 2012, where low-ability students scored lower than high-ability students (conditional on camera implementation). However, this difference is still half as big as the difference for these groups on the high-stakes exam. This is very reassuring and suggests that the difference in year effects across groups for the Romanian written exam can credibly be said to reflect the overall effect of the campaign, as there are no comparable differences between pre- and post-campaign years for the low-stakes exam.

[36]Poor students are defined according to MHS recipient status; see the discussion in Appendix B.

From the estimates reported in Table 5 we can therefore conclude that the drop in Baccalaureate high-stakes exam scores caused by i) the camera installation is about 0.18 SD larger for the low-ability students, 0.13 SD units larger for poor students, and about 0.08 SD units larger for males; ii) the corruption-fighting initiatives, such as threat of punishment (net of the CCTV monitoring), is about 0.20 SD larger for the low-ability students, 0.06 SD units larger for poor students, and about 0.05-0.10 SD units larger for males; iii) the combined effects of the anti-corruption campaign one year after implementation (in 2012 vs. 2010) is about 0.43 SD larger for the low-ability students, 0.19 SD units larger for poor students, and about 0.18 SD units larger for males.[37] Hence, both the monitoring and the punishment component of the anti-corruption campaign reduced the corruption opportunities in the high-stakes tests more for poor and low-ability students, driving their larger drop in scores between 2010 and 2012. We conclude that our estimates show that disadvantaged students became even worse off following the corruption-fighting initiatives. While in the case of ability the campaign revealed the true standing of students, in the case of poor students the campaign may have had adverse effects an issue explore further below.

Figure C1 (Appendix C) provides an alternative way of presenting the heterogeneous effects, by displaying the written Romanian exam score distributions separately by group (and by subgroup), for 2010 (unhindered corruption) compared with 2012 (little or no corruption). The distribution by ability (in Figure C1[a]) shows the high-ability students to the far-right of the distribution scores; the score distribution for this group is slightly flatter in 2012 than in 2010. In contrast, the low-ability students'score distribution becomes flatter and also less spread out, with a larger spike at 5 (the passing threshold) in 2012 relative to 2010. Figure C1(b) shows that males are worse off in 2010 and their situation further deteriorates in 2012. Next we consider the differences between poor and non-poor students. The score distributions by poverty status (Figure C1[c]) display a large frequency shift from high to low scores in 2012 relative to 2010 for both poor and non-poor, but more pronounced for the already disadvantaged poor students.

To conclude, the heterogeneity results shown in this section indicate some interesting findings. First, in line with our initial hypothesis, the high-ability students, even after controlling for their poverty status, seem to benefit relatively more from a system with little or no corruption. This is not surprising as high-ability students should be less reliant on cheating or paying bribes to pass the exam. The

[37]We do not show specifications with only year effects for 2012, but the total effects of the campaign in 2012 vs. 2010 can simply be obtained from the table by adding the camera and year 2012 estimates for each group and then comparing these sums across groups.

differences in CCTV monitoring and the year effects between students of different abilities are quite large and statistically significant, hinting that the low-ability students had relied to a larger extent on cheating and/or other means of fraud before the campaign.

Secondly, we show that the poor students may perform worse in a less corrupt system. This is actually not in line with our prior that, while both poor and non-poor students would benefit from collective bribes and cheating, the non-poor would additionally benefit from individual bribes as this form of corruption requires sizeable bribes (money) and access to the corruption networks (see discussion in Section 2.2). Intuitively, the non-poor should be able to afford the required payments, as well as gifts and private tutoring with in-class teachers. Moreover, those from a privileged economic background typically also enjoy a high social standing, which should grant them easy access to the nepotistic networks.[38] Overall, if the campaign would eliminate both forms of corruption, we would expect the non-poor students to lose more in a non-corrupted environment.

So what could explain the wider score gap between poor and non-poor students? One potential concern is that the proxy for poverty reflects not only socioeconomic status but also some potential effect of the MHS program (used to define poverty status) on the recipients. In order to ensure that this is not the case, we compare students just below with students just above the cutoff income for receiving MHS in 2005-2006, which was the only year when funds were short of the demand and some eligible students did not receive the money (to save space, the details on the data and strategy for this test are shown in Appendix B). The RD estimate of the treatment effect is insignificant, indicating that a potential MHS treatment is not a concern here. We therefore proceed to discuss some other potential channels leading to the observed increased score gap:

i) *Increased private tutoring or parental investment for the non-poor.* To rule out the private tutoring channel we consider additional data from the 2010-2012 Romanian Household Budget Survey and observe no increase in private tutoring for high school students in 2010-2012. It is also possible that parents of non-poor students may have substituted bribes for more time spent working on homework or exam preparation. This is less likely to have generated a large effect in 2011, as the camera policy was implemented in May, leaving very little time for extra preparation. However, in 2012 this behavioural effect could partly account for our result.

[38]Note that this is also in line with the generally lower performance of poor students relative to non-poor before 2011.

ii) *Stronger cheating norms for the poor.* One way to dismiss this channel is to look at the share of students eliminated from the exam (caught cheating) (see Table C2), which before 2011 shows no difference between rich and poor students.

iii) *Collective vs. individual bribes.* We believe that one key to understanding the detrimental effects of the campaign on the poor lies in the various mechanics of the bribing process. If poor cheat as much as rich, without being able to afford bribes, the poor students'ability to take part in the fraud can only come as a result of free-riding. A good candidate explanation for this opportunity lies in the mechanism of *collective bribing*, which is essentially used to provide a "public good." If some students contribute, the benefit is collective and everyone, including poor students, can take advantage of the slack proctoring. Given some level of ability, the annihilation of cheating practices (likely coupled with particular unobserved traits, like motivation and the educational investment of poor students throughout high school) generates lower results for the poor students. This implies that monitoring and punishment reveal wide pre-existing inequalities, previously concealed by corruption. A complementary explanation may lie in that only richer students can afford *individual bribes*. Recognizing the existence of a well-developed market for bribes, the poor student could not afford the required amounts or services. Moreover, following the implementation of the anti-corruption campaign it is likely that teachers could have substituted collective for more *individual bribes*, pricing out the poor students.[39] It is unlikely that monitoring and punishment threats can fully eradicate this form of corruption, as revealed by further anecdotal evidence from crackdowns on corruption in some exam centers in 2012 and 2013.

7 The Short-Term Impact of the Anti-Corruption Campaign: Preliminary Evidence from Admission into an Elite University

As revealed in the heterogeneity analysis, the corruption-fighting campaign led to a separation of ability types and a reshuffling of the students in the score distribution, by income. These changes may have direct implications for the selection of students into higher education.[40] In this section we document the short-term consequences

[39]This displacement effect has been documented in the CCTV and crime literature. See Priks (2015).

[40]The total number of students in higher education (university) decreased from 775,319 in 2009 to 464,592 in 2012. The biggest drop took place at the private universities (from more than 300,000 to less than 100,000 students in four years), while the number of students enrolled (regardless of year of study) at public universities decreased from about 452,892 in 2009 to 364,916 in 2012.

of the anti-corruption campaign by using admission data from an elite university in Romania.[41] This university admits about the same number of students every year; all admitted students are ranked according to an overall score and the top 55 to 65% are exempt from the tuition fee (*la buget*), while the rest pay a monthly fee.[42] We have the following information for the admitted students at this elite university from 2009 to 2012: the overall Baccalaureate grade, the overall high school grade, name and date of birth, the county and school they come from, whether or not they are tuition exempt, and home department at the university. We merge this information with our main data by name, date of birth, gender, county and the Baccalaureate grade and end up with a sample of 15,395 admitted students with a full education history.[43] In what follows we label the group of tuition-exempt students "top students" and the group of tuition-paying students "good students."

To understand whether there is any change in the composition of students admitted in the elite higher education due to corruption-reducing measures, in Table 6 (Panel A all students; Panel B top students; Panel C good students), we provide estimates from regressions based on equation (1), but where the dependent variable is student's *ability* (columns 1-3), *poverty* (columns 4-6), and *male* (column 7). We cannot infer anything from the changes in these outcomes across years for two reasons. First, since the mean 5th-8th grade score, which we, as before, use as a proxy for ability, increased from 2010 to 2012, there is an increased likelihood that among those admitted to the elite university, we will observe a higher mean ability score over time. The same argument holds when we look at the changed composition of the top 20% Baccalaureate performers below. There is also a slight change in the fractions of poor and males over time. Second, the admission rules changed slightly every year.[44] We therefore standardize the ability score to have mean zero and SD

[41]This is one of the oldest and highly regarded universities in Romania, with a long tradition of attracting elite students from all over the country. Students admitted here are usually in the top 15% of the overall high school scores and Baccalaureate grades. The proportion of accepted students coming from CCTV-monitored counties is about 77%.

[42]The number of students admitted to the university was relatively constant across years: 4,742 (in 2009), 3,792 (in 2010), 4,937 (in 2011), and 4,648 (in 2012); students are exempt from the tuition fee (la buget) contingent on the Ministry of Educations budgetary allocation each year; the remaining students need to pay a tuition fee of roughly 85 USD/month.

[43]We cannot fully merge the two data sets because of some duplicates. We do not have the Baccalaureate (they are from older cohorts), the poverty and/or ability measure for about 2,400 students. The attrition rate is however fairly constant across years, at less than 10%. Note that our final sample of 15,395 students includes 660 students who took the Baccalaureate before the university admission year (i.e., about 85% took the Baccalaureate in 2009 and 2010 and applied in 2011 and 2012, respectively). This may signal that our results are contaminated with students who got accepted with inflated Baccalaureate grades. In the regressions below we control for these students, although the results are very similar if we exclude them from the regressions.

[44]While the Baccalaureate grade remains the most important piece of the final admission score, its share changed from 50% of the admission score (in addition to 25% high school grades and 25%

one for each year in the estimations. Hence, we can only credibly separate out the effect of camera monitoring on the composition of admitted students.

The insignificant camera estimate in column 1 (Panel A) indicates that the admitted students are on average of the same ability, regardless of whether or not they were treated with additional monitoring (in addition to punishment threat). This holds both for poor and non-poor students (columns 2-3). Admitted CCTV-monitored students are 3.3% less likely to be poor than those not monitored (column 4), while a breakdown of this effect by groups below and above the 50th ability percentile (columns 5-6) shows that the disadvantage of poor students arises mainly in the low-ability group. The composition in terms of gender has not changed (column 7).

Interestingly, the results in Panel B for the top students show a clearer pattern, especially for the composition in terms of ability (column 1): the admitted students from CCTV-monitored counties seem to have a higher ability, for both poor and non-poor students (columns 2-3), even though the results are less precise for the poor due to small sample size. Also, similar to Panel A, among the top admitted students, those who were CCTV-monitored are less likely to be poor (column 4). This effect comes from the lower ability poor students (column 5). The results in Panel C for good students mirror the results in Panel A, but are not significant.

As an additional exercise, we run the same regressions on a subsample of Baccalaureate students who were in the top 20% of the final Baccalaureate scores each year. We expect these students to be the top contenders for elite universities. The estimates, displayed in Table 7, convey the same effects of the campaign on student composition that we see for the university admission sample (particularly the top, tuition-exempt students). The results reassure that monitoring contributed to an improvement in ability, but also confirm that the poor students'chances to snag the top places were significantly reduced.

Taken together, these estimates strengthen the finding that the anti-corruption campaign resulted in increased inequality between poor and non-poor students. The poor students with low-ability had significantly reduced chances of entering higher education, especially those with tuition-exempt status. Interestingly, the ability is

the university's own admission exam) in 2009 and 2010 to 67% (and 33% high school grades) in 2011 and 100% of the admission score in 2012. This change implies that the 2011 and particularly the 2012 admission scores were far less inflated than earlier, due to both the anti-corruption policy and the change in admission rules, reflecting the true composition of students. This should have led to a better composition in terms of admitted students'ability. However, another effect works in the opposite direction: the elimination of the very competitive own admission exam (potentially to attract more students) may have meant that lower ability students stood a better chance to be admitted. This may bias the camera and the 2011 and 2012 year effects downwards.

more important now, especially for the non-poor students admitted on a tuition-exempt basis.

8 Conclusions and Discussion

This paper adds a new building block to the understanding of corruption in two dimensions. Firstly, it provides additional evidence that punishment coupled with monitoring are effective in reducing corruption even in settings where the potential gains from corruption are very large. Second, it analyzes the ramifications of fighting corruption from a distributional perspective - an issue largely overlooked in previous studies.

We make use of a setting where corruption in education is rampant and has large gains for students, i.e., the Romanian national school-leaving exam, the Baccalaureate. We exploit a nationwide anti-fraud campaign that began in 2011 featuring both increased credible threat of punishment (for teachers and students) and increased monitoring during the exam. We make use of the variation across years and counties in closed-circuit TV (CCTV) exam monitoring to calculate the effect of the campaign on Baccalaureate exam scores. Our results indicate that the campaign was more effective when the probability of detection was higher, i.e., in the presence of CCTV monitoring. While the punishment component was implemented in the whole country at the same time, because of its strict implementation and the use of a placebo exercise, we can say that increased punishment brought about by the campaign has curbed corruption, as seen in the drop in test scores. We conclude that monitoring and punishment interact with each other and, more specifically, monitoring enables and enhances the effectiveness of punishment.

After having established that the campaign had a drastic effect on the test scores and on the average student's probability of passing the Baccalaureate high-stake exam, we show the efficiency and equity side effects of curbing corruption and, in particular, look at the heterogeneous impact by students' ability, poverty and gender. Not surprisingly, we find that high-ability students seem to benefit relatively more from a system with little or no corruption, as the low-ability students relied to a larger extent on cheating and/or other means of fraud before the campaign. Yet, when it comes to the poor, disadvantaged students, we show that they perform even worse in a non-(less) corrupt system, an ex-ante unexpected pattern.

Finally, we also look at the composition of students at an elite university. The results strengthen the finding that the anti-corruption campaign revealed a greater inequality between poor and non-poor students than the apparent pre-campaign

level. More exactly, we find that poor students' (with low ability) chances of entering higher education went down significantly, especially with regard to tuition-exempt admission.

An important lesson from these results is that anti-corruption programs are not a cure for all ills. In terms of inequality of opportunity, the finding that poor students do worse in a non-corrupt state is especially important for policy makers. This result uncovers the wide pre-existing inequalities between the poor and the well-off students, which corruption had only concealed. The implication is that, in addition to maintaining the anticorruption strategies, there is a need for more in-depth investigation of the differences in achievement between poor and non-poor. The implications of these findings extend to other countries, such as Moldova or Cambodia, where, similar anti-corruption measures for high-stake exams are currently being discussed and implemented, and where the initial inequality level is already very high.

References

[1] Baicker, K. and Staiger, D. (2005) "Fiscal Shenanigans, Targeted Federal Health Care Funds, and Patient Mortality," Quarterly Journal of Economics 120 (1): 345-386.

[2] Banerjee, A., Duflo, E. and Glennerster, R. (2008) "Putting a Band-Aid on a Corpse. Incentives for Nurses in the Indian Public Health Care System," Journal of the European Economics Association, 6(2-3), pp. 487-500.

[3] Banerjee, A., Mullainathan, S. and Hanna, R. (2012) "Corruption," NBER Working Paper No. 17968, April.

[4] Becker, G. S. and George J. Stigler (1974) "Law Enforcement, Malfeasance, and Compensation of Enforcers," The Journal of Legal Studies, University of Chicago Press, vol. 3(1), pp. 1-18.

[5] Becker, G. S. and Tomes, N. (1986) "Human Capital and the Rise and Fall of Families," Journal of Labor Economics, University of Chicago Press, 4(3), pp. S1-39, July.

[6] Bjrklund, A. and Jntti, M. (2012) "How Important is Family Background for Labour-Economic Outcomes?," Labour Economics, 19, pp. 465-474.

[7] Black, S. E. & Devereux, P. J. (2011) "Recent Developments in Intergenerational Mobility," Handbook of Labor Economics, Elsevier.

[8] Bertoni, M., Brunello, G, Rocco, L. (2013) "When the cat is near, the mice won't play: The effect of external examiners in Italian schools," Journal of Public Economics 104, pp. 65-77.

[9] Borcan, O., Lindahl, M. and Mitrut, A. (2014) "The Impact of an Unexpected Wage Cut on Corruption: Evidence from a "Xeroxed" Exam," Journal of Public Economics, Volume 120, December, pp. 32-47.

[10] Chapell, M.S., Blanding, Z.B., Silverstein, M.E., Takahashi, M., Newman, B., Gubi, A. (2005) "Test anxiety and academic performance in undergraduate and graduate students," Journal of Educational Psychology, 97, pp.268-274.

[11] Choe, C., Dzhumashev, R., Islam, A. and Khan, Z. H. (2013) "The Effect of Informal Networks on Corruption in Education: Evidence from the Household Survey Data in Bangladesh," Journal of Development Studies, 49(2), pp. 238-250.

[12] Di Tella, R. & Schargrodsky, E. (2003) "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires,"Journal of Law and Economics, University of Chicago Press, vol. 46(1), pp. 269-92.

[13] Duflo, E., Hanna, R. and Ryan, S. (2012) "Incentives Work: Getting Teachers to Come to School," American Economic Review, 102(4), pp. 1241-78.

[14] Emran, M.S. Islam, A. and Shilpi, F. (2013) "Admission is Free Only if Your Dad is Rich! Distributional Effects of Corruption in Schools in Developing Countries," Monash Economics Working Papers 11-13, Monash University, Department of Economics.

[15] Ferraz, C. and Finan, F. (2008) "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," The Quarterly Journal of Economics, MIT Press, 123(2), pp. 703-745.

[16] Ferraz, C. and Finan, F. (2011) "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments," American Economic Review, American Economic Association, 101(4), pp. 1274-1311, June.

[17] Ferraz, C., Finan, F. and Moreira, D. B. (2012) "Corrupting learning," Journal of Public Economics, 96(9-10), pp. 712-726.

[18] Glewwe, P., Ilias, N. and Kremer, M. (2010). "Teacher Incentives," American Economic Journal: Applied Economics, American Economic Association, 2(3), pp 205-27, July.

[19] Hunt, J. (2007). "How corruption hits people when they are down," Journal of Development Economics, 84(2), pp. 574-589.

[20] King, J., Mulligan, D. K. and Raphael, S. (2008) "The San Fransciso Community Safety Camera Program: An Evaluation of the Effectiveness of San Francisco's Community Safety Cameras," CITRIS Report, University of California, Berkeley

[21] Lance Lochner & Alexander Monge-Naranjo, 2012. "Credit Constraints in Education," Annual Review of Economics, Annual Reviews, vol. 4(1), pages 225-256, 07.

[22] Mauro, P. (1995) "Corruption and growth," Quarterly Journal of Economics, 110(3), pp. 681-712.

[23] Nagin, D.S., Rebitzer, J. B., Sanders, S. and Taylor, L. J. (2002). "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment," American Economic Review, vol. 92(4), pp. 850-873.

[24] Olken, B. (2006) "Corruption and the Costs of Redistribution: Micro Evidence from Indonesia" Journal of Public Economics, 90(4-5), May, pp. 853-870.

[25] Olken, B. (2007) "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," Journal of Political Economy, 115(2), April, pp. 200-249.

[26] Olken, B. and Pande, R. (2012) "Corruption in Developing Countries," Annual Review of Economics, 4, pp. 479-505.

[27] Pop-Eleches, C. (2009) "Abortion and Child Cognitive Outcomes," mimeo, Columbia University.

[28] Pop-Eleches, C., Urquiola, M. (2013) "Going to a Better School: Effects and Behavioral Responses," American Economic Review, 103(4), pp. 12891324.

[29] Priks, M (2014). "Do Surveillance Cameras Affect Unruly Behavior? A Close Look at Grandstands," The Scandinavian Journal of Economics, 116(4), pp. 11601179.

[30] Priks, M. (2015). "The Effects of Surveillance Cameras on Crime: Evidence from the Stockholm Subway," Economic Journal, forthcoming.

[31] Reinikka, R. and Svensson, S. (2004). "Local Capture: Evidence From a Central Government Transfer Program in Uganda" The Quarterly Journal of Economics, MIT Press, 119(2), pp. 678-704.

[32] Reinikka, R. and Svensson, J. (2005) "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda," Journal of the European Economic Association, MIT Press, 3(2-3), pp. 259-267.

[33] Rothstein, B. and Uslaner, E. (2005) "All for All: Equality, Corruption, and Social Trust Source," World Politics, 58(1), pp. 41-72.

[34] Solon, G. (1999) "Intergenerational Mobility in the Labor Market," in Orley C. Ashenfelter and David Card (editors), Handbook of Labor Economics, 3A, Amsterdam.

[35] Svensson, J. (2005) "Eight Questions about Corruption," Journal of Economic Perspectives, 19(3), pp.19-42

[36] Welsh, B. C. and Farrington, D. P. (2003). "Effects of Closed-Circuit television on Crime," The ANNALS of the American Academy of Political Science, 587, pp. 110-135.

[37] Welsh, B. C. and Farrington, D. P. (2009), "Public Areas CCTV and Crime Prevention: An Updated Systematic Review and Meta-Analysis," Justice Quarterly, 26, pp. 716-745.

[38] You, J., Khagram, S. (2005) "A Comparative Study of Inequality and Corruption." American Sociological Review 70(1), pp.136-157.

Figures and Tables

Figure 1: Baccalaureate score evolution 2004-2012, by early and late camera installation

Figure 1A. Romanian Scores

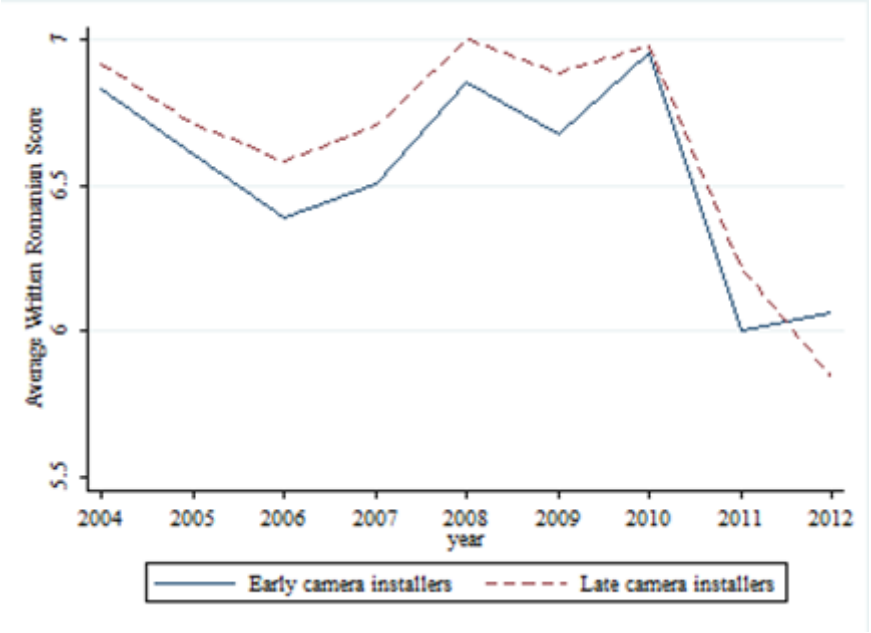
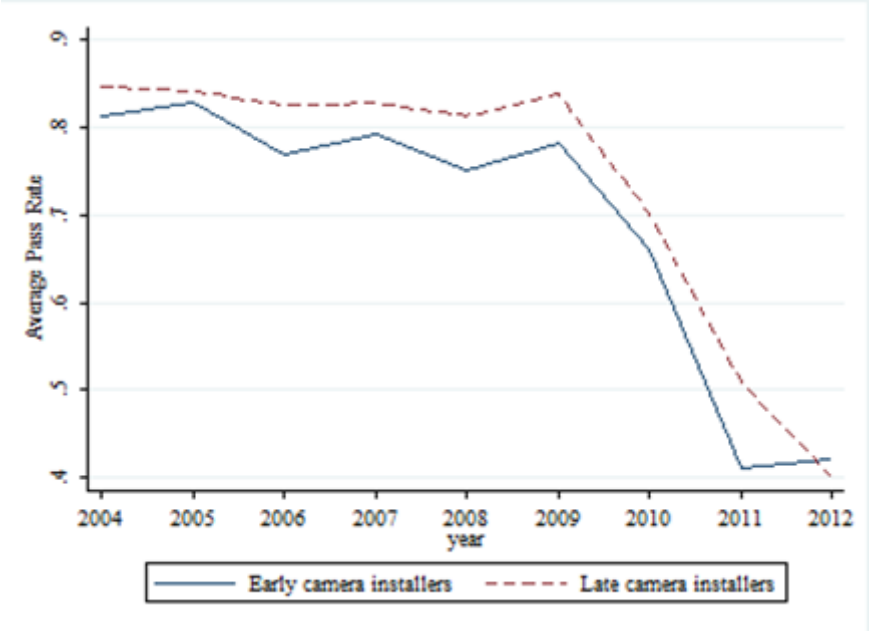


Figure 1B. Pass Rates



Notes: The figure displays the average Romanian written exam scores (left) and overall pass rates (right) separately for counties that did and did not implement the camera in 2011. The average scores are displayed on the y-axis, while the x-axis displays the years from 2004 to 2012.

Figure 2: Baccalaureate average scores, by early and late installation, and by groups

Figure 2A. By Ability

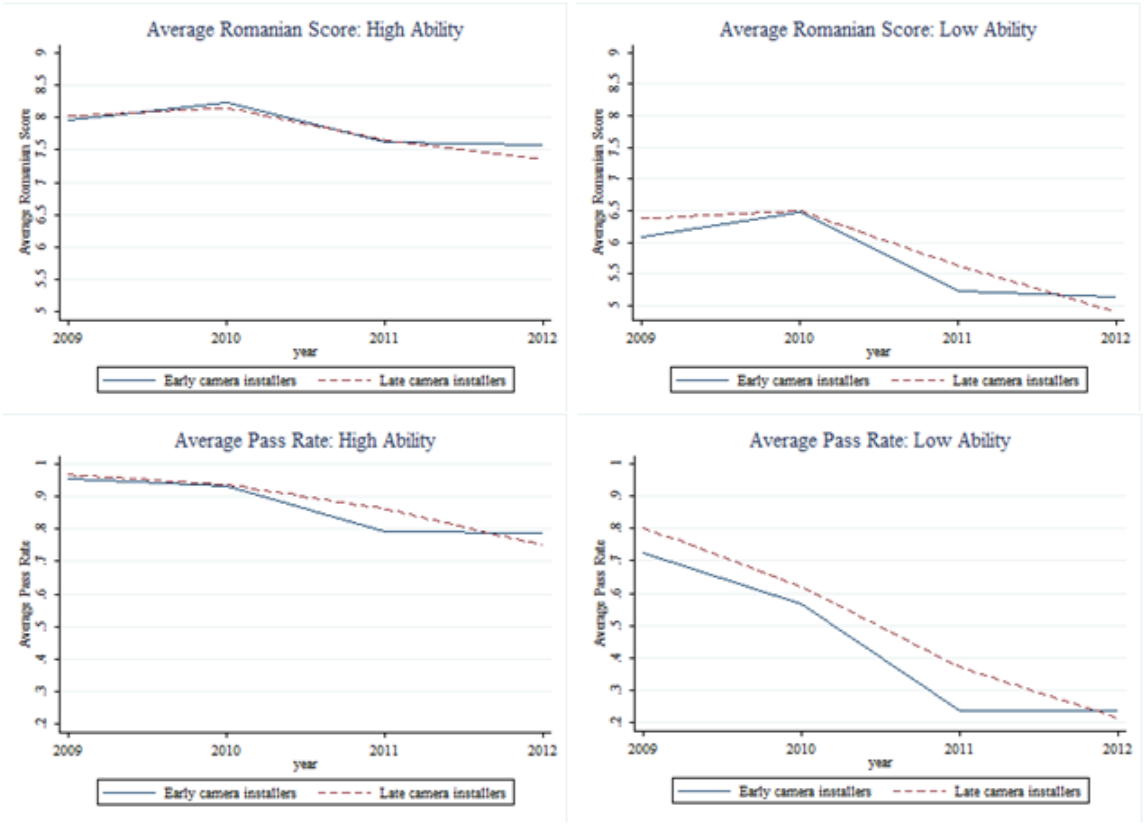


Figure 2B. By Gender

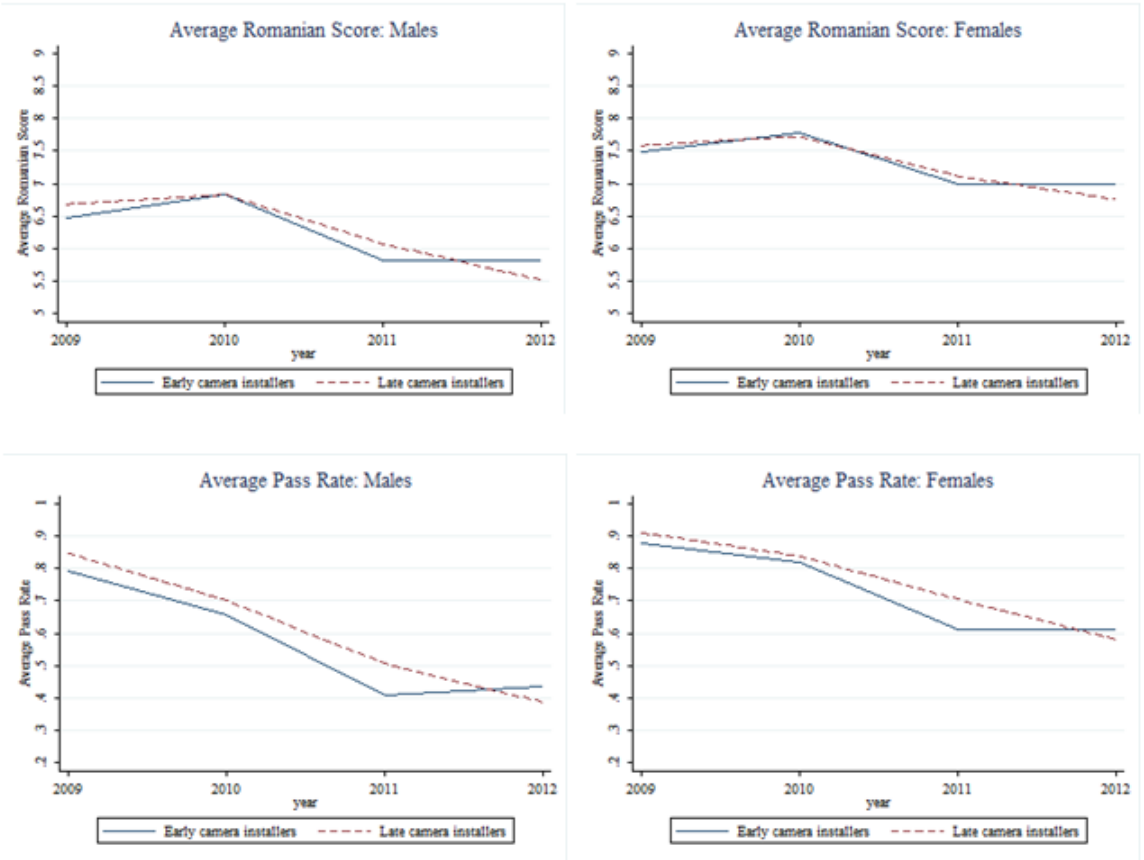
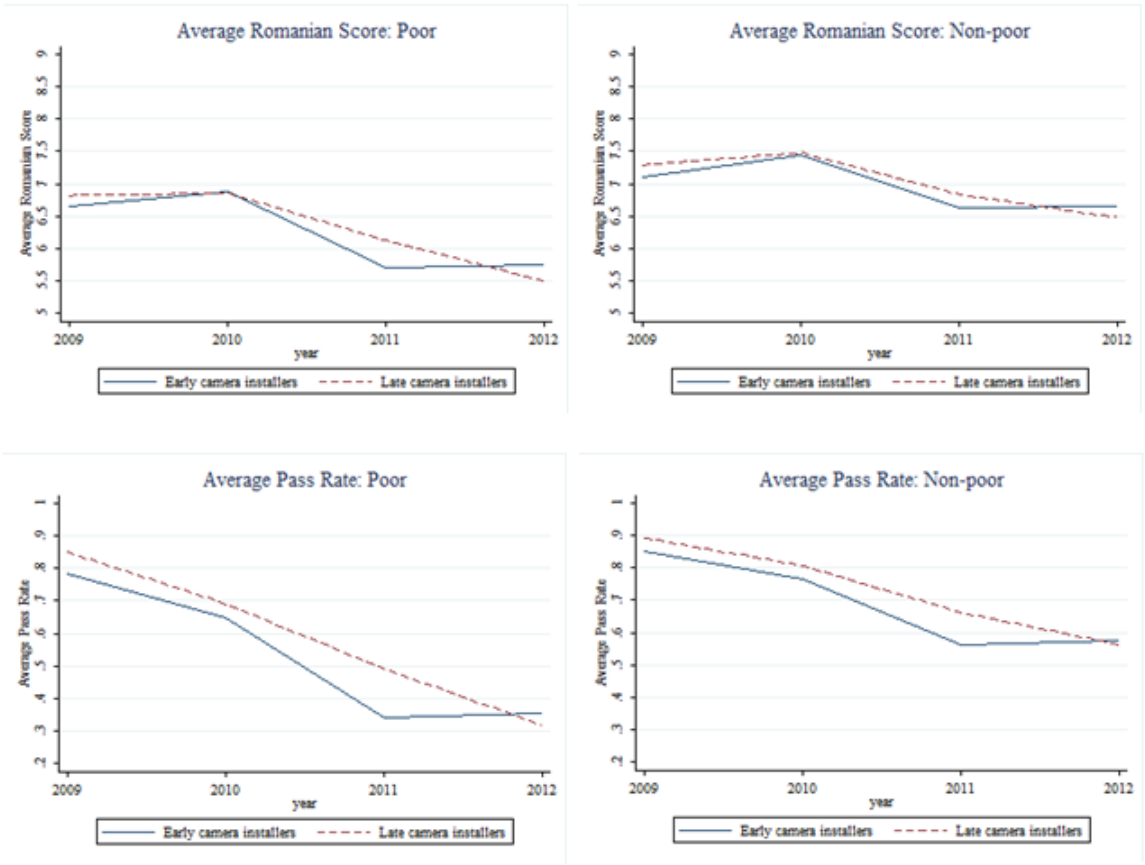
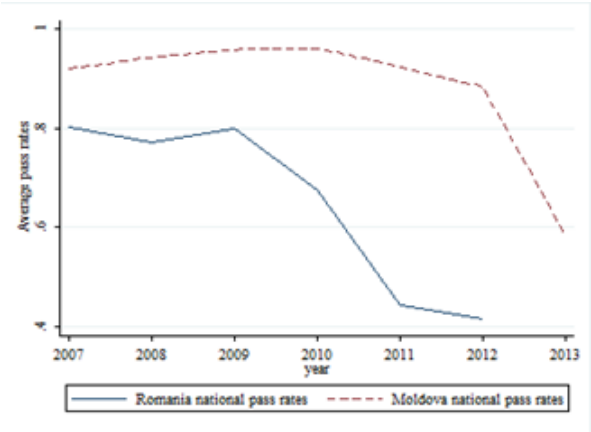


Figure 2C. By Poverty



Notes: The figures display the average Romanian written exam scores (top) and overall pass rates (bottom) by groups (by ability Figure 2A, by gender Figure 2B, and by poverty status Figure 2C) and separately for counties that did and did not implement the camera in 2011. The average scores are displayed on the y-axis, while the x-axis displays the years from 2009 until 2012.

Figure 3: Baccalaureate National Pass Rates in Romania and Moldova, 2007-2013



Notes: The figure displays the average national pass rates in Romania 2007-2012 (blue) and Moldova 2007-2013 (red dashed). The figures for Moldova are retrieved from the government website www.bloguvern.md. The figures for Romania are the authors own calculations using the available individual-level datasets (hence the 2013 figure for Romania is missing). The average pass rates are displayed on the y-axis, while the x-axis displays the years from 2007 to 2013.

Table 1: Summary statistics

	2009		2010		2011		2012	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Written Romanian score	7.073	1.769	7.323	1.570	6.510	2.007	6.377	2.065
Pass	0.852	0.354	0.753	0.431	0.549	0.498	0.519	0.500
Oral Romanian score			2.545	0.661	2.584	0.654	2.560	0.672
Percentile rank oral **			-0.038	1.007	0.033	0.998	0.006	1.002
Percentile rank written **			-3.202	0.131	-3.265	0.153	-3.273	0.153
Poor	0.200	0.400	0.222	0.415	0.229	0.420	0.227	0.419
Ability (Score 5-8th grade)	8.635	0.927	8.619	0.939	8.650	0.934	8.732	0.897
Male	0.455	0.498	0.466	0.499	0.465	0.499	0.451	0.498
Theoretical track	0.509	0.500	0.485	0.500	0.501	0.500	0.530	0.499
Rural	0.043	0.202	0.048	0.214	0.051	0.220	0.053	0.223
N*	146,576		143,380		136,902		127,045	

Notes: The table displays descriptive statistics by year for the overall sample. *The number of observations for the Romanian written and oral exams is slightly smaller; **Standardized numbers

Table 2: The impact of the anti-corruption campaign: main results; 2009-2012 academic years

	Written Romanian		Baccalaureate pass	
	(1)	(2)	(3)	(4)
Camera	-0.222**	-0.237***	-0.083***	-0.084***
	(0.088)	(0.084)	(0.026)	(0.025)
Year12	-0.719***	-0.888***	-0.151***	-0.180***
	(0.068)	(0.057)	(0.021)	(0.018)
Year11	-0.667***	-0.707***	-0.150***	-0.158***
	(0.051)	(0.052)	(0.015)	(0.014)
Year09	-0.253***	-0.297***	0.099***	0.090***
	(0.051)	(0.056)	(0.009)	(0.009)
Male		-0.483***		-0.039***
		(0.015)		(0.002)
Poor		-0.299***		-0.061***
		(0.018)		(0.004)
Ability		1.038***		0.184***
		(0.019)		(0.007)
Theoretical		0.602***		0.186***
		(0.030)		(0.010)
Rural		-0.163***		-0.037**
		(0.048)		(0.015)
County FE	Yes	Yes	Yes	Yes
Observations	547,447	547,447	553,903	553,903
R-squared	0.061	0.497	0.102	0.376

Notes: The table displays the estimates from the baseline Difference-in-Differences specifications for the changes in the Romanian exam scores and exam pass probability when the students were treated with the camera monitoring, relative to before the monitoring was introduced. Standard errors are clustered at county level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Placebo test: Written vs. oral Romanian score, standardized; 2010-2012 academic years

	High-stakes exam - Written Romanian exam				No-stakes exam - Oral Romanian exam			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Camera		-0.119***	-0.090***	-0.093***		-0.012	0.013	0.007
		(0.030)	(0.030)	(0.029)		(0.018)	(0.020)	(0.020)
After11	-0.328***	-0.230***			0.059***	0.069***		
	(0.017)	(0.026)			(0.010)	(0.019)		
Year 12			-0.271***	-0.284***			0.033	-0.032
			(0.026)	(0.027)			(0.025)	(0.026)
Year 11			-0.236***	-0.240***			0.064***	0.053***
			(0.025)	(0.025)			(0.019)	(0.019)
Male				-0.280***				-0.152***
				(0.008)				(0.010)
Poor				-0.217***				-0.150***
				(0.010)				(0.012)
Ability				0.789***				0.474***
				(0.016)				(0.014)
Theoretical				0.287***				0.164***
				(0.016)				(0.024)
Rural				-0.141***				-0.092***
				(0.020)				(0.030)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	400,088	400,088	400,088	400,088	400,088	400,088	400,088	400,088
R-squared	0.065	0.067	0.067	0.315	0.025	0.025	0.025	0.296

Notes: The table displays the estimates from the baseline Difference-in-Differences specifications for the Romanian written exam performance (columns 1-4) and the Romanian oral exam performance (columns 5-8). Both dependent variables are expressed in standardized percentile rank scores, using data from 2010-2012. The variable After11 is an indicator equal to 1 for years 2011-2012 and 0 for 2010. Standard errors in parentheses are clustered at county level.*** p<0.01, ** p<0.05, * p<0

Table 4: Heterogeneous effects of the anti-corruption campaign

	Written Romanian		Baccalaureate Pass		Written Romanian		Baccalaureate Pass		Written Romanian		Baccalaureate Pass	
	High Ability	Low Ability	High Ability	Low Ability	Poor	Non-Poor	Poor	Non-Poor	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Camera	-0.117*	-0.354***	-0.076***	-0.095***	-0.423***	-0.189**	-0.122***	-0.078***	-0.294***	-0.187**	-0.080***	-0.088***
	(0.060)	(0.110)	(0.017)	(0.032)	(0.092)	(0.083)	(0.025)	(0.024)	(0.101)	(0.072)	(0.028)	(0.022)
Year12	-0.591***	-1.124***	-0.084***	-0.269***	-0.990***	-0.854***	-0.236***	-0.159***	-1.012***	-0.787***	-0.207***	-0.157***
	(0.044)	(0.076)	(0.011)	(0.025)	(0.066)	(0.060)	(0.024)	(0.017)	(0.075)	(0.048)	(0.022)	(0.015)
Year11	-0.493***	-0.901***	-0.066***	-0.246***	-0.783***	-0.688***	-0.202***	-0.143***	-0.778***	-0.647***	-0.187***	-0.133***
	(0.040)	(0.062)	(0.008)	(0.019)	(0.056)	(0.052)	(0.017)	(0.014)	(0.062)	(0.045)	(0.016)	(0.013)
Year09	-0.231***	-0.364***	0.024***	0.152***	-0.275***	-0.296***	0.118***	0.085***	-0.353***	-0.245***	0.124***	0.062***
	(0.039)	(0.074)	(0.004)	(0.014)	(0.071)	(0.055)	(0.015)	(0.009)	(0.065)	(0.050)	(0.011)	(0.008)
Poor	-0.497***	-0.305***	-0.088***	-0.067***					-0.320***	-0.274***	-0.073***	-0.051***
	(0.020)	(0.019)	(0.005)	(0.006)					(0.018)	(0.019)	(0.005)	(0.005)
Male	-0.671***	-0.619***	-0.060***	-0.069***	-0.561***	-0.461***	-0.061***	-0.032***				
	(0.016)	(0.018)	(0.004)	(0.004)	(0.015)	(0.015)	(0.003)	(0.003)				
Theoretical	0.875***	1.011***	0.178***	0.301***	0.533***	0.606***	0.207***	0.176***	0.650***	0.558***	0.195***	0.179***
	(0.041)	(0.044)	(0.011)	(0.014)	(0.047)	(0.028)	(0.015)	(0.009)	(0.029)	(0.030)	(0.011)	(0.010)
Rural	-0.353***	-0.340***	-0.084***	-0.065***	-0.119**	-0.227***	-0.029*	-0.053***	-0.196***	-0.128**	-0.035**	-0.039**
	(0.044)	(0.065)	(0.019)	(0.017)	(0.056)	(0.051)	(0.016)	(0.018)	(0.052)	(0.049)	(0.014)	(0.018)
Ability					0.921***	1.076***	0.160***	0.192***	0.952***	1.136***	0.175***	0.194***
					(0.016)	(0.020)	(0.005)	(0.007)	(0.017)	(0.023)	(0.006)	(0.008)
County FE	Yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Obs.	276,905	270,542	277,618	276,285	118,812	428,635	121,374	432,529	250,222	297,225	254,360	299,543
R-squared	0.200	0.244	0.137	0.294	0.449	0.500	0.384	0.365	0.451	0.468	0.370	0.355

Notes: The table displays the estimates from baseline Difference-in-Differences specifications for the Romanian written exam and exam pass probability, for different subgroups of students, divided by: ability (columns 1-4), poverty status (columns 5-8), gender (columns 9-12). Standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Placebo test by ability, poor, and gender. Written Romanian vs. percentile rank oral Romanian score, standardized; 2010-2012 academic years

	High-stakes exam		**Low-stakes exam**	
	Written Romanian exam		**Oral Romanian exam**	
	(1)	(2)	(3)	(4)
PANEL A: Ability	**High ability**	**Low ability**	**High ability**	**Low ability**
Camera	-0.046	-0.228***	0.053	0.086
	(0.042)	(0.066)	(0.039)	(0.071)
Year12	-0.331***	-0.581***	-0.041	-0.150**
	(0.045)	(0.067)	(0.044)	(0.074)
Year11	-0.271***	-0.445***	0.017	0.008
	(0.032)	(0.047)	(0.030)	(0.050)
Poor	-0.222***	-0.120***	-0.141***	-0.168***
	(0.014)	(0.018)	(0.014)	(0.018)
Ability	0.787***	0.502***	0.389***	0.504***
	(0.017)	(0.011)	(0.021)	(0.019)
Theoretical	0.283***	0.304***	0.174***	0.170***
	(0.018)	(0.020)	(0.022)	(0.035)
Male	-0.283***	-0.211***	-0.142***	-0.167***
	(0.009)	(0.010)	(0.009)	(0.013)
County FE	Yes	Yes	Yes	Yes
Observations	202,032	198,056	202,032	198,056
R-squared	0.289	0.341	0.082	0.134
PANEL B: Poverty	**Poor**	**Non-poor**	**Poor**	**Non-poor**
Camera	-0.273***	-0.146***	-0.000	0.012
	(0.048)	(0.039)	(0.027)	(0.019)
Year 12	-0.466***	-0.402***	-0.055	-0.029
	(0.038)	(0.031)	(0.035)	(0.025)
Year 11	-0.378***	-0.328***	0.054**	0.050***
	(0.036)	(0.031)	(0.025)	(0.018)
Ability	0.509***	0.588***	0.490***	0.472***
	(0.008)	(0.009)	(0.009)	(0.017)
Theoretical	0.281***	0.315***	0.131***	0.174***
	(0.023)	(0.015)	(0.023)	(0.027)
Male	-0.291***	-0.232***	-0.229***	-0.131***
	(0.009)	(0.008)	(0.015)	(0.008)
County FE	Yes	Yes	Yes	Yes
Observations	89,375	310,713	89,375	310,713
R-squared	0.476	0.522	0.274	0.286

	High-stakes exam		Low-stakes exam	
	Written Romanian exam		Oral Romanian exam	
	(1)	(2)	(3)	(4)
PANEL C: Gender	**Male**	**Female**	**Male**	**Female**
Camera	-0.216***	-0.140***	0.014	0.002
	(0.048)	(0.036)	(0.026)	(0.016)
Year12	-0.471***	-0.370***	-0.036	-0.032
	(0.037)	(0.029)	(0.033)	(0.023)
Year11	-0.366***	-0.310***	0.062**	0.044***
	(0.037)	(0.030)	(0.025)	(0.015)
Poor	-0.184***	-0.152***	-0.201***	-0.126***
	(0.010)	(0.010)	(0.013)	(0.011)
Ability	0.523***	0.619***	0.476***	0.474***
	(0.008)	(0.010)	(0.015)	(0.014)
Theoretical	0.342***	0.287***	0.194***	0.138***
	(0.015)	(0.015)	(0.026)	(0.024)
County FE	Yes	Yes	Yes	Yes
Observations	182,943	217,145	182,943	217,145
R-squared	0.485	0.488	0.272	0.265

Notes: The table displays the estimates from the baseline Difference-in-Differences specifications for the Romanian written exam performance (columns 1-4) and the Romanian oral exam performance (columns 5-8), for different subgroups of students, divided by: ability (Panel A), poverty status (Panel B), and gender (Panel C). Both dependent variables are expressed in standardized percentile rank scores, using data from 2010-2012. Standard errors in parentheses are clustered at county level.*** p<0.01, ** p<0.05, * p<0.1

Table 6: Composition of cohorts admitted to an elite university 2009-2012

	Ability (Standardized)			Poor			Male
	All	Poor	Non-poor	All	Low ability	High ability	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: ALL ADMITTED STUDENTS							
Camera	0.057	0.059	0.058	-0.033***	-0.051**	-0.016	-0.025
	(0.049)	(0.198)	(0.045)	(0.012)	(0.024)	(0.020)	(0.022)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	15,395	813	14,582	15,395	7,762	7,633	15,395
R-squared	0.039	0.085	0.040	0.039	0.048	0.039	0.019
Panel B: TUITION-EXEMPT STUDENTS (the top students)							
Camera	0.130***	0.161	0.128**	-0.035**	-0.060**	-0.019	-0.040
	(0.045)	(0.154)	(0.048)	(0.013)	(0.024)	(0.022)	(0.029)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	9,777	546	9,231	9,777	3,973	5,804	9,777
R-squared	0.045	0.079	0.046	0.038	0.052	0.037	0.020
Panel C: TUITION-PAYING STUDENTS (good students)							
Camera	0.082	-0.015	0.086	-0.032	-0.038	-0.023	-0.042
	(0.086)	(0.470)	(0.072)	(0.027)	(0.036)	(0.038)	(0.040)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	5,618	267	5,351	5,618	3,789	1,829	5,618
R-squared	0.027	0.207	0.026	0.059	0.064	0.080	0.023

Notes: The table displays the baseline Difference-in-Differences specifications for the composition of admitted university students in terms of ability (columns 1-3), poverty status (columns 4-6) and gender (columns 7). The estimates for the changes in composition in terms of ability are further divided by poverty status (columns 2 and 3) the estimates for the changes in terms of poverty status are further divided by students ability (columns 5 and 6). All regressions include a dummy indicator for students who took the Baccalaureate before the year of university admission. Results are similar if we do not include this indicator. Standard errors clustered at county level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 7: Composition of students in the top 20% of the final Baccalaureate score distribution.

	Ability Standardized			**Poor**			**Male**
	All	Poor	Non-poor	All	Low ability	High ability	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Camera	0.047**	0.092***	0.038*	-0.023***	-0.020***	-0.024	-0.008
	(0.020)	(0.030)	(0.020)	(0.007)	(0.007)	(0.018)	(0.008)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	108,461	11,681	96,780	108,461	100,383	8,078	108,461
R-squared	0.030	0.041	0.028	0.046	0.043	0.074	0.003

Notes: The table displays the baseline Difference-in-Differences specifications for the composition of students in the top 20% of the Baccalaureate final score distribution, in terms of ability (columns 1-3), poverty status (columns 4-6) and gender (columns 7). The estimates for the changes in composition in terms of ability are further divided by poverty status (columns 2 and 3) the estimates for the changes in terms of poverty status are further divided by students'ability (columns 5 and 6). Results are similar if we do not include this indicator. Standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.10.

A Appendix – Robustness and further tests

Table A1 demonstrates (in columns 1-5 for the written Romanian test and columns 6-10 for the probability of passing the Baccalaureate) that our results in Table 2 are robust to different specifications. First, since we saw some evidence from the graphical analysis that there was a tendency for the written Romanian exam scores to converge between 2009 and 2010, we want to investigate if controlling for this pattern changes the conclusions. Columns (2) and (7) include the county-specific trends, which does not change the main results (shown in columns 1 and 6). This accounts for potential selection of counties due to pre-campaign performance or corruption trends (assuming these would be linear). Columns (3) and (8) add a placebo camera indicator (equal to 1 in 2010 for the counties that were first monitored in 2011 and in year 2011 for the counties that were first monitored in 2012, and 0 otherwise), which is not significant, while the magnitude of the main coefficients does not change, even though the camera indicator in column (3) is not statistically significant.

We also exclude observations in 2010 and hold as benchmark the year 2009. This is done to rule out concerns about the estimates of interest being driven by the contrast to the exceptional events in the 2010 "Xeroxed exam." The results shown in Table A2 confirm that this is not the case. Moreover, when restricting the sample to 2011 and 2012 (hence the variation in monitoring comes only from late implementers), we find that counties that implemented the camera later sustained a larger drop in scores than the early implementers.

Second, one might worry that our controls are not sufficient to adjust for compositional differences between counties that were early or late camera implementers. In columns (4) and (9) we replace the county indicators with school indicators and find that the estimates and standard errors are almost identical to the baseline ones. Lastly, using the location, family name, and father's initial, we detect a sample of about 90,000 sibling students. In this sample, the exogenous variation in scores stems from a monitored and an un-monitored sibling, after netting out everything common to the siblings (e.g., family investment in children's education).[45] The

[45]Based on intra-class correlations of 5th-8th grade performance, we keep the groups of two assumed siblings (for whom the intra-family correlation is 30%, a typical estimate from the literature on sibling correlations in educational achievement; see Bjrklund and Jntti, 2012). Thus, the most popular surnames (seemingly yielding larger groups of siblings) are automatically excluded, thereby increasing the likelihood that we indeed identify siblings. A critique to this approach is that the exclusion of most popular names could entail the systematic exclusion of low-income students. We therefore face a trade-off between precision of sibling pairing and the extent to which the sibling sample is representative. Yet, the analysis using the extended sample of siblings (allowing for up to four students per "family") yields very similar results. At worst we have a random sample of

estimates shown in columns (5) and (10) do not depart from the baseline results, supporting that the pre-2011 scores were artificially inflated and that the sharp drop in scores is the impact of the anti-corruption intervention.

We have also checked whether our results are affected by the fact that our main sample excludes exam re-takes (47,910 observations) and students for whom we do not have ability as proxied by the 5th-8th grade scores (201,000 observations). Including re-takes and repeating the analysis without controlling for ability yield similar results as our baseline analysis.

students, and the results should be similar to the baseline estimates if the anti-corruption campaign had an effect on exam outcomes.

Table A1: Sensitivity Analysis

	Written Romanian exam					Baccalaureate pass				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Camera	-0.237***	-0.321***	-0.205	-0.244***	-0.288**	-0.084***	-0.101***	-0.108***	-0.088***	-0.102**
	(0.084)	(0.083)	(0.142)	(0.084)	(0.137)	(0.025)	(0.023)	(0.039)	(0.024)	(0.042)
Placebo camera			0.024					-0.018		
			(0.063)					(0.025)		
Year12	-0.888***	-0.692***	-0.904***	-0.888***	-0.867***	-0.180***	-0.123***	-0.168***	-0.181***	-0.167***
	(0.057)	(0.074)	(0.083)	(0.057)	(0.100)	(0.018)	(0.016)	(0.023)	(0.018)	(0.036)
Year11	-0.707***	-0.593***	-0.720***	-0.707***	-0.661***	-0.158***	-0.127***	-0.148***	-0.161***	-0.144***
	(0.052)	(0.065)	(0.075)	(0.052)	(0.094)	(0.014)	(0.014)	(0.024)	(0.014)	(0.029)
Year09	-0.297***	-0.352***	-0.281***	-0.296***	-0.282***	0.090***	0.071***	0.078***	0.091***	0.084***
	(0.056)	(0.030)	(0.080)	(0.054)	(0.095)	(0.009)	(0.011)	(0.019)	(0.010)	(0.016)
Male	-0.483***	-0.483***	-0.483***	-0.480***	-0.508***	-0.039***	-0.039***	-0.039***	-0.036***	-0.042***
	(0.015)	(0.015)	(0.015)	(0.012)	(0.030)	(0.002)	(0.002)	(0.002)	(0.002)	(0.006)
Poor	-0.299***	-0.298***	-0.299***	-0.224***	-0.236***	-0.061***	-0.060***	-0.061***	-0.053***	-0.064***
	(0.018)	(0.018)	(0.018)	(0.011)	(0.042)	(0.004)	(0.005)	(0.004)	(0.003)	(0.012)
Ability	1.038***	1.039***	1.038***	0.854***	1.035***	0.184***	0.184***	0.184***	0.134***	0.182***
	(0.019)	(0.019)	(0.019)	(0.026)	(0.035)	(0.007)	(0.007)	(0.007)	(0.005)	(0.012)
Theoretical	0.602***	0.601***	0.602***	0.348***	0.602***	0.186***	0.185***	0.186***	0.116***	0.187***
	(0.030)	(0.029)	(0.030)	(0.029)	(0.048)	(0.010)	(0.010)	(0.010)	(0.010)	(0.016)
Rural	-0.163***	-0.161***	-0.163***			-0.037**	-0.036**	-0.037**		
	(0.048)	(0.049)	(0.048)			(0.015)	(0.015)	(0.015)		
County FE	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
County trends	No	Yes	No	No	No	No	Yes	No	No	No
School FE	No	No	No	Yes	No	No	No	No	Yes	No
Family FE	No	No	No	No	Yes	No	No	No	No	Yes
Observations	547,447	547,447	547,447	547,447	89,967	553,903	553,903	553,903	553,903	90,915
R-squared	0.497	0.510	0.497	0.543	0.819	0.376	0.388	0.377	0.439	0.759

Notes: The table displays the estimates from alternative specifications for the Romanian written exam and exam pass probability. Columns 1-3 and 6-8 include county fixed effects. Columns 2 and include county specific linear time trends. Columns 3 and 8 include a placebo camera indicator equal to 1 in 2010 for the counties that were first monitored in 2011 and in year 2011 for the counties that were first monitored in 2012, and 0 otherwise. Columns 4 and 9 include school fixed effects. Columns 5 and 10 include family fixed effects. Standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Sensitivity check: the main results without the 2010 academic year

	Written Romanian			Baccalaureate pass		
	(1)	(2)	(3)	(4)	(5)	(6)
Camera	-0.230**	-0.229**	-0.239**	-0.092***	-0.090***	-0.093***
	(0.098)	(0.095)	(0.095)	(0.026)	(0.025)	(0.025)
Year12	-0.457***	-0.607***	-0.605***	-0.241***	-0.265***	-0.266***
	(0.093)	(0.094)	(0.093)	(0.022)	(0.021)	(0.020)
Year11	-0.407***	-0.414***	-0.413***	-0.243***	-0.243***	-0.245***
	(0.080)	(0.084)	(0.084)	(0.019)	(0.019)	(0.019)
Male		-0.504***	-0.500***		-0.037***	-0.034***
		(0.018)	(0.015)		(0.002)	(0.002)
Poor		-0.326***	-0.240***		-0.068***	-0.056***
		(0.019)	(0.011)		(0.004)	(0.003)
Ability		1.099***	0.897***		0.186***	0.135***
		(0.021)	(0.029)		(0.006)	(0.005)
Theoretical		0.643***	0.369***		0.205***	0.130***
		(0.032)	(0.032)		(0.010)	(0.010)
Rural		-0.169***			-0.040**	
		(0.059)			(0.018)	
County FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	No	No	No
Observations	405,046	405,046	405,046	410,523	410,523	410,523
R-squared	0.044	0.491	0.542	0.114	0.397	0.456

Notes: The table displays estimates from the baseline Difference-in-Differences specifications, for the written Romanian exam scores and Baccalaureate pass probability, excluding the year 2010. Standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

B Appendix – How good is our poverty proxy?

In this digression we scrutinize the quality of our poverty proxy. Firstly, we need to clarify what part of the income distribution the MHS status represents. Using the Romanian Household Budget Survey we have identified these students in households situated in the 10%-40% quantiles. This means that our analysis does not capture students living in extreme poverty, nor Roma children of the age of these cohorts, since these are the most likely to be high school dropouts. This is bound to slightly reduce the external validity of our finding.

Secondly, we try to rule out the concern that the effects of the MHS program on the beneficiaries' performance might confound our interpretation of the interaction estimates. We extract some evidence from a special feature of the MHS program. The disbursement of MHS funds has been carried out every year since 2004. However, in the beginning of the program, the funds fell short of the demand. This meant that from a total of about 76,500 eligible students (income below 150 RON, equivalent to 35 EUR, per household member) in the academic year 2005-2006, 31,547 were omitted from the program.[46] Some of these students applied and received the MHS funds in subsequent years, but 19,743 students never benefited from the MHS. We therefore use a regression discontinuity design to estimate the treatment effect of receiving money on exam scores, for the marginal student just receiving money, relative to the marginal student who never received the money The cutoff for receiving the money was set within each county, but varied only marginally around 30 RON. However, this means that as long as we include county fixed effects in the regression, we are able to use a sharp RD design. Hence, we estimate the effect for a weighted average of marginal students just receiving money, where the weights are given by the number of students at each cutoff. In order to capture all targeted students' exam outcomes (i.e., students who were eligible and applied for MHS in 2005-2006), we make use of the 2006-2010 Baccalaureate sample. The drawback with this sample is that we do not have corresponding data about the 5th-8th grade score, nor other background variables, apart from high school track.

[46] In our sample, these students who were not allotted the MHS in 2005-2006 despite being eligible, report incomes between 30 and 150 RON per family member, and the mean income is 82.6 RON. In the subsequent years the funds allocated from the national budget for MHS were adjusted at the beginning of each year in response to the demand, leaving no more eligible requests unsatisfied. The schools where the applications were registered had to submit their lists of applicants to the Ministry, which disbursed the funds, and typically they ranked the students by income, drawing the line according to the funds available. However, because of rising demands, from 2009 to 2010 a new criterion was introduced demanding that the student must have a very good school attendance rate. A little over 100 students were denied the allowance because of low attendance in 2010-2011.

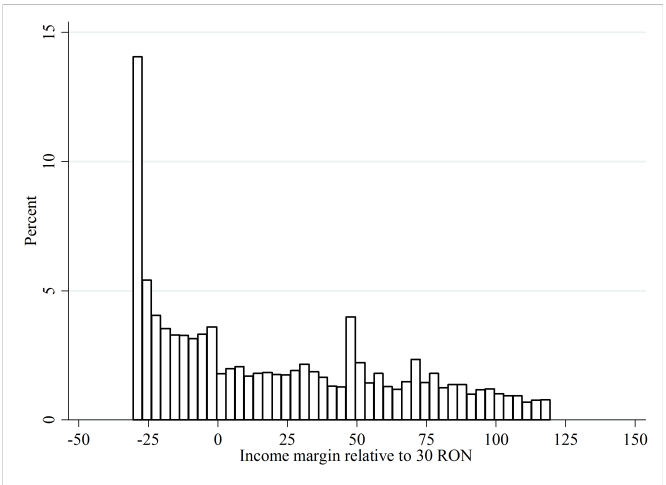
We estimate the following equation:

$$y_{ict} = \alpha + \beta_0 NMHS_{ict} + \beta_1 inc06_{ict} + \gamma' \cdot X_{ict} + \theta_c + \epsilon_{ict}, \tag{2}$$

where $NMHS_{ict}$ is an indicator equal to 1 if the student is a non-beneficiary, $inc06_{ict}$ is the family income in 2006, and X_{ict} is an indicator for theoretic track. The coefficient of interest, which yields the effect of the program, is β_0.

When we estimate this model, we get virtually no effects from the program once we control for income (Table B1). We interpret this as evidence that the MHS program did not affect the performance of the recipients relative to their comparable peers, and thus it can be used as a proxy for poverty status. The caveat is that some students may have underreported income, making some sorting around the cutoff a possibility (see Figure B1). The results hold also when we exclude those with close to or zero income, the easiest to misreport. Nonetheless, we interpret the RD estimate as suggestive rather than causal here.

Figure B1: Income margin density of the MHS applicants.



Notes: The figure displays the density bar chart of the MHS applicants income margin relative to the 30 RON cutoff in 2005-2006. The figure excludes applicants who reported 0 income.

Table B1: The MHS treatment effect. RD regressions

	Written Romanian		Baccalaureate pass	
	(1)	(2)	(3)	(4)
NMHS	0.146***	-0.020	0.021***	-0.002
	(0.023)	(0.042)	(0.005)	(0.008)
Income 2006		0.210***		0.029***
		(0.044)		(0.010)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Track control	Yes	Yes	Yes	Yes
Observations	64,506	64,506	64,511	64,511
R-squared	0.159	0.160	0.180	0.180

Notes: The table displays estimates from a sharp Regression Discontinuity in exam scores around the cutoff of income below which students are treated with the "Money for Highschool" financial support. NMHS is an indicator equal to 1 if the student did not receive the financial support. Standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.1

C Appendix – Supplementary Figures and Tables

Figure C1: Romanian written exam scores density 2010 vs. 2012

Figure C1.a. By ability

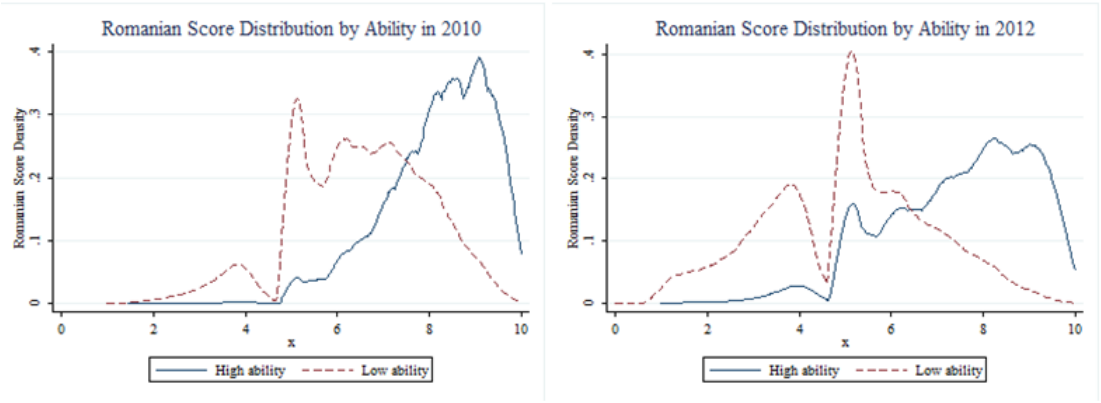


Figure C1.b By gender

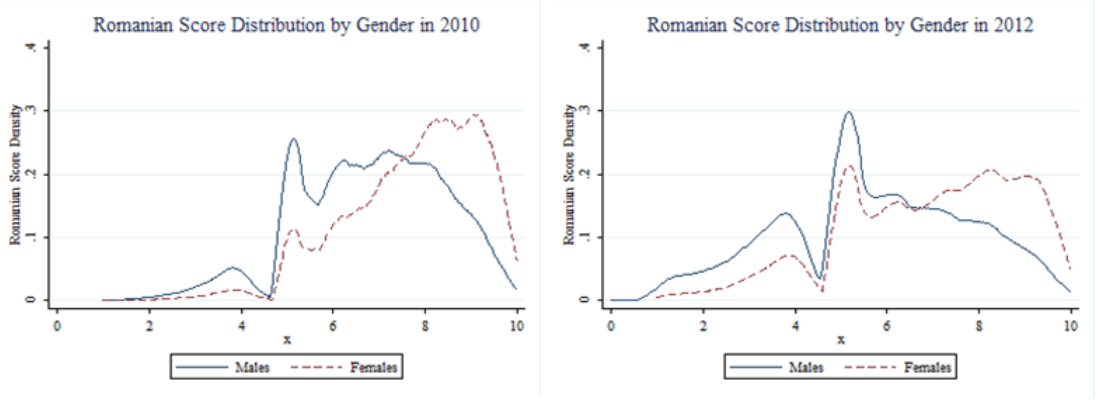
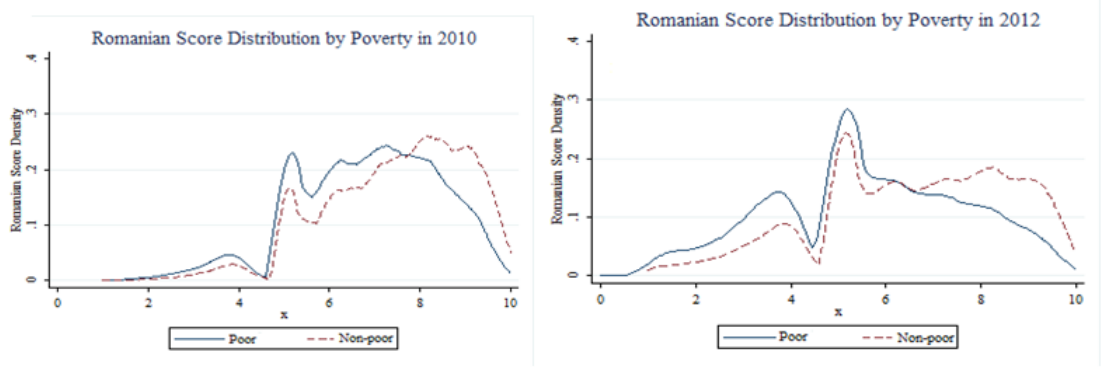


Figure C1.c By poverty status



Notes: The figure displays written Romanian exam score distributions in 2010 (left) and 2012 (right), for different subgroups of students, divided by: ability, gender and poverty status.

Figure C2: Changes at the 2010 exam. All test score distributions in 2009 and 2010

Notes: The figures display the score distributions for each written test in 2009 (blue) and 2010 (red dashed): 1) the written Romanian exam (top-left); 2) the track-specific exam (top-right); 3) the first elective exam (bottom-left); and 4) the second elective exam (bottom-right). Note that the second elective was removed in 2010, and before that, around 75% of the students chose physical education as their second elective test.

Table C1: Self-selection into camera treatment

	Early installation	Late installation	Difference	County clustered SE p-value
Pass	0.791	0.828	-0.037	0.220
Romanian exam score	7.175	7.239	-0.064	0.593
Ability	8.632	8.618	0.014	0.742
Male	0.459	0.462	-0.003	0.618
Poor	0.184	0.263	-0.079	0.033**
Theoretical	0.504	0.483	0.021	0.387
Rural	0.038	0.060	-0.022	0.247
Log county population	13.343	12.957	0.386	0.040**
Trust in justice	1.866	2.032	-0.166	0.103
Corruption BOP	0.558	0.379	0.179	0.331
Unemployment April	8.02	9.019	-0.999	0.348
County share Romanians	0.852	0.802	0.050	0.367
N	191970	97986		

Notes: The figure displays individual and county summary statistics for the joint years 2009-2010, separately by early and late camera installation. The trust in justice variable is an average county score calculated by us using the answers to the question "Can justice courts be trusted?," from the Romanian Barometer of Public Opinion 2007, Soros Foundation. The variable Corruption BOP is a proxy developed by our calculations using the same Public Opinion Barometer. We use the question: "Is there anyone (i.e., informal network) that could "help" you solve (i.e., informally): issues in court/trials, medical problems, city hall, police, or issues related to the local authorities?" P-values are based on standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Share of students eliminated from the exam due to in-class cheating

	Share of Eliminated Students			
	(1)	(2)	(3)	(4)
Camera	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Year12	-0.002**	-0.002**	-0.002**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Year11	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Year09	0.000	0.000*	0.000*	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Male		0.000***	0.000***	0.000***
		(0.000)	(0.000)	(0.000)
Poor		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
Poor x Camera			0.001	0.000
			(0.001)	(0.001)
Poor09			0.000	0.000
			(0.000)	(0.000)
Poor11			-0.000	0.000
			(0.000)	(0.000)
Poor12			-0.000	-0.000
			(0.001)	(0.001)
Ability		-0.001***	-0.001***	-0.000***
		(0.000)	(0.000)	(0.000)
Theoretical		-0.000	-0.000	0.000
		(0.000)	(0.000)	(0.000)
Rural		0.001	0.001	
		(0.001)	(0.001)	
County FE	yes	yes	yes	no
School FE	no	no	no	Yes
Observations	553,903	553,903	553,903	553,903
R-squared	0.002	0.003	0.003	0.020

Notes: The table displays estimates from the baseline Difference-in-Differences specifications, for the probability to be eliminated from the exam due to cheating. In addition to the standard specifications in columns 1 and 2, columns 3 and 4 display the estimated parameters of all treatment interactions with poverty status. Standard errors clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

Chapter IV

The Benefits of Local Party Alignment in National Elections

Oana Borcan*
University of Gothenburg

Abstract

This paper provides robust evidence that local officials deliver votes for their parties in national elections. I use a sharp regression discontinuity design with closely-contested Romanian local elections in June 2012. I find up to 5.4 percentage points increased turnout in government-aligned localities at the July 2012 referendum launched by the governing coalition to dismiss the president. Turnout was crucial in the referendum, as a minimum participation of 50% of all voters was required. Instead, I find no direct electoral alignment advantage in terms of turnout or vote shares in subsequent parliamentary elections. The referendum alignment effect is driven by rural areas, with less educated and more manipulable voters. This along with the contrasting results at legislative elections, and extra heterogeneity tests suggest that local politicians mobilize voters successfully when: i) the voter commitment problem is overcome (unlike the vote, turnout is observable); ii) vote buying is common; iii) there is weak local competition and monitoring of incumbents. I also show suggestive evidence that after the referendum, government transfers increase in aligned localities and higher referendum turnout also drives higher legislative elections turnout and vote shares for the government coalition.

JEL Codes: D72, H77, P26

Keywords: party alignment, elections, vote-buying, intergovernmental transfers

*Email: oana.borcan@economics.gu.se. I thank Andreea Mitrut, Ola Olsson, Amrish Patel, Måns Söderbom, Randi Hjalmarsson, Barbara Petrongolo, Eric Battistin, Marco Manacorda, Andrea Tesei, Albert Solé-Ollé, Pilar Sorribas-Navarro, James Snyder, Francesco Fasani, Hannes Mueller, Gabriel Kreindler and Ragnar Torvik for valuable discussions. Seminar and conference participants at the University of Gothenburg, Queen Mary University, IEB Workshop on Politican Economy and ERMAS 2015 provided useful comments. I am grateful to the Romanian Central Electoral Bureau for providing data. All errors are my own.

1 Introduction

A growing number of studies have brought under empirical scrutiny the issue of political alignment at different tiers of government and associated electoral gains. These studies have all produced the same intriguing find: locally aligned municipalities may exhibit incumbency effects in municipal elections, but they don't award the parties in government with more votes in national elections. This line of inquiry emerged from the vast literature on intergovernmental grants, which are clearly allocated preferentially to municipalities controlled by the party in central government (e.g. Bracco et al., 2015; Brollo and Nannicini, 2012; Solé-Ollé and Sorribas-Navarro, 2008).

The ensuing question is what are the gains for central governments. The working assumption in these studies is that local politicians represent important "political capital" for their parties, manifested in vote delivery in national polls, alongside policy congruence and campaign efforts (e.g. Brollo and Nannicini, 2012; Grossman, 1994). However, although it was "widely accepted that federal politicians allocate own-purpose expenditures for the purpose of enhancing their reelection chances" (Grossman, 1994, p. 296) the anticipated national electoral bonuses from local alignment did not surface so far.

This paper is the first to provide robust evidence of an alignment effect in national polls. To get around the endogeneity of mayor alignment, I use a regression discontinuity design, and compare national polls outcomes in localities where an aligned candidate narrowly won vs. narrowly lost the mayoral race. I use data from the Romanian June 2012 local elections, and for the main outcomes I use data from a 2012 nationwide referendum and immediate parliamentary elections. Using these two elections, along with a battery of heterogeneity tests, I bring evidence that alignment-driven electoral bonuses occur when: 1) the voter's commitment problem can be overcome (see Robinson and Verdier, 2013); 2) vote-buying practices are common; 3) local politicians are left unchecked.

The first of these is evidenced by the referendum results. The July 2012 referendum was launched by the coalition in government to decide the impeachment of the president, who was affiliated with the opposition party.[1] Importantly, a quorum requirement of 50% voter presence made the referendum turnout the crucial outcome ahead of the vote shares. Hence, any campaign promises or vote-buying attempts were more likely to yield results since turnout is easily observed in contrast to the actual vote. I document a bias of up to 5.4 percentage points higher

[1] For details surrounding the referendum see http://www.theguardian.com/world/2012/jul/29/romanians-unlikely-impeach-president-traian-basescu (in English)

turnout in localities with mayors from the governing coalition than unaligned localities. By contrast, I find small and insignificant alignment effects on turnout at the parliamentary elections, which is expected since turnout is irrelevant in that context. Moreover, there are no significant alignment gains to vote shares neither in the referendum nor in legislative elections. This is in line with the logic of the voter commitment problem.

The second favourable circumstance is under widespread vote-buying practices. Note that the very short time between local elections and the referendum rules out pork-barrel spending as a way to attract votes (except perhaps promises, but the parliamentary results seem to dismiss this channel too). Vote-buying is therefore a plausible competing explanation. The results are suggestive of this channel, as the alignment effect on turnout is fully driven by rural areas, where the RD estimate is between 3.5 and 6.2 percentage points. This is in line with recent studies and the abundant Romanian contextual evidence that identify higher social pressures, more tightly-knit clientelistic networks, and a higher prevalence of vote buying in rural areas (Funk, 2010; Vicente, 2014; Volintiru, 2012). To further substantiate this channel, I also present heterogeneity tests which reveal larger alignment effects in counties with higher perceived incidence of vote buying and in counties with higher shares of out-migrants (the latter test is inspired by survey and trial evidence that fraudulent ballot stuffing in migrants'names also occurs). That mayors respond directly for these actions is evident from their role as leaders of the local administrations and, in particular at the referendum, from undercover journalist investigation reports (Biro, 2012). Another argument against pork-barrel spending is that localities aligned even before 2012 have, if anything, lower turnout bonuses than localities where the governing coalition is new in the local office.

The third condition identified above, namely lax checks and balances on mayors, is evidenced by a heterogeneity analysis on races where the challenger has different incentives to monitor the winner. In this case, the president's party had high stakes in the referendum, whereas other parties had low stakes. Consequently, the estimates show zero alignment effects in narrow races between the governing coalition and the president's platform, and very large (up to 14 percentage points) effects in races between the governing coalition and its allies. This points to the role of local competition in keeping electoral misconduct in check. Since most national elections have relatively more similar incentives for all parties, and since narrow races by default reflect a balance of political influence at local level, this may account for why the alignment effect in higher tier elections has not been found before.

It should be noted that the main alignment estimate on turnout is robust to various specifications, from OLS in small regions around the victory margin threshold, to various polynomial approximations using the full sample. The tests for an extensive set of covariates and vote margin density continuity at the threshold rule out the sorting bias issues raised in Caughey and Sekhon (2011) as a threat to the validity of RDD in electoral contexts.

Finally, I also inquire into post-referendum benefits from alignment. I show that aligned locality revenues are larger than those of unaligned localities in 2013, which reverses the 2011 pattern for these localities. The evidence suggests that government transfers are preferentially directed towards localities of the governing parties, although this cannot be causally linked to the local officials' efforts in the referendum. Going further, under fairly strong assumptions, I also use alignment as an instrument for referendum turnout, and I show that localities with higher turnout at the referendum also have significantly higher turnout and incumbent vote shares at parliamentary turnout. This effect is clearly not large enough to show in the reduced form, but it suggests that the referendum served as a mobilization exercise.

The paper is structured as follows: section 2 presents an overview of the related literatures and my contributions; section 3 describes the institutional setting; section 4 discusses the data; section 5 outlines the identification strategy and provides validity tests; section 6 displays the main results, the heterogeneity analysis and some robustness checks; section 7 presents a brief heterogeneity analysis and a discussion of mechanisms; section 7 presents the results from local revenues and parliamentary elections; section 8 concludes.

2 Literature Review

This paper contributes to at least three strands of literature.

Firstly, it fills a gap in the literature surrounding partisan alignment, and the political and economic advantages it may entail. A part of this literature investigates the role of local officials' partisan alignment in discriminate intergovernmental transfers. A few earlier correlation studies (Grossman, 1994 and Snyder and Levitt, 1995 for the U.S. and Worthington and Dollery, 1998, for Australia) and more recent causal analyses (Solé-Ollé and Sorribas-Navarro, 2008; Brollo and Nannicini, 2012; Migueis, 2013; Bracco et al., 2015) found that localities aligned with the party in power receive more transfers than those unaligned. A common assumption in these studies is that local officials provide an important source of political capital for their parties, partly because they can mobilize voters and provide electoral advantages

also for higher-tier politicians. This assumption is also present in theoretical papers that explain partisan transfers (e.g. Zudenkova, 2011 and Persico et al., 2011). Yet to date there is no hard evidence in support of this assumption.

Earlier studies on reverse coattail effects (i.e. electoral success of lower-tier politicians attracts the party's success in higher-tier polls)[2] showed mixed and no more than correlational evidence (Ames, 1994, Samuels, 2000a, Samuels, 2000b; Broockman, 2009).

Only a few studies have tested this assumption in a causal sense, allowing the possibility that local politicians may attract votes also through electoral corruption. Brollo and Nannicini (2012) use data from Brazil in a RDD to show that the president's party penalizes unaligned mayors through smaller federal funds before local elections. However, they find a minor, but not robust, electoral gain from alignment in presidential elections.

Ade and Freier (2013) also use a close elections design to study the dependency between mayoral and town council elections in Germany. They find that mayoral winners attract a vote bonus for their co-partisans in council when the two elections are simultaneous. This advantage is lost if elections are sequential.[3] However, they do not find any local alignment effects in European and German parliament elections.

The same pattern of preferential transfers is documented also in Portugal by Migueis (2013), but again, when looking at the electoral outcomes in national elections using a similar RD strategy, no alignment effect was found.

Lastly, Martinez-Bravo (2014) presents a theoretical model and data from the democratization years in Indonesia, showing that appointed local officials may have gone to great lengths (using heavy campaigning and vote-buying) to attract votes for the district mayoral elections. While controlling thoroughly for observables, her identification leaves room for potential selection.

Thus, I contribute to this literature with robust evidence that directly elected lower-tier politicians provide electoral advantages at higher tiers. Moreover, the setting I exploit allows for a more nuanced discussion than was possible in previous studies, about the mechanisms and conditions under which this alignment effect emerges. Contrasting a turnout quorum referendum with typical parliamentary elections I show that an effect is picked up only when voters' choices are observable, which helps to overcome the voter commitment problem (Robinson and Verdier,

[2]For documented presidential coattail effects in the U.S., see Cohen et al. (2000), Mattei and Glasgow (2005), Gélineau and Remmer (2006) and Golder (2006).

[3]One explanation is that when elections are on the same day, voters incur smaller cognitive costs if they choose the same party, while if elections are sequential, voters update and adjust their choice to ensure a power balance through divided government

2013).[4] The small lag between the local and the first nationwide elections rules out pork-barrel spending and mitigates the issue of voters' time-inconsistent preferences. I show also that vote-buying norms can be conducive to the alignment effects found. Moreover, I put forth the role of local competition in preventing incumbents from monopolizing the vote-rigging machine. Overall, the evidence suggests local officials provide the often invoked political capital and provides some explanations for why this type of alignment effect has not been found before.

A second related literature investigates the voter participation decision. Why and how people vote are central questions in political economics. The voter paradox (we observe positive turnouts despite the near-zero benefits of the individual vote), has been explained through concepts like the utility from voting (e.g. fulfilling a civil or moral obligation, see Riker and Ordeshook, 1968) and social norms (e.g. social pressure, see Funk, 2010). As for externally enforced voting decisions, votes can be gained through targeted spending (e.g. Manacorda et al., 2011; Pop-Eleches and Pop-Eleches, 2012), exploiting media biases (e.g. Durante and Knight, 2012; Durante et al., 2014) and, importantly, through active mobilization and vote buying (e.g. Finan and Schechter, 2012). I show that local officials exert a direct influence on turnout in favor of their parties, with potential ripple effects in subsequent elections.

The paper also contributes to the literature on political processes in young democracies by showing that even in countries well into transition, the misconduct of directly elected local politicians can impinge on democratic freedoms, by means of aggressive voter mobilization, vote buying and electoral fraud (Keefer and Vlaicu, 2008, Finan and Schechter, 2012, Martinez-Bravo, 2014).

3 Institutional Setting

3.1 The Romanian Electoral System and Politics in 2012

Romania is a young semi-presidential democracy, ruled by a government accountable to the Parliament. As head of state, the president is an active player in internal politics: he oversees the balance of powers, has the right to appoint the prime minister and to veto laws. The president is directly elected every five years in a runoff majority vote. The parliamentary elections run on a list system of independent candidatures, based on a closed list proportional representation system, with a min-

[4]The voter commitment problem refers to the possibility that voters may receive favours or money from politicians in exchange for their promise to vote in a certain way, but may default on that promise when the vote is secret.

imum vote share of 5% required for any party to get seats.[5] A multiparty system has been in place since 1990, with as many as 39 parties and alliances registered for the 2012 parliamentary elections. However, the most prominent parties in 2012 were the former communist Social Democratic Party (PSD), *the liberals* - National Liberal Party (PNL) and the *centrist* Liberal Democratic Party (PDL).

The incumbent president in 2012, Traian Basescu, although *de jure* politically unaffiliated, enjoyed strong support from the centrist PDL and was generally regarded to be its unofficial leader.[6] From December 2008 until April 2012, all the cabinets represented coalitions with the *centrists*. For most of this time, the *social-democrats* and *the liberals* were in the opposition. In February 2011, the latter forged a coalition - the *Social-Liberal Union* (USL), also joined by the Conservative Party - PC. Presided by the PSD and PNL leaders Victor Ponta and Crin Antonescu, USL made it a first priority to remove the *centrist* government and the president Basescu from office.[7]

The year 2012 was an electoral year: local elections were held on June 10, and parliamentary elections on December 9. In light of these electoral events, the two rivals, *the centrist party* and *the social-liberal coalition*, had large stakes from staying in power and taking power, respectively. The coalition gained momentum in April 2012, when a newly formed *centrist* cabinet fell short of majority support in the Parliament and was dismissed through a motion of no confidence.[8] Shortly after, Victor Ponta took office and formed the *social-liberal coalition government*, which immediately initiated the legal procedures to suspend president Basescu from office. These entailed a quick succession of controversial institutional reforms, leading to a full-blown political crisis.[9] The impeachment of the president was to be decided at the national referendum in July 2012.

[5]In the proportional representation system, legislative seats get allocated to each party in proportion to the number of votes the party receives. The closed-list system means that each party has an internal method for deciding the candidates put forth for elections, and each voter effectively casts only one vote for the party in each chamber, for the candidate decided *a priori* by the party.

[6]Article 84 in the Romanian Constitution stipulates that the president cannot be a member of any political party while in office. However, before taking office in 2004, Basescu had been been the president of one of the parties that merged in 2007 to found PDL.

[7]USL Founding Document issued on 5 February 2011 outlines the alliance's political agenda. The first objective reads: 'To remove from power, in a democratic manner, the current clientelistic, corrupt and inefficient regime'

[8]In February 2012 the prime minister who, at that time, was a favourite of the president, was forced by popular protests to step down from office. His centrist government had lost popular support following the austerity measures in 2010 (e.g. 25% public sector wage cuts).

[9]Constitutional court processes were changed and the national ombudsman was replaced with another, who had ties to the *social-liberals*. This made it easier to pass government ordinances facilitating the president's impeachment. Ponta's initiatives and the June 2012 political crisis are described in Politeanu (2012) and some accounts of it are also available at:
http://www.economist.com/blogs/easternapproaches/2012/06/romanian-politics-0 (in English)

3.2 The 2012 Impeachment Referendum

On 4th July, the *social-liberal union* submitted to the Parliament an official request to impeach the president on grounds of unconstitutional conduct in office.[10] On 6th July, a Parliament majority voted in favor of impeachment, with his recall from office to be decided in a national referendum.[11] The referendum was held on the 29th July, when Basescu's popularity was at an all-time low, owing to the draconian austerity cuts by at least 25% in public sector wages in 2010. Opinion polls just ahead of the referendum anticipated his removal from office.[12]

Importantly, the referendum law in Romania stipulated a *quorum* rule: a minimum turnout of 50% was necessary to validate the referendum.[13] On the day of the referendum, a staggering 87.52% voted "YES" for having Basescu removed.[14] However, only 46.24% of the 18 million registered voters cast their ballot, and therefore the referendum was ruled invalid by the constitutional court. The president resumed his duties shortly.

The president had withstood an impeachment referendum before, in 2007, when 74.48% of voters agreed to keep him in office. The turnout in the first referendum was 44.45%, but no quorum rule was in place at the time. Figure 1 shows a brief chronology of political and electoral events leading up to the referendum.

3.3 The Role of Mayors in National Polls

Romania's local administration is organized into 41 counties (and the capital, further divided into urban localities (cities and towns) and rural localities (communes and villages). The local administration falls in the remit of mayors and local councils. As head of the local public administration, the mayor enjoys the highest status in the civil servants hierarchy and in the community (particularly in rural areas).[15]

http://www.theguardian.com/world/2012/jul/29/romanians-unlikely-impeach-president-traian-basescu (in English)

[10]Basescu was accused of having taken over too many of the government's attributions and to have attempted influencing the justice courts. The official document's title is 'Solicitare privind suspendarea din functie a presedintelui Romaniei, Traian Basescu'

[11]Basescu was immediately suspended from his attributions and Crin Antonescu, leader of the *liberals*, became interim president.

[12]Around 67-70% of respondents would vote him out, and just over 50% would cast their vote (53% in rural and 52% in urban areas) according to a survey conducted by the Group for Social-Behavioral Studies "Avangarde": "Operations Research" Survey 23-25th July; See http://bit.ly/1O8r35T (In Romanian)

[13]Law No.3/2000, article 5(2).

[14]The exact question on the ballots was "Do you agree with the dismissal of the president Traian Basescu?"

[15]Article 66 of Law No. 215/2001 Art. stipulates that "the mayor has a position of public authority. He/She is the head of public local administration and of the locality-specific public administration apparatus, which he/she manages and controls"

Mayors are directly elected every four years through a first-past-the-post system.[16] Once elected, mayors assume office almost immediately.[17]

Importantly, mayoral candidates can run for a separate party, a coalition, or as independent candidates. For instance, at the 2012 local elections, 42.39% of the seats were won by the *government coalition*, but an additional 11.9% of the seats were taken by candidates representing *the social-democrats* separately, and 8.3% by separate *liberal* candidates. The local competition between parties *within the governing coalition* was allegedly forced by quarrels over local administration seats. This has implications for the treatment definition (see section 4.1).

The mayor's political alignment flags the locality's political leaning, and is thought to predict the parties' local performance in national elections (Buti, 2012; IRES, 2012).[18]

Mayors are believed to have an active role in higher-tier elections, where they mobilize voters through: 1) campaigning and "get-out-the-vote" strategies (Seceleanu, 2009); 2) pork-barrel spending (e.g. Pop-Eleches and Pop-Eleches, 2012 investigate the EURO 200 program for supplying the needy with computers, Pop-Eleches and Pop-Eleches, 2012; EFOR, 2013 document the clientelist allocation of infrastructure funds). 3) facilitating vote-buying and electoral fraud, since as heads of the local administration they have direct access to the vote-rigging apparatus.

Vote buying is a common practice in Romania, particularly in rural, more tightly-knit and less educated communities. Votes are bought for eggs and buckets all the way to mobile phones, livestock or cash (Volintiru, 2012).[19] The 2012 elections, particularly the impeachment referendum, were fraught with allegations of electoral fraud (Freedom House, 2013). An undercover journalist published his phone conversations with small town mayors, passing for a government coalition representative just ahead of the polls (Biro, 2012). The mayors hinted to unorthodox practices: "You give them [the voters] a snaps, a sandwich, a pie. We'd rather you [the party] sent us money.", or "Evening after evening we went with the people in taverns. We'll have 75% [turnout]". Conversely, the *president's party's* mayors mentioned organized distractions to keep voters away from the polling stations.

[16]This system was first applied in the local elections on 10th June 2012. Until 2008, mayors were selected in a majority two-round election. The Law No. 129/2011 changed this to a first-past-the-post-system. Local councillors are also directly elected on the same day as the mayor, up to a number of seats determined by the population size.

[17]Mayors' mandate is validated in court in maximum 20 days after the local elections, as stipulated in article 63, Law No. 215/2001. Mayors can assume office right after taking the oath of duty, as soon as their mandate is deemed valid by the court.

[18]In 2000 legislative and local elections, the vote share correlation was 40% (Klasnja, 2014).

[19](Volintiru, 2012) shows details on these and other practices documented from face-to-face interviews with local officials and party members.

Additionally, several mayors were indicted in a most controversial referendum corruption trial against a minister, member of *the social-democratic party*, sentenced for electoral corruption to reach a 60% turnout target (National Anticorruption Directorate Press Release 2013).[20] Court files revealed a variety of fraudulent practices: ballot-stuffing (including votes attributed to out-migrants and deceased), violent threats to get or to impede votes, double-counting and the fraudulent use of the mobile ballot box which circumvented monitoring devices.[21] Several mayors from the *president's party* were also indicted for electoral misconduct.[22]

Overall, the rich anecdotal evidence reveals the (likely illicit) influence of local officials in national polls. The formal analysis below sets out to establish whether and how this influence was turned into a significant alignment bonus in the referendum and parliamentary elections.

4 Data

In order to estimate the impact of partisan alignment of mayors on the referendum outcomes and parliament election outcomes, I combine several sources of data aggregated at locality level:

i) Electoral data from 2012 local elections, the national impeachment referendum and legislative elections. This data is publicly available from the Romanian Electoral Authority (AEP). From this data I extract: the vote shares for mayoral candidates, underlying my running variable in the RDD; the referendum outcome variables: the turnouts by locality (defined as the share of ballots cast from the total number of registered voters) and the 'YES' vote shares (the percentage of votes in favour of dismissing the incumbent president of the total number of votes cast); parliamentary election outcomes: turnouts by locality and party vote shares by locality. I also use the turnout in the 2007 impeachment referendum and the number of mayoral candidates at the 2012 local elections as electoral controls in the regressions.

ii) Census data from 2011 and 2002 Romanian Population Censuses. This data is publicly available online from Statistics Romania. Using this source I construct the

[20] http://www.bbc.com/news/world-europe-24437209 (In English)
See also "Top Romanian minister Liviu Dragnea receives suspended prison sentence of 1 year for fraud in 2012 referendum to dismiss ex-president Basescu" http://english.hotnews.ro/, accessed 15 May 2015.

[21] Princeton political scientist Grigore Pop-Eleches describes the context of the referendum and the fraud means on U.S. political science blog http://themonkeycage.org/ in the article "Post-Election Report: Romanian's Presidential Impeachment Referendum, and a Request for Help in Identifying Potential Fraud" posted on 9th August 2012.

[22] See e.g. http://bit.ly/1Or9UCg, Romania Libera, published 30 July 2012 (in Romanian).
(In Romanian)

first set of pre-treatment locality-level covariates: the locality log population size, the share of voting-age population, the share aged over 65, the share of Romanians, the gender shares, the share of high school and university educated, the unemployment rate. For a slightly smaller sample, I also have a number of additional covariates: the shares of illiterate, the share working in agriculture and in public administration, the proportion of migrants working abroad.

iii) Fiscal data from 2011. This data from the Ministry of Regional Development is available at locality level, including: total income (own taxes, intergovernmental transfers, subsidies), total and split public expenditure (on education, health, and public services). These form the second set of covariates.

4.1 Defining treatment

A locality is treated if the newly elected mayor is aligned with the *governing coalition*. Alignment here assumes the support for the *social-liberal union*'s manifesto, i.e. interest in ousting the president and the representation of the coalition in all elections.

To reiterate, 1979 top-two mayoral candidates represented the *governing coalition*, while some ran for a separate party within it: 706 for *the social-democrats*, 554 for the *liberals*, and 47 for the *conservatives*. It is known that these these separate candidates were in disagreement with the coalition over local administration offices, but it is unclear to what extent they still had their interests aligned with *the governing coalition*. While the *social-democrats* were the main drivers of the referendum, the position of the independent *liberals* and *conservatives* was relatively ambiguous.[23] Therefore, to compare the level of engagement in the referendum, in Table 1 I present the turnout in narrow (comparable) races between the different parties, including independent *liberals* and *conservatives*. Not only is the only significant difference in turnout in races between *social-democrats* and *liberals/conservatives*, but the average turnout in localities where a *social-democrat* took office is 64.6% compared to 56.2% in *liberal/conservative* localities. Given the much lower turnouts in *liberal/conservative* narrowly won localities, I place these candidates in the control group in the main specifications.

Therefore a locality is aligned with the *governing coalition* (henceforth G-aligned) if the mayor is either from the governing coalition or from the *social-democrats* and G-unaligned otherwise.

[23]Also, they were the strongest party in politics. The *social-democrats* alone won 11.9% seats in the local elections, while the other two combined secured 9.65% of seats. Moreover, PSD leader Victor Ponta was head of the cabinet and also the main proponent of the reforms facilitating the impeachment.

In the robustness section 6.3 I also present the results with *liberals/conservatives* in the treatment group (where, in light of the contrasting electoral behaviour of *liberal/conservative* led localities, I expect the alignment effect to be underestimated). In alternative specifications, I also control for races against *liberals* and *conservatives*, and I also exclude these races from the sample.

4.2 Sample selection

The treatment definition above has some implications for the sample selection. The complete dataset contains 3181 localities with information about the 2012 elections. Firstly, I exclude localities with only one local candidate and those where the winning candidate obtained a vote share above 80%, which is unproblematic since assessing the effect of mayors' partisan alignment relies on identification from close electoral races. These account for 11.4% of all races.[24] This is also useful because estimates using higher order polynomial control functions used later on in the RD are sensitive to extreme values of the assignment variable, and may therefore be biased (Gelman and Imbens, 2014).

Secondly, I restrict the sample to those localities where a candidate aligned with the *governing coalition* was either winner or runner-up, irrespective of the number of candidates running in that locality. This leads to the exclusion of an additional 421 races, or 13.2% of all localities. I also drop 9 observations due to missing data for the covariates.

This sample definition comes with a caveat: localities will select into the sample based on the degree of popular support for the aligned candidate, i.e. depending on the running variable. The internal validity problem is that, while still using narrow races, I might compare a G-winner with, for instance, 47% vote share with a G-runner up of 37% share (this will depend on how many candidates split the votes in local elections), which is not a narrow enough comparison.

Brollo and Nannicini (2012) choose a cleaner identification from two- or three-candidate races over sample size. However, the pool of two-candidate races is also a potentially selected sample, and a much smaller one. In this case this strategy is less feasible, because in Romania elections are traditionally disputed amongst many parties. In the 2012 mayoral elections, some races had up to fifteen competitors and there were merely 154 (6.4%) two-candidate races and 479 (20%) three-candidate races against G, which means estimations based on these samples might have low

[24]However, their inclusion does not change the results qualitatively, since the identification of the alignment effect comes from localities close to the cutoff, as explained in detail in the next section. Ade and Freier (2013), for instance, exclude races where the victory margin is larger than 60%.

statistical power. The question is essentially how to resolve this trade-off between internal validity and precision.

Appendix Table A1, showing comparative statistics between the races with a victory margin within 5 percentage points and the two-candidates races, reveals that two-candidate localities have slightly lower population, higher per capita revenue, and, importantly, lower education levels and expenditures (around 1 percentage point significant difference in share with high education). Thus, apart from the higher per capita revenue, the other observables suggest two-candidate localities could be more susceptible to manipulation. Consequently, the results from these samples, which I present in section 5.2, could display an upward selection bias. Thus, my main strategy is to use the entire sample, where, in some specifications, I also control for the number of candidates and for the joint vote share of the first two ranked candidates (as this should help to pick up the alignment estimate only from the closest races). The advantage with this approach is enhanced statistical power and external validity against the restricted alternatives.

4.3 Descriptive statistics

The sample thus defined contains 2386 localities with mayoral races against a G-aligned candidate, of which 2116 are small rural localities (with 3263 inhabitants on average), and 270 are large urban localities.

Table 2 displays some comparative statistics in national elections outcomes and locality characteristics in the sample of races against G-aligned candidates, for all localities and separately for rural and for urban ones. Referendum turnout is markedly larger in localities where the G-aligned candidates won (58.3%) than where they lost (47.2%). This difference is larger in rural localities (12 percentage points) and smaller in urban localities, which also have generally lower turnout - below the quorum rule. Interestingly, this contrasts with the voters' electoral behaviour at the impeachment referendum in 2007, where turnout was significantly lower in G-aligned localities. However, the share of votes in favour of impeaching the president is very similar in all groups, around 86%.

In terms of parliamentary outcomes, it is evident that localities where G won the mayoral race have larger turnout and vote shares for G. The turnout difference is close to 3 percentage points.

Table 2 also shows that G-aligned and unaligned localities are largely similar in pre-treatment characteristics. While most p-values associated with the differences in means are large, non-aligned localities are slightly more ethnically diverse and

have larger per capita fiscal revenues and expenditures owing to somewhat larger subsidies (by a margin of 40 RON, approximately 9 EUR per capita).

These fiscal revenue differences could influence voters' political attitudes, and thus drive both the probability of electing a certain candidate as well as the electoral outcomes. However, they are less problematic in the RDD if the fiscal variables don't change discontinuously at the victory threshold.

Therefore, in Table 3, I zoom in on those races within 5 percentage points of the victory threshold. The average difference in referendum turnout remains significant in rural areas. However, almost all the differences in pre-treatment covariates, including the previous referendum turnout, between G-aligned and unaligned rural localities vanish, except the share of people in higher education, which is larger in G-aligned localities. A few differences in revenues and expenditures persist in the urban sample. Therefore, in order rule out that the few small differences confound the RDD estimates, I test formally the discontinuities in pre-treatment characteristics in the validity section 5.2. I also present the main estimates including these covariates on the right hand side in the results section.

5 Identification Strategy

I use closely-contested mayoral elections to identify the impact of mayor alignment with the *governing coalition* on referendum and parliamentary outcomes: locality turnout and locality vote shares. Hence, I exploit the sharp regression discontinuity design as in Lee (2008), comparing referendum outcomes in localities where an aligned candidate barely won with those where an aligned candidate barely lost the mayoral race. The vote margin between the aligned and the unaligned candidate is the running variable, based on which treatment is assigned, and the locality is treated if the aligned candidate vote margin is larger than the threshold 0. The vote margin is determined by: 1) the localities' characteristics (e.g. voters' preferences and choices); 2) pure chance. Lee (2008) showed that, as long the conditional probability density function of the running variable (conditioning on individuals' characteristics) is continuous, the pre-treatment characteristics are independent of treatment status in a tight neighbourhood around the threshold (local independence). This means that the variation in treatment status is due to chance, as in a natural randomized experiment.

In the case of voting, the identification strategy is based on the fact that, as long as this assumption holds, in closely contested elections, the electorates' (localities')

characteristics are similar in all respects except for treatment status. This is due to the inherent uncertainty about the final vote count, which, in principle, makes it impossible for the candidates (or voters) to control the vote margin perfectly. In simple terms, in very closely-contested elections, the partisan alignment treatment is as good as randomly assigned. Therefore, the average treatment effect can be estimated as the expected difference in referendum outcomes between treated and non-treated localities around the threshold. Below I expand on the econometric specifications and I also provide validity tests to confirm the random assignment.

5.1 Econometric specifications

In two-party races against aligned candidates, let dG_i be the vote share difference in locality i between the G-aligned and unaligned candidates. This vote margin is the *running variable*: if the dG_i is positive, then the G-aligned candidate wins the elections; if it is negative, he is the runner-up. The victory threshold is then $dG_i^* = 0$. For races in close proximity to this cutoff, the Average Treatment Effect (ATE) of partisan alignment is gauged from the discontinuity in observed outcomes at the cutoff. This can be estimated using the following simple linear regression model in tight intervals around the cutoff:

$$y_i = \alpha + \beta \cdot G\_wins_i + \varepsilon_i, \tag{1}$$

$$\text{with} \;\; G\_wins_i = 1\left[dG_i \geq 0\right] \text{ and } E\left[\varepsilon_i | dG_i\right] = 0,$$

where i indexes the locality and the dependent variable y_i is one of two outcomes: 1) referendum turnout and 2) "YES" vote share, in locality i. $\hat{\beta}$ is the estimated ATE of partisan alignment and $1\left[.\right]$ is the identity function.

I restrict the sample to intervals where the vote margins lie in the intervals [-2.5; +2.5] and [-5; +5] percentage points. Since I use data from one election year, these intervals yield quite small samples, trading off precision for validity.

An alternative estimation method which allows me to use the entire sample is a spline polynomial approximation with different parameters on the left and right of the threshold:

$$y_{ic} = \alpha + \beta_0 \cdot G\_wins_{ic} + \sum_{k=1}^{p} \delta_k \cdot dG_{ic}^k + G\_wins_{ic} \sum_{k=1}^{p} \beta_k \cdot dG_{ic}^k + \gamma' X_{ic} + \theta_c + \varepsilon_i, \tag{2}$$

where i indexes the locality and c indexes the county; y_{ic} is either the referendum turnout, or the share of "YES" votes, or the turnout and share votes for

G in parliamentary elections; X_{ic} is a vector of locality covariates and θ_c includes county fixed effects, included in some regressions to account for the fact that the victory thresholds may be county-specific;[25] p is some order of the polynomial in the victory margin dG_i (control function), which accounts for voters' preferences away from the victory threshold. Following the literature, I add the following in the list of covariates: 1) a set of demographic characteristics including the (log) size of population of the locality, the age, gender, ethnicity, education and unemployment rate by locality; 2) a set of fiscal pretreatment characteristics, including locality per capita revenues and expenditures; 3) pre-treatment electoral characteristics: the number of candidates and the turnout in the previous impeachment referendum, in 2007 (the Data section contains a more detailed account of the covariates).[26] Standard errors are clustered at county level.

5.2 Validity analysis

The RDD in elections has been used extensively in the literature, most notably to investigate incumbency advantages. A few recent studies have raised concerns with respect to the validity of the design, particularly in the U.S. close elections, where, e.g. Caughey and Sekhon (2011) have shown that stronger competitors are more likely to win elections by a narrow margin. Eggers et al. (2015) have surveyed a large number of electoral contexts and refuted the incumbency advantage in various countries, concluding that the U.S. case is an exception, and that the RDD is a sound approach in electoral analysis, provided that the main assumptions withstand thorough testing.

The coefficient $\hat{\beta}$ provides the unbiased estimate of the impact of partisan alignment on referendum outcome y if: 1) there is a discontinuity in treatment at the zero cutoff vote margin. 2) potential outcomes are a continuous function in the running variable at the threshold (Hahn et al., 2001). While this is an untestable assumption, Lee (2008) has shown that this can be replaced by a milder assump-

[25]Moreover, they account for the fact that counties have traditionally supported preponderantly certain political parties, and for the political alignment of the elected county council president.

[26]The electoral RD literature includes similar characteristics. E.g. Pettersson-Lidbom (2008) estimates the impact of left- vs. right- wing party control on economic outcomes, controlling for a number of predetermined characteristics e.g. income, population size, proportion of people below 15, and proportion of people above 65. Ferreira and Gyourko (2009) also use a RDD to estimate the difference in economic outcomes between localities with a Republican or a Democrat mayor. Their covariates include percentage white households, percentage with a college degree or more, household income, as well as predetermined fiscal outcomes in year t-1: total revenues per capita, total taxes per capita, total current expenditures per capita, total full-time employees per 1000 residents.

tion, namely that the probability density function of the running variable should be continuous at the cutoff, implying:

i) Locality unobservables do not vary discontinuously at the cutoff $dpx_l^* = 0$. While this cannot be tested directly, one can test the continuity of locality observed pre-treatment covariates at the threshold.

ii) The assignment mechanism is perfectly followed. This requires that candidates are not able to sort perfectly around the threshold (i.e. they cannot perfectly manipulate the vote share so as to win or lose elections). It is unlikely that candidates have perfect foresight on the number of ballots needed to tip the elections in their favor. However, below I perform standard tests to confirm the validity of this and the first assumption.

Going back to the treatment discontinuity, this condition is clearly satisfied because all winning candidates take office immediately after elections.[27]

Testing assumption i) requires looking for jumps in the locality observable characteristics around the threshold. Smooth locality covariates around the threshold would be reassuring that the RD estimate measures the average treatment effect of mayor alignment. A condensed test is to predict one of the outcomes using only the set of covariates, and to estimate the RD model with the predicted, instead of the realized outcomes, on the left-hand side. Insignificant RD estimates would confirm that covariates vary continuously around the cutoff.

This is confirmed in appendix Table A2, which presents the results from races against G, for predicted referendum turnout. I present the results from estimating model 1 (using a bandwidth of 5 percentage points in columns 1, 4 and 7) and model 3, without fixed effects (columns 2, 5 and 8) and also including county fixed effects (columns 3, 6 and 9). The outcomes are predicted by including successively: i) the subset including age, gender, education and ethnicity, unemployment status covariates (columns 1-3); an additional subset of fiscal covariates (columns 4-6); iii) an additional subset containing electoral covariates (columns 7-9). Panels A, B and C present races in all, rural, and urban localities, respectively. None of the RD coefficients are significant, suggesting that the observable characteristics should not confound the treatment effects. RD estimates for each individual covariate also suggest that almost all the pretreatment characteristics are continuous around the threshold (Table A3 in the appendix, for rural localities only). There are some significant differences in the share in higher educationand health expenditure, which remind that the close races disputed amongst any candidates may not be narrow

[27] In very few cases, mayors have given up their seat well into their mandate, so this does not affect the identification, because the outcomes are realized shortly after the elections.

enough to ensure a good counterfactual. When including these covariates in the main estimations, where, as shown in the results section, they do not change the treatment effect estimate much. In addition, the two-candidate races estimations should provide complementary evidence to mitigate this concern. While not irrefutable proof that unobservable characteristics vary continuously at the threshold, this evidence lends more credibility to the second assumption above.

The assumption of random assignment into treatment would be breached if mayoral candidates had perfect control over vote margins. Suppose aligned mayors manipulated the ballots to obtain an otherwise unattainable positive vote margin. Then the partisan alignment treatment effect would be confounded by the characteristics of candidates or places that enabled the manipulation, if they also altered national elections. One 'symptom' of such manipulation could be the discontinuity in the probability density function of running variable, i.e. the vote margin (McCrary, 2008). As pointed out in Lee (2008), with imperfect manipulation the treatment effect can still be identified, under the assumption of continuous conditional probability density function. McCrary (2008) developed a complementary test, based on the fact that Lee's assumption implies the continuity of the density function for the running variable. McCrary obtained smooth approximations of the running variable probability density functions to the right and to the left of the cutoff. The estimate of the discontinuity of probability density functions at the threshold can be used to test the null hypothesis of zero discontinuity.

In the case of mayoral elections, even if some candidates turn to vote buying to influence their winning probability, it is unlikely that they have complete control over final vote shares. Moreover, in closely-disputed elections, both candidates have the interest and the means to monitor each other. Nonetheless, graphical evidence of the distribution of vote margins is customary in testing the assumption of no perfect sorting around the threshold. Appendix Figure A1 displays the histograms for the running variable (Vote margin for G-aligned candidates), for all localities and separately by rural and urban localities. There is a visible, albeit small, difference in percentage localities with a G-aligned candidate just below and above the threshold. This is salient particularly in rural electoral races. To understand whether this is a problematic discontinuity, Figure A1 also displays McCrary's polynomial approximations of the probability density functions for all races of interest. A slight difference in the vote margin density for G-aligned candidates resembles the patterns in the histograms. However, the confidence intervals on the two sides of the threshold overlap, and the test statistics (t=1.03 for all, 1.21 for rural, 0.96 for ur-

ban localities) confirm that the zero discontinuity cannot be rejected and alleviates concerns of perfect sorting.

6 Results

6.1 Main estimates

This section presents the main estimations of the impact of mayors' partisan alignment on outcomes in the national impeachment referendum and parliamentary elections, identified from closely contested elections.

Figures 2 and 3 illustrate the essence of my findings. The figures plot regression function approximations for the referendum outcomes (Figure 2) and parliamentary outcomes (Figure 3) in races against G-aligned candidates. The approximations of the underlying regression functions use the data-driven selection of bins in Calonico et al. (2015).[28] Only Figure 2 displays a discontinuous increase in referendum turnout above the zero vote margin threshold, in races where a G-aligned candidate narrowly wins (Figure 2a, right). The sharp increase in turnout above the threshold is even more pronounced in rural localities, with the discontinuity estimate around 5 percentage points (Figure 2b, left). In urban areas, however, turnout is lower in barely successful G-aligned localities.[29] No discontinuous jump in turnout is present around the cutoff in legislative elections, which is consistent with the fact that turnout was not an objective in these polls (Figure 3).

In terms of vote shares, not only is there no discontinuity in the "YES" vote share at the referendum (Figure 2, right), but G vote shares in senate elections seem to be almost smooth around the cutoff.[30] The figures are suggestive of fundamental differences in incentives and voter mobilization at the two elections, which I discuss further below after estimating the alignment effects from RD regressions.

The main estimation results are displayed in Tables 4 (referendum outcomes) and 5 (parliamentary outcomes). The tables display a range of RD estimators, structured in three panels: Panel A shows results from all localities, Panel B restricts the sample to rural, and Panel C to urban localities.

The RD estimates from six specifications in both tables are displayed as follows: i) Columns (1)-(2) and (7)-(8) present OLS estimations in the tight intervals [-2.5;

[28]The Stata command from the robust data-driven regression discontinuity package developed by Calonico et al. (2015) is *rdbinselect*, using the evenly-spaced bins method.

[29]However, note that there is a lower density of winning G-aligned candidates in cities

[30]The same graphs with vote shares for G in the Lower Chamber are very similar and are available on request.

+2.5] and [-5; +5] percentage points around the threshold; ii) Columns (3) and (9) present estimates from robust local linear approximations with the optimal bin selection as in Calonico et al. (2014); iii) columns (4)-(6) and (10)-(12) present the results from 3rd order polynomial approximations with full samples, where the polynomial parameters are allowed to differ on the two sides of the cutoff; columns (5)-(6) and (11)-(12) include locality covariates and columns (6) and (12) further include county fixed effects.[31]

In Table 4, Panel A, the first estimate from the most restricted bandwidth indicates that G-aligned localities have 3.1 percentage points higher turnout that unaligned localities. Increasing the bandwidth decreases this estimate to around 2 percentage points, and none of the estimates in columns (1) - (4) are statistically significant. The effect of mayor alignment on turnout is large and significant in the specifications that allow for covariates and county fixed effects. The turnout premium from alignment with the governing coalition is 5.4 percentage points in the sixth column. On the other hand, all RD estimates of the difference in the share of "YES" votes are close to zero and insignificant. This evidence is in line with a turnout maximizing objective of the *governing coalition* in the presence of a participation quorum rule, as the theories of quorum referenda predict (e.g. Herrera and Mattozzi, 2010).

Panel B zooms in on races in rural localities. The turnout differences between communities with aligned and non-aligned mayors are clearly wider. All RD estimates are significant, and are around 5.3 percentage points in the tightest interval around the threshold and 6.2 percentage points in specifications with county fixed effects, and controlling for locality pre-treatment characteristics. The inclusion of covariates changes the estimate slightly from 4 to 5 percentage points. This is likely not a significant difference, but it does raise the question of whether treatment is truly randomly allocated. To reiterate, the issue is that a narrow top two candidates race is not the same as a narrow just two candidates race, and selection of G-aligned winner could play a role. Table 2 showed that most characteristics in the rural sample are the same in G-aligned and unaligned localities. However, as seen in appendix Table A3 which shows separate RD estimation for all covariates, the share in high education is slightly larger in G-aligned narrowly won localities. Assuming what the literature typically shows, that lower educated people are more easily manipulated,

[31]The estimation results with varying polynomial orders are similar, and are displayed in Table 9. Note also that introducing an additional set of covariates (share illiterate, share migrants, share working in agriculture and in public administration) makes the sample slightly smaller, but the results from that sample do not change when including these covariates. These results are not reported, but are available upon request.

one would expect the alignment estimate with no control for covariates to be underestimated. This is consistent with what columns 4 and 5 display. Nevertheless, just-two candidates races are discussed in section 6.3.

The county fixed effects also increase the coefficient's size. This is almost entirely driven by three counties. once these counties are excluded from the sample, the estimates are more consistent across specifications and the inclusion of county fixed effects does not alter the RD estimate size.[32]

The RD estimates in Panel C, from urban races show the opposite: there seems to be a negative turnout premium for G-aligned localities, although imprecisely estimated because of the small sample size. However, the reduction in magnitude in columns (5)-(6) suggests that the difference could be explained by municipality or county characteristics. Thus, the assumption of random assignment into alignment clearly does not hold in the urban sample (as anticipated by the significant differences in predetermined characteristics in Table 2). As in the full and rural samples, the RD estimates for the share of "YES" votes are not significant. Overall, there is a marked positive impact of partisan alignment on turnout. This effect is in line with the interests of the *governing coalition* to obtain the validating quorum, and it is driven exclusively by rural localities.[33]

Turning to the parliamentary elections in Table 5, the striking difference is that all RD estimates of mayoral alignment are close to zero and insignificant in all samples. These polls used closed lists, meaning that in each chamber and college (part of a city and/or a group of villages), the voter had one vote for the candidate from his preferred party. Hence, the voting decisions were arguably more complex than in the referendum. The absence of an alignment premium for turnout is expected given that turnout is not an objective. One possibility would be that, if government-aligned winning parties have the means to mobilize more voters, both the vote shares in their favor and turnout in their localities may exceed those in unaligned localities. However, the alignment estimates are also insignificant for vote shares in all specifications and samples. More specifications with varying polynomials are reported in

[32]The results excluding these three counties are available upon request. These counties do not seem to differ from the rest in terms of average number of local competitors, the vote share of top two parties in local elections, nor composition of races.

[33]Since the president's party had opposing interests, they may have tried to lower turnout. Treatment can thus also be defined as P-alignment. Appendix Table A4 reveals that the president's party close winners have lower turnout than close runner-ups, driven by rural localities. The magnitude of discontinuity estimates in slightly smaller than the analogous ones in Table 4. The estimates are only significant when controlling for covariates and county fixed effects. However, in this sample, most close races are between P and G, a competition which can result in lower alignment benefits, as discussed further in section 6.2. The usual validity checks hold for the P-alignment treatment, and available upon request.

Table 12, columns (1)-(3), where the alignment estimates for parliamentary turnout and vote shares are similarly small (slightly larger and significant only when fitting a linear control function). The conclusion from these estimations is that there is a much smaller impact, if at all, of alignment with the *governing coalition* on outcomes in the legislative elections that took place six months after the local elections. In section 7 below I explore these results further, looking at whether a larger turnout at the referendum may have persisted in parliamentary elections.

The positive alignment effect at the referendum may not be entirely surprising given the lack of popular support for the president at the time, and the 50% participation target, which the *governing coalition* needed for an almost certain victory. The more intriguing issue is how this mobilization was achieved in just over a month after the local elections and why, like in all previous studies, it is absent from subsequent national elections. In the space of one month pork-barrel spending on visible public projects would not have been feasible, but promises and expenditures on public projects may have ensued before the parliamentary elections. Vote buying and active "get-out-the-vote"strategies may have been used for both elections, as observers' reports show.

So what explains the asymmetry in these results? The most straightforward explanation could be that promises and vote buying can only overcome the voter commitment problem in the referendum, where turnout is easily observed, especially in small rural communities (see Nichter, 2008, for a discussion on why turnout buying is more likely to occur than vote buying). In the case of legislative elections, the voters may receive the benefits offered by the parties, but may still vote as they wish, since the vote is secret. This, of course, assumes the existence of vote buying practices, which, are widespread in many developing and transition countries. Below I conduct some heterogeneity tests to further highlight the role of vote-buying and also the local competition, or the power of local parties to keep the party in office in check.

6.2 Mechanisms

Exploring the differences across narrow races can give some idea about what drives the results and why the alignment effects have not been picked up in previous studies. In delivering votes, the corruption environment may matter, as well as the the party's experience in office in the past. For instance, as trial evidence and election monitoring reports showed, one method used was ballot stuffing in the names of people that could not vote, either because they were working abroad at the time of the election, or were deceased. Thus, in Table 6 I conduct a heterogeneity analysis

based on: the pre-treatment share of people who worked abroad (Panel A), pre-existing stated vote-buying norms (Panel B); party incumbency (Panel C). For space considerations, I focus only on referendum turnout in rural localities as the outcome, since all RD estimates for the vote shares and parliamentary outcomes are close to zero.

The results in panel A suggest that localities with migrant shares above the median 5% display larger and significant alignment effects in the expected directions in two out of three specifications. Note that the results are opposite to what one would expect in the absence of manipulation: an additional migrant is one less vote, so turnout should naturally be smaller.

The same holds for localities with more widespread vote buying practices. I use the 2011 Romanian Electoral Surveys data which asks individuals about electoral practices, including vote-buying. Not all localities are represented in the survey, hence I construct a county level measure of vote-buying norms. The share of people who report vote buying ranges from 6.25% to 87% with a standard deviation of 21%, which offers a good amount of variation. I divide counties by the median share of voters who admitted to have been asked to sell their vote (25%). The alignment premium seems higher in counties where vote buying is perceived as a more common practice (and insignificant in two specifications in low-vote buying prevalence counties).

Finally, the RD estimate is also consistently larger in localities where G was not in office in the previous mandate.[34] While the differences between localities where G was the incumbent before and those where G was a newcomer may not be statistically significant, this at least rules out differential voter responses to past pork-barrel spending in places where G was already in office. One potential explanation for this result is that the new G leadership needs to prove their loyalty to the party, which could attract future grants that help keeping their position. This may also flag a tendency to target swing voters (à la Lindbeck and Weibull, 1987), rather than strongly supportive localities where G is incumbent.

If one believes electoral fraud and vote buying account for the alignment effect, the party in office still needs unhindered access to the vote rigging apparatus. This may not be straightforward if the other parties, particularly the challengers for local office, have opposing interests and enough power to restrain the incumbent's misconduct. This points to the role of local competition (which brings about checks

[34]To capture the interaction between G's victory and G incumbency, here the treatment is 1 if G wins the race and the locality had a G mayor before 2012, and 0 if G loses the race regardless of party incumbency in their locality; in the case of non-incumbency, the treatment is 1 if G wins and the locality had a mayor from a different party before 2012

and balances) for the effect of party alignment. However, since RD estimation relies by default on races closely disputed, how should one go about finding a variation in local competition? The referendum setting brings the advantage that while all parties competed for local offices, only G and P actually competed at the referendum. Moreover, the *liberals/conservatives* represent a special case of local competition, but national alliance with G. Thus, P had high incentives to monitor G's activity around the referendum, and vice-versa, while *the liberals/conservatives* may have been more slack (note that this was preempted by the turnout differences in the various races in Table 1).

In Table 7 I show RD estimates from close races between: i) G and P (panel A); ii) G and any parties except P (panel B); iii) G and *liberals/conservatives* (panel C). The results in the three panels are fairly different: in races between G-aligned and P, the two parties with the largest stakes, the effects are very small and insignificant (except in the specification with county fixed effects). By contrast, when G wins over other parties than P, and in particular the *liberals/conservatives*, they get a significant turnout advantage (large and positive estimates). This is in line with the logic of tighter checks in the G versus P race, and potentially offsetting efforts on one another's electoral activities (be they vote rigging or voter mobilization). On the other hand, while the other parties may not influence turnout themselves, they may not have kept the G-mayors in check.

These results do not have a causal interpretation because the top two competitors' parties are selected based on voter's preferences. However, they are suggestive of the role of competition in vote delivery, an issue previously overlooked. This competition is even more important when all parties have strong stakes in national elections, such as legislative elections. Thus, it is likely that the parties' efforts to attract votes may cancel each other out in constituencies where there is a balance of powers, but not when a party dominates. Hence, in general, the RDD, which relies on balanced races, may not pick up alignment effects not because they do not occur, but because they occur away from the threshold.

6.3 Robustness Checks

In this section I return to the concerns about sample selection and treatment definition enunciated in section 4. The main estimates are based on a sample of all races where a G-aligned candidate comes first or second, regardless of how many other candidates there are. This choice reflects a trade-off between internal validity (sample selection based on voter preferences) and estimates' precision (larger sample). The risk is that a narrowly won and narrowly lost race by G candidates may

actually be quite different in unobserved characteristics (e.g. G wins with 47% to 45% votes and loses with 35% to 37%; the overall difference in support for G in the two races is 12 percentage points).

In what follows I describe various sensitivity and robustness checks, all focused mainly on the rural samples at the referendum, where the main alignment estimate was identified.

Firstly, I run the RD regressions in a restricted sample of rural localities where only two candidates (or up to three) candidates competed (an approach used in the main estimations in Brollo and Nannicini, 2012). The results shown in Table 8 confirm that the estimates are qualitatively similar to the baseline results, although imprecisely estimated due to the small sample size. The small sample estimates are generally larger in magnitude, with the exception of the county fixed effects specification (6) in Panel A. This, along with the descriptives in Table A1 displaying lower education levels in the small sample, raise concerns that the two-candidate races represent themselves a selected sample. Thus these estimates might be (potentially upward) biased. Another approach to use more of the original sample is to also include in the main sample races where a G-aligned candidate came third (an approach used in Migueis, 2013). Indeed, this reclaims roughly 200 additional observations compared to the baseline sample and the estimates are similar, only slightly larger than those in Table 4.

Secondly, I report the baseline RD estimates including control functions of varying polynomial orders, and the estimates displayed in Table 9 are consistent across all the different specifications. In addition, there is a concern that excess vote shares for G in local elections, may have generated a larger turnout if G supporters are politically more active. The parliamentary election turnout results suggest this is not the case, but as an additional test to dismiss a mechanical effect I proceed as follows. I predict referendum turnout using the local election vote shares, and I introduce this predicted variable as a control in the baseline RD regressions. The RD alignment estimate is reassuringly unaffected (results available upon request).

Finally, I revisit the treatment definition, where I now include the winning independent *liberals/conservatives* in the treatment rather than the control group. This automatically excludes narrow races between *social-democrats* and *liberals/conservatives*. As shown in Table 1 and also below in the heterogeneity tests, precisely these races displayed a large difference in turnout between winning and losing G localities. Hence, the redefined treatment is expected to at least deliver lower RD estimates than the baseline specifications. Table 10 results show smaller effects on referen-

dum turnout, significant only upon inclusion of controls and county fixed effects.[35] Panels B and C in Table 10 look at the redefined treatment effect on parliamentary outcomes, which remains insignificant. Overall these results again suggest that races within the governing coalition are the main drivers of the alignment effect.

7 Persistence of alignment: government transfers and following elections

7.1 Government transfers

The electoral advantage of alignment supports the political *quid pro quo* invoked in the intergovernmental transfers literature. Study after study have shown that transfers to local governments unequivocally follow the party in power. Accordingly, in this section I replicate the RD strategy for investigating the alignment effect on local government revenues after the 2012 local elections, consisting primarily of government redistributed funds.[36]

In Table 11 I report RD estimates from the baseline specifications, where the dependent variables are: Total locality per capita revenue (panel A); Locality-level per capita revenue from local taxation (own income, panel B); Locality-level per capita transfers for road infrastructure (panel C); Locality-level subsidies per capita (panel D). In the first three columns these dependent variables are measured in 2012, and in the last three columns they are measured in 2013.

The RD estimates are almost all positive and large, in particular for total locality revenue per capita. The most striking result is that the alignment effect on total revenue is larger and only significant in 2013. This is in stark contrast to total revenue in 2011, the pre-treatment year, when if anything, they were lower for localities won over by G in later local elections (see Tables 2 and 3, as well as appendix table A2). This is not driven by an increase in local tax revenue, which is not significantly larger in aligned localities. The difference can only be accounted for by government subsidies and discretionary transfers. Since I have only local budget data on subsidies, I look at the two potentially most visible to voters: roads and other subsidies, which do not display significant alignment effects (although note that again, at least in average terms, the balance of road subsidies has changed in

[35] I also redo the estimations with the original treatment, but only for localities where no candidate represented the coalition parties separately (perfect alignment between local and national parties). The results are very similar to those in Table 10 and are available upon request.

[36] Data on discretionary government transfers is not available for 2012-2013, therefore I use different categories of local revenues, which contain discretionary transfers.

favour of aligned localities). It is perhaps not surprising that the alignment premium for total revenue is smaller in 2012 than in 2013, since the G government only took office in May and the reshuffling of local administration seats only in June.

These results, at least on total revenue per capita, suggest a gradual reversal of the previous distribution of government funds, favouring G-aligned localities from 2012 onwards. Note also that these might underestimate the alignment bonus if transfers are strategic, and higher in the years just before new local elections, i.e. after 2013.[37]

7.2 Parliamentary elections

The core result in this paper is that local officials deliver votes for their parties in national polls, plausibly through vote buying under weak competition and easy enforcement of voter's promises. This effect is not detectable in elections with strong incentives for all parties and where voter's commitment cannot be secured, like in legislative elections. Yet, one could follow up asking whether the electoral advantage earned in the referendum persisted in parliamentary elections. Voter habituation and turnout anchoring through referenda have been documented before (Melton, 2014; Górecki, 2013), therefore I use the two national polls to investigate an indirect alignment effect.

OLS regressions in appendix table A5 indicate larger average turnouts in aligned localities in legislative elections. However, the alignment coefficient all but vanishes when controlling for referendum turnout, which is strongly and significantly correlated with legislative election turnout. This indicates that raising the turnout once might make it less costly to mobilize voters again. However, OLS incurs selection bias, which can be overcome with the RD.

As seen in Table 5, the direct alignment effect on parliamentary turnout and votes share was insignificant. In Table 12 columns 1-3, using the full rural sample and varying polynomials, I show at best modest alignment effects in parliamentary elections (significant only with the linear control function).

Then, I attempt to test whether the excess voters in aligned localities made any difference in legislative elections. Hence, I use the sharp RD from close mayoral race between aligned and unaligned localities as a source of exogenous variation in referendum turnout, conditional on locality and county characteristics. Hence, I estimate a two stage least squares model as in Van Der Klaauw (2002), where the baseline RD is the first stage, and the predicted referendum turnout $E[Turnout_i|dG_i] = \beta G\_wins_i + f(dG_i)$ is then used as the explanatory variable in

[37] This data is currently unavailable.

the second stage:

$$Yparl_i = \gamma_0 + \gamma_1 E[Turnout_i|dG_i] + k(dG_i) + \varepsilon_i, \quad (3)$$

where i indexes the locality and $Yparl_i$ is either the parliamentary election turnout or the vote share for G, $k(dG_i)$ is a control function that accounts for voter's preferences away from the local election victory threshold.[38]

The assumptions the narrowly-won alignment to be a good instrument are: i) instrument relevance (clearly holds, as seen in the first stage reported in Table 4); ii) instrument validity; i.e. alignment effects work exclusively through their impact on referendum turnout. The insignificant RD results in Table 5 are reassuring in this respect, but parties' campaigns and power of persuasion may have made a difference in referendum turnout and vote shares. The IV estimations will yield an unbiased estimate at best if the exclusion restriction conditional on locality characteristics holds.

Table 12, columns (4)-(6), displays the 2SLS estimates of referendum turnout on parliamentary election turnout and vote shares. The estimates are somewhat larger than the OLS estimates and are all significant, with 0.26 percentage points excess turnout and 0.55 percentage points increase in G vote share in legislative elections for 1 percentage point additional referendum turnout. Overall, the results support the persistence, or "stickiness" of referendum turnout. The implication is that the referendum with it's asymmetric incentives was a successful mobilization exercise for G, whose excess voters shifted the G vote share in the legislative elections, albeit by a minor margin.

8 Concluding Remarks

To date there is no clear evidence on whether and how local officials provide political capital for their parties, a frequent assumption in the voting and intergovernmental transfers literatures.

I provide robust evidence from an RD with closely-contested local elections that mayors bring electoral advantages for their parties in national polls. This effect may occur and may be detected in some favorable conditions: when the voter commitment problem can be easily overcome (hence, I identify an effect in a participation

[38] In Van Der Klaauw (2002) this identification strategy is used to examine the impact of college financial aid offers on students' decisions to enrol into university, where the authors explores financial aid offer discontinuities along a college aid score calculated for every students

quorum referendum but not in legislative elections); when vote buying and electoral fraud are entrenched norms (as effects are picked up only in rural areas and areas where vote buying is more widespread); and when other local competitors have incentives and power to monitor the incumbent. These partisan alignment effects may ensue in other contexts, but may not have been detected by previous studies due to the aggregation of results and the fact that they may occur in less competitive environments, i.e. away from the victory threshold in the RD.

I also present some evidence that partisan alignment produced an increase in aligned local government revenues, although I cannot causally link this to the increase in turnout. Furthermore, aligned localities seem to have increased revenues in the year following the referendum. Also, they generated higher voter mobilization in the referendum, which was partly transferred to the following national election.

These findings have important policy implications along at least two dimensions. Firstly, these findings raise concerns about the legitimacy of the political capital that local officials provide for the national parties, particularly in young inexperienced democracies. To the extent that voter mobilization involves illicit means like vote buying and electoral fraud, it undermines the most vital democratic freedom. This is the case in many developing countries, as well as countries in transition, where the legacy of the former autocratic regimes is still felt to this day, raising questions about how to institute healthy democracy (Keefer and Vlaicu, 2007; Martinez-Bravo, 2014). I contribute to this ongoing debate by showing that directly elected local officials can tamper with voters' choices. In this case, the referendum, the very expression of direct democracy, was turned into an instrument of manipulation, which may have propagated to the following national elections

Secondly, these results raise questions about the use and design of referenda as instruments for exercising healthy direct democracy. Referenda are nowadays increasingly used as policy-making tools in Europe and in the United States (Casella and Gelman, 2008). In line with the theory (see e.g. Herrera and Mattozzi, 2010), the high quorum requirement of 50% introduced asymmetric incentives for voter mobilization. As a result, the aligned localities boosted participation, while unaligned localities may have encouraged vote apathy. One way to address this issue is by balancing mobilization incentives, through setting a low participation quorum. This is precisely what happened in Romania following the 2012 electoral year, when the Parliament voted a new law that reduced the referendum participation quorum to 30%.

References

Ade, F. and R. Freier (2013). Divided government versus incumbency externality effectquasi-experimental evidence on multiple voting decisions. *European Economic Review 64*(0), 1 – 20.

Ames, B. (1994, 3). The reverse coattails effect: Local party organization in the 1989 brazilian presidential election. *American Political Science Review 88*, 95–111.

Biro, A. (2012, July). Referendum 2012. reportaj sub acoperire din mijlocul armatei psd. *Gandul*. Online; posted 25-July-2012 (In Romanian).

Bracco, E., B. Lockwood, F. Porcelli, and M. Redoano (2015). Intergovernmental grants as signals and the alignment effect: Theory and evidence. *Journal of Public Economics 123*, 78 – 91.

Brollo, F. and T. Nannicini (2012). Tying your enemy's hands in close races: The politics of federal transfers in brazil. *American Political Science Review 106*, 742–761.

Broockman, D. E. (2009). Do congressional candidates have reverse coattails? evidence from a regression discontinuity design. *Political Analysis 17*(4), 418–434.

Buti, D. (2012). Local election. domination and imbalance. *Sfera Politicii*, 34– 42.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal 14*(4), 909–946(38).

Calonico, S., M. D. Cattaneo, and R. Titiunik (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*.

Casella, A. and A. Gelman (2008). A simple scheme to improve the efficiency of referenda. *Journal of Public Economics 92*(1011), 2240 – 2261.

Caughey, D. and J. S. Sekhon (2011). Elections and the regression discontinuity design: Lessons from close u.s. house races, 19422008. *Political Analysis 19*(4), 385–408.

Cohen, J. E., J. R. Bond, R. Fleisher, and J. A. Hamman (2000). State-level presidential approval and senatorial support. *Legislative Studies Quarterly 25*(4), 577 – 590.

Durante, R. and B. Knight (2012). Partisan control, media bias, and viewer responses: Evidence from berlusconi's italy. *Journal of the European Economic Association 10*(3), 451–481.

Durante, R., P. Pinotti, and A. Tesei (2014, May). No News, Big News. The political consequences of entertainment TV. Working Papers 063, "Carlo F. Dondena" Centre for Research on Social Dynamics (DONDENA), Universit Commerciale Luigi Bocconi.

EFOR (2013). Political clientelism. Online report, http://expertforum.ro/en/clientelism-map/, last updated in February 2013 (In English).

Eggers, A. C., A. Fowler, J. Hainmueller, A. B. Hall, and J. M. Snyder (2015). On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science 59*(1), 259–274.

Ferreira, F. and J. Gyourko (2009). Do political parties matter? evidence from u.s. cities. *The Quarterly Journal of Economics 124*(1), 399–422.

Finan, F. and L. Schechter (2012). Vote-buying and reciprocity. *Econometrica 80*(2), 863–881.

Freedom-House (2013). Nations in Transit, Romania. Technical report. Accessed 5 February 2015 at https://freedomhouse.org/report/nations-transit/2013/romania.

Funk, P. (2010, 09). Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System. *Journal of the European Economic Association 8*(5), 1077–1103.

Gélineau, F. and K. L. Remmer (2006, 1). Political decentralization and electoral accountability: The argentine experience, 19832001. *British Journal of Political Science 36*, 133–157.

Gelman, A. and G. Imbens (2014, August). Why High-order Polynomials Should not be Used in Regression Discontinuity Designs. (20405).

Golder, M. (2006). Presidential Coattails and Legislative Fragmentation. *American Journal of Political Science 50*, 34–48.

Górecki, M. A. (2013). Electoral context, habit-formation and voter turnout: A new analysis. *Electoral Studies*, 140152.

Grossman, P. (1994). A political theory of intergovernmental grants. *Public Choice 78*(3-4), 295–303.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Herrera, H. and A. Mattozzi (2010). Quorum and turnout in referenda. *Journal of the European Economic Association 8*(4), 838–871.

IRES (2012). The referendum for impeachment of romania's president, 29 july 2012 - the analysis of electoral behaviour and post-referendum analyses.

Keefer, P. and R. Vlaicu (2008). Democracy, credibility, and clientelism. *Journal of Law, Economics, and Organization 24*(2), 371–406.

Klasnja, M. (2014). Corruption and incumbency disadvantage: Theory and evidence. *Working paper*.

Lee, D. S. (2008, February). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics 142*(2), 675–697.

Lindbeck, A. and J. W. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice 52*(3), 273–297.

Manacorda, M., E. Miguel, and A. Vigorito (2011). Government transfers and political support. *American Economic Journal: Applied Economics 3*(3), 1–28.

Martinez-Bravo, M. (2014). The role of local officials in new democracies: Evidence from indonesia. *American Economic Review 104*(4), 1244–87.

Mattei, F. and J. Glasgow (2005). Presidential coattails, incumbency advantage, and open seats: A district-level analysis of the 19762000 u.s. house elections. *Electoral Studies 24*(4), 619 – 641.

McCrary, J. (2008, February). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Melton, J. (2014). Why is vote habit-forming? *mimeo*.

Migueis, M. (2013). The effect of political alignment on transfers to portuguese municipalities. *Economics and Politics 25*(1), 110–133.

National-Anticorruption-Directorate (2013). Press release No. 843/VIII/3. Last updated 7 October 2013: http://www.pna.ro/comunicat.xhtml?id=4342 (In Romanian).

Nichter, S. (2008, 2). Vote buying or turnout buying? machine politics and the secret ballot. *American Political Science Review 102*, 19–31.

Persico, N., J. C. R. Pueblita, and D. Silverman (2011). Factions and Political Competition. *Journal of Political Economy 119*(2), 242 – 288.

Pettersson-Lidbom, P. (2008). Do parties matter for economic outcomes? a regression-discontinuity approach. *Journal of the European Economic Association 6*(5), 1037–1056.

Politeanu, M. (2012). Ponta and the rule of law. *Hungarian Review 5*, 16 – 21.

Pop-Eleches, C. and G. Pop-Eleches (2012, June). Targeted Government Spending and Political Preferences. *Quarterly Journal of Political Science 7*(3), 285–320.

Riker, W. and P. Ordeshook (1968). A Theory of the Calculus of Voting. *American Political Science Review 62*, 25 – 43.

Robinson, J. A. and T. Verdier (2013). The political economy of clientelism. *Scandinavian Journal of Economics 115*, 260–291.

Samuels, D. (2000b). Concurrent elections, discordant results: presidentialism, federalism, and governance in brazil. *Comparative Politics*, 1–20.

Samuels, D. J. (2000a). The gubernatorial coattails effect: Federalism and congressional elections in brazil. *Journal of Politics 62*(1), 240–253.

Seceleanu, A. (2009). Mediatization of political life in conditions of electoral campaign–a retrospection of the elections in romania. *Central European Journal of Communication* (03), 267–278.

Snyder, J. M. and S. Levitt (1995, 11/1995). "political parties and the distribution of federal outlays". *American Journal of Political Science 39*, 958–980.

Solé-Ollé, A. and P. Sorribas-Navarro (2008). The effects of partisan alignment on the allocation of intergovernmental transfers. differences-in-differences estimates for spain. *Journal of Public Economics 92*(12), 2302–2319.

Van Der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on College Enrollment: A RegressionDiscontinuity Approach*. *International Economic Review 43*(4), 1249–1287.

Vicente, P. C. (2014). Is vote buying effective? evidence from a field experiment in west africa. *The Economic Journal 124*(574), F356–F387.

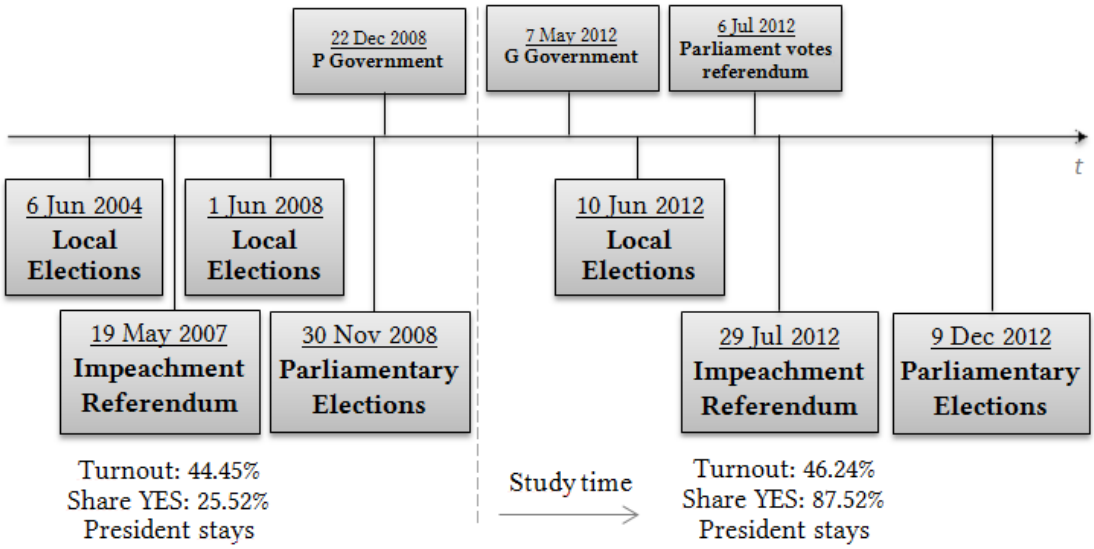
Volintiru, C. (2012). Clientelism: electoral forms and functions in the romanian case study. *Romanian Journal of Political Sciences* (01), 35–66.

Worthington, A. C. and B. E. Dollery (1998). The political determination of intergovernmental grants in australia. *Public Choice 94*(3-4), 299–315.

Zudenkova, G. (2011). A political agency model of coattail voting. *Journal of Public Economics 95*(11), 1652–1660.

Figures and Tables

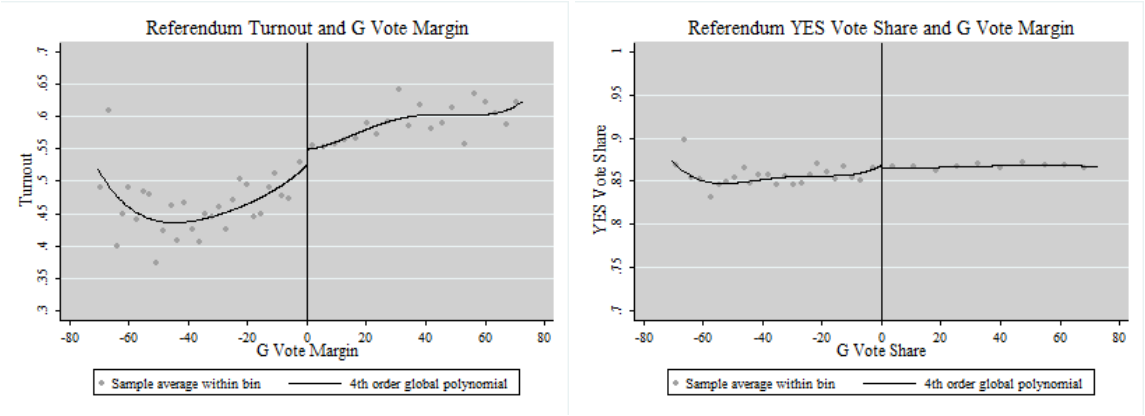
Figure 1 Timeline of electoral events



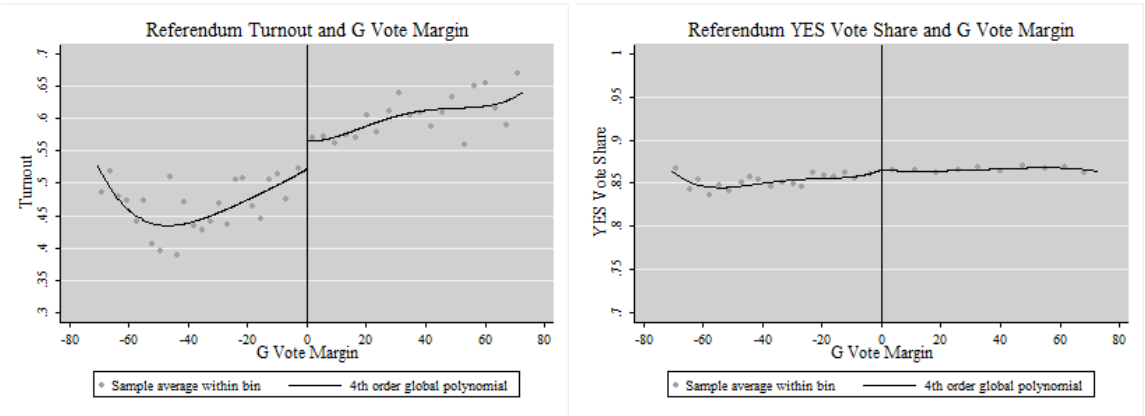
Note: The P Government was formed by the president's party (*the democrat-liberals*). The G Government was formed by the *social-liberal union*, the main opponents to the president's party rule.

Figure 2: G-alignment and referendum outcomes

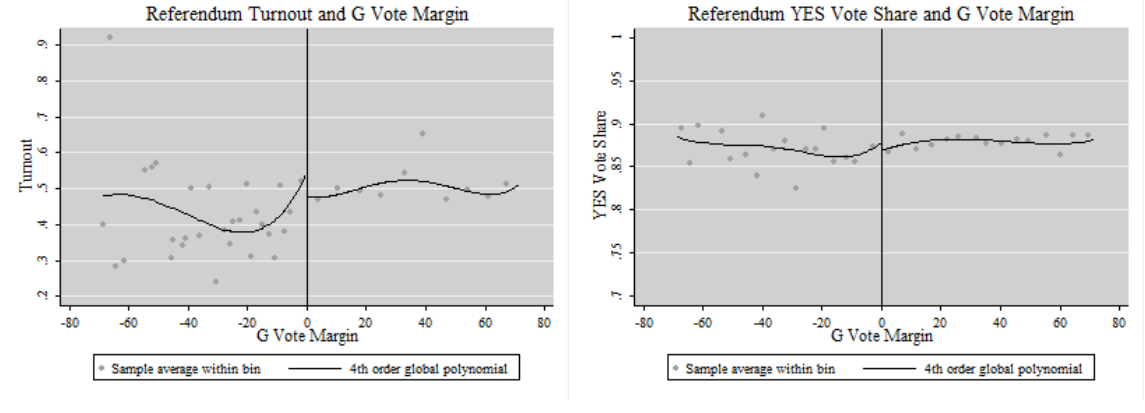
(a) All localities



(b) Rural localities



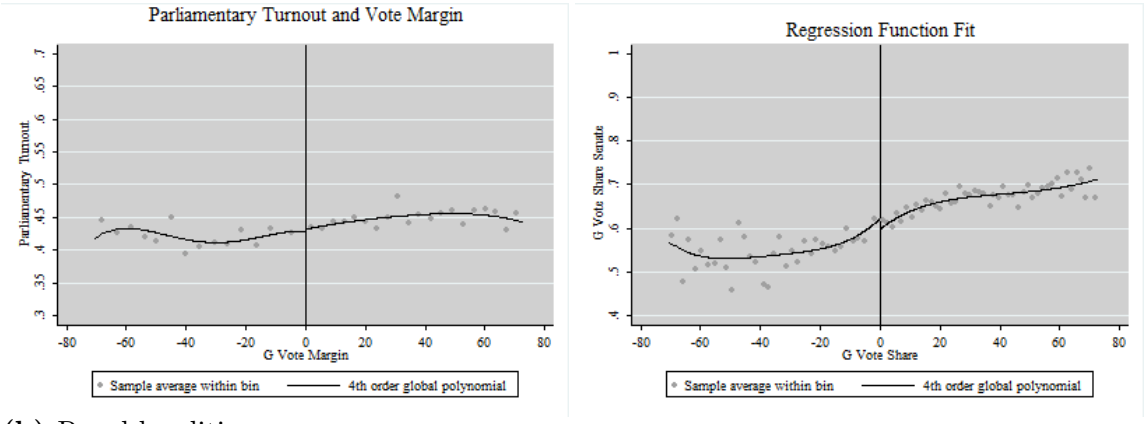
(c) Urban localities



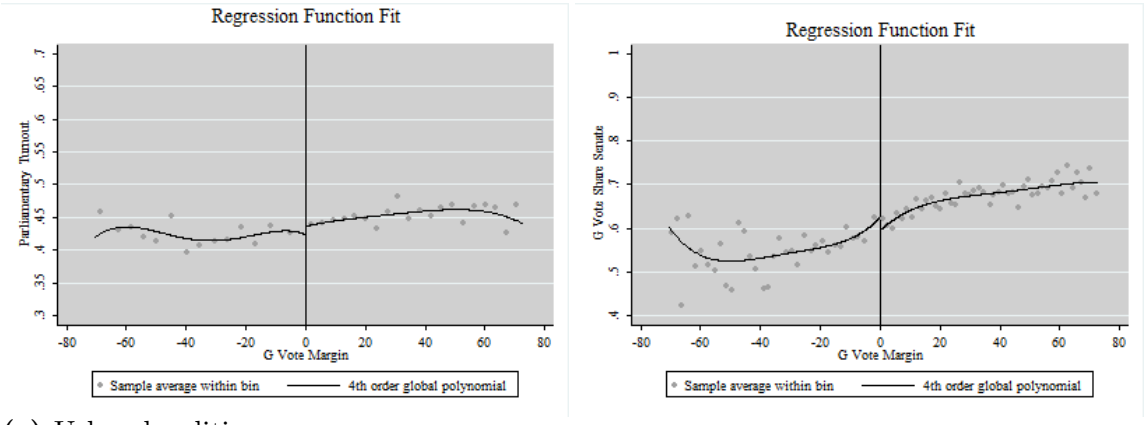
Note: The figure displays the polynomial approximations of the referendum turnout (left) and "YES" vote share (right) plotted against the running variable on the X-axis (Victory Margin for G in races against G).

Figure 3: G-alignment and Parliamentary election outcomes

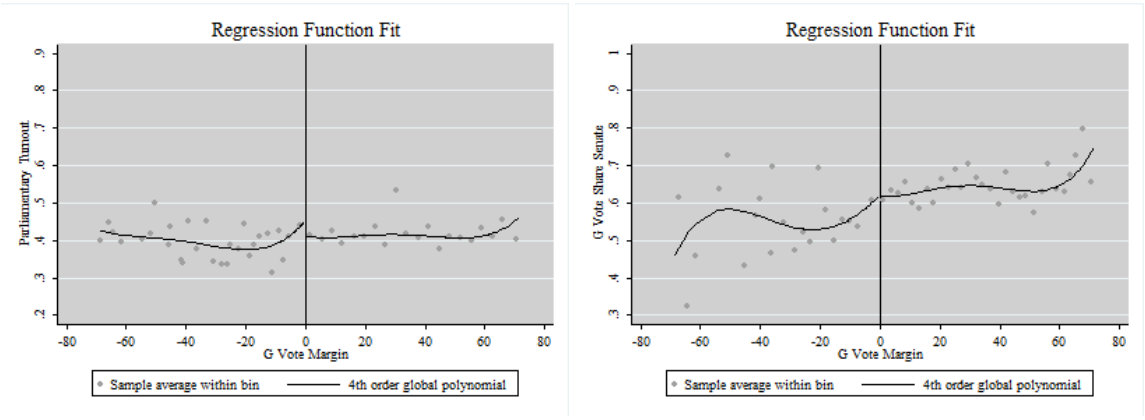
(a) All localities



(b) Rural localities



(c) Urban localities



Note: The figure displays the polynomial approximations of the referendum turnout (left) and "YES" vote share (right) plotted against the running variable on the X-axis (Victory Margin for G in races against G).

Table 1: Referendum Turnout in races within 5 percentage points from the victory cutoff

Race (Top 2 parties)	G wins	SD wins	P wins	L/C wins	Difference
Panel A: Races between G and P, SD and P, P and L/C, SD and L/C					
G(=SD+L/C) vs P	0.547		0.528		0.019
	(0.152)		(0.159)		
	n=82		n=68		
SD vs P		0.541	0.507		0.034
		(0.118)	(0.179)		
		n=20	n=17		
P vs L/C			0.451	0.453	-0.002
			(0.191)	(0.117)	
			n=12	n=13	
SD vs L/C		0.649		0.562	**0.087***
		(0.152)		(0.133)	
		n=25		n=17	
Panel B: Main G treatment= party SD or coalition SD+L/C					
G(=SD or SD+L/C) vs P	0.546		0.524		0.022
	(0.145)		(0.162)		
	n=102		n=85		
G(=SD or SD+L/C) vs L/C	0.650			0.551	**0.101****
	(0.149)			(0.139)	
	n=26			n=18	

Note: The table displays the comparison in means in referendum turnout between localities for different *narrow* races. SD stands for *social-democrats*, L/C is *liberals/conservatives*, P stands for *the president's party*. Standard deviations are reported in parantheses. The differences in means are reported for each type of race in the last column. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Descriptive statistics. Full samples of races against G

	All sample			Rural			Urban		
	G	**non-G**	**P-value**	**G**	**non-G**	**P-value**	**G**	**non-G**	**P-value**
Panel A: Outcomes									
Turnout	0.583	0.472	**0.000**	0.595	0.476	**0.000**	0.497	0.439	**0.003**
Share YES	0.867	0.857	**0.000**	0.866	0.855	**0.000**	0.878	0.869	**0.049**
Turnout Parliamentary Elections	0.447	0.422	**0.000**	0.452	0.425	**0.000**	0.412	0.402	0.243
G Vote Share in Senate	0.661	0.562	**0.000**	0.665	0.562	**0.000**	0.636	0.561	**0.000**
Panel B: Covariates									
No. candidates in local elections	4.650	4.517	**0.055**	4.443	4.344	0.128	6.208	5.969	0.387
Sum % votes top 2 local candidates	85.139	86.297	**0.011**	85.718	86.802	**0.023**	80.783	82.061	0.363
Turnout referendum 2007	44.746	45.213	**0.000**	44.943	45.273	0.471	43.265	44.706	0.186
Log Population	8.157	8.115	0.227	7.957	7.957	0.987	9.660	9.444	0.102
Share adult population	0.771	0.769	0.439	0.770	0.768	0.340	0.779	0.783	0.435
Share over 65	0.204	0.203	0.833	0.212	0.209	0.296	0.141	0.152	**0.005**
Share males	0.495	0.496	0.458	0.497	0.497	0.703	0.486	0.487	0.449
Share high education	0.046	0.045	0.398	0.037	0.037	0.903	0.116	0.110	0.379
Share high school	0.159	0.156	0.155	0.147	0.146	0.491	0.250	0.239	**0.097**
Share Romanians	0.893	0.875	**0.002**	0.898	0.881	**0.005**	0.854	0.824	**0.061**
Unemployment rate	0.049	0.049	0.844	0.051	0.051	0.963	0.033	0.035	0.380
Per capita revenue	1366.051	1465.576	**0.010**	1335.907	1443.941	**0.009**	1592.916	1647.126	0.628
Per capita own revenue	485.207	492.102	0.646	439.949	463.897	**0.079**	825.824	728.785	0.156
Per capita roads funds	11.227	11.603	0.799	12.228	12.559	0.841	3.694	3.585	0.932
Per capita subsidies	98.412	140.310	**0.000**	95.892	127.683	**0.003**	117.373	246.270	**0.001**
Per capita expenditures	1306.659	1396.277	**0.012**	1270.828	1365.327	**0.010**	1576.322	1655.995	0.504
Per capita expenditures education	357.241	368.331	0.137	331.726	350.470	**0.003**	549.264	518.215	0.394
Per capita health expenditures	9.582	9.316	0.804	6.837	7.594	0.433	30.240	23.769	0.211
Per capita public expenditures	175.527	200.354	**0.077**	170.342	192.366	0.148	214.552	267.392	0.119
N	1475	911		1302	814		173	97	

Note: The table displays the comparison in means in outcome variables (Panel A) and locality characteristics (Panel B) between localities aligned and unaligned with G, for all races against G, and separately for rural and urban localities. The p-values for the differences in means are reported (significant differences marked in bold).

Table 3: Descriptive statistics. Samples of races against G within 5% margins around cutoff

	All sample			Rural			Urban		
	G	**non-G**	**P-value**	**G**	**non-G**	**P-value**	**G**	**non-G**	**P-value**
Panel A: Outcomes									
Turnout	0.553	0.528	0.176	0.566	0.531	**0.088**	0.479	0.514	0.403
Share YES	0.864	0.868	0.457	0.863	0.866	0.613	0.868	0.877	0.412
Turnout Parliamentary Elections	0.435	0.430	0.694	0.437	0.429	0.514	0.423	0.438	0.526
G Vote Share in Senate	0.611	0.613	0.892	0.609	0.614	0.775	0.624	0.612	0.677
Panel B: Covariates									
No. candidates in local elections	5.222	5.244	0.917	5.036	5.076	0.850	6.292	6.045	0.700
Sum % votes top 2 local candidates	79.985	79.669	0.851	80.210	79.722	0.795	78.688	79.413	0.
Turnout referendum 2007	44.783	44.478	0.529	44.751	43.974	0.544	44.965	46.882	0.450
Log Population	8.152	8.177	0.819	7.926	7.943	0.826	9.453	9.293	0.657
Share adult population	0.768	0.772	0.523	0.767	0.768	0.869	0.773	0.788	0.191
Share over 65	0.203	0.209	0.400	0.213	0.219	0.472	0.145	0.163	**0.045**
Share males	0.496	0.495	0.884	0.497	0.497	0.933	0.488	0.488	0.827
Share high education	0.050	0.048	0.775	0.040	0.034	**0.082**	0.106	0.117	0.586
Share high school	0.163	0.160	0.663	0.149	0.141	0.261	0.247	0.251	0.795
Share Romanians	0.883	0.880	0.884	0.885	0.889	0.850	0.868	0.840	0.464
Unemployment rate	0.051	0.053	0.730	0.054	0.056	0.743	0.033	0.037	0.510
Per capita revenue	1390.731	1405.986	0.890	1400.887	1338.739	0.624	1332.330	1726.939	**0.022**
Per capita own revenue	483.665	455.977	0.443	447.069	409.855	0.313	694.090	676.106	0.850
Per capita roads funds	11.005	14.989	0.506	11.980	16.919	0.485	5.396	5.780	0.932
Per capita subsidies	113.401	149.042	0.335	125.291	109.617	0.690	45.035	337.206	**0.004**
Per capita expenditures	1281.949	1346.232	0.505	1276.639	1267.771	0.935	1312.485	1720.706	**0.019**
Per capita expenditures education	364.864	375.396	0.631	337.243	345.181	0.715	523.681	519.604	0.945
Per capita health expenditures	7.884	12.802	0.101	6.278	10.699	0.184	17.117	22.839	0.358
Per capita public expenditures	137.354	162.911	0.304	138.022	137.195	0.975	133.511	285.646	**0.025**
N	162	127		138	105		24	22	

Note: The table displays the comparison in means in outcome variables (Panel A) and locality characteristics (Panel B) between localities aligned and unaligned with G, for all, rural and urban localities within 5 percentage points victory margins. The p-values for the differences in means tests are reported (significant differences marked in bold).

Table 4: Alignment and referendum outcomes. RD Estimates

	Referendum Turnout						**Share YES**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	OLS	Local LR	Poly.	Approx.	Full sample	OLS	OLS	Local LR	Poly.	Approx.	Full sample
					Panel A: All localities							
G wins	0.031	0.025	0.020	0.021	0.037***	0.054***	-0.005	-0.005	-0.006	0.002	0.003	0.006
	(0.021)	(0.015)	(0.021)	(0.020)	(0.013)	(0.014)	(0.008)	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)
Obs.	143	289	1,057	2,386	2,386	2,386	143	289	921	2,386	2,386	2,386
					Panel B: Rural localities							
G wins	0.053**	0.035*	0.041*	0.040*	0.050***	0.062***	-0.002	-0.004	-0.002	0.003	0.004	0.006
	(0.024)	(0.018)	(0.022)	(0.022)	(0.016)	(0.015)	(0.009)	(0.007)	(0.008)	(0.008)	(0.009)	(0.008)
Obs.	118	243	949	2,116	2,116	2,116	118	243	942	2,116	2,116	2,116
					Panel C: Urban localities							
G wins	-0.078	-0.034	-0.106**	-0.079	-0.029	0.017	-0.016	-0.009	-0.015	-0.007	-0.005	-0.001
	(0.051)	(0.040)	(0.053)	(0.049)	(0.054)	(0.035)	(0.012)	(0.011)	(0.013)	(0.013)	(0.011)	(0.011)
Obs.	25	46	113	270	270	270	25	46	119	270	270	270
County FE	No	No	No	No	No	Yes	No	No	No	No	No	Yes
Controls	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Specification	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on referendum turnout (columns 1-6) and the "YES" vote share (columns 7-12). Results are based on all localities (Panel A), rural (Panel B) and urban (Panel C). Estimates from simple linear regression in a small interval around the cutoff in columns (1)-(2) and (7)-(8). Estimates from local linear regression using the Calonico et al. (2014) robust bias-corrected confidence intervals in columns (3) and (9). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (4)-(6) and (10)-(12). The controls in columns 5-6 and 11-12 include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Alignment and parliamentary election outcomes. RD Estimates

	Parliamentary Turnout						G Share Votes Senate					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	OLS	Local LR	Poly.	Approx.	Full sample	OLS	OLS	Local LR	Poly.	Approx.	Full sample
						Panel A: All localities						
G wins	0.006	0.004	0.000	-0.002	0.004	0.006	-0.003	-0.002	-0.021	-0.017	-0.009	0.002
	(0.012)	(0.010)	(0.013)	(0.012)	(0.011)	(0.009)	(0.019)	(0.013)	(0.016)	(0.017)	(0.016)	(0.015)
Obs.	143	289	1,048	2,386	2,386	2,386	143	289	999	2,386	2,386	2,386
						Panel B: Rural localities						
G wins	0.016	0.008	0.009	0.007	0.010	0.011	-0.002	-0.004	-0.023	-0.015	-0.007	0.003
	(0.015)	(0.012)	(0.015)	(0.014)	(0.013)	(0.011)	(0.022)	(0.014)	(0.018)	(0.020)	(0.018)	(0.017)
Obs.	118	243	897	2,116	2,116	2,116	118	243	888	2,116	2,116	2,116
						Panel C: Urban localities						
G wins	-0.040	-0.015	-0.048	-0.045*	-0.013	0.017	-0.013	0.012	-0.002	-0.026	-0.028	-0.011
	(0.036)	(0.021)	(0.033)	(0.025)	(0.023)	(0.022)	(0.036)	(0.027)	(0.035)	(0.030)	(0.032)	(0.032)
Observations	25	46	121	270	270	270	25	46	134	270	270	270
County FE	No	No	No	No	No	Yes	No	No	No	No	No	Yes
Controls	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Specification	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on parliamentary elections turnout (columns 1-6) and the senate G vote share (columns 7-12). Results are based on all localities (Panel A), rural (Panel B) and urban (Panel C). Estimates from simple linear regression in a small interval around the cutoff in columns (1)-(2) and (7)-(8). Estimates from local linear regression using the Calonico et al. (2014) robust bias-corrected confidence intervals in columns (3) and (9). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (4)-(6) and (10)-(12). The controls in columns 5-6 and 11-12 include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Rural races heterogeneity: migration, vote buying and incumbency.

	Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Poly.	Approx. Full sample	OLS	Poly.	Approx. Full sample
	Panel A: Migration					
	Share work abroad <5%			Share work abroad >= 5%		
G wins	-0.008	0.031	0.070**	0.051**	0.040	0.061***
	(0.027)	(0.035)	(0.029)	(0.025)	(0.028)	(0.020)
Obs.	98	906	906	145	1,210	1,210
	Panel B: Vote Buying Norms					
	County share vote buying <= 25%			County share vote buying > 25%		
G wins	0.025	0.012	0.050**	0.052***	0.068**	0.073***
	(0.026)	(0.031)	(0.020)	(0.015)	(0.025)	(0.023)
Obs.	155	1,283	1,283	88	833	833
	Panel C: Incumbency					
	G mayor in 2008			Other mayor in 2008		
G wins	0.014	0.008	0.043*	0.040*	0.056**	0.066***
	(0.028)	(0.035)	(0.022)	(0.022)	(0.025)	(0.017)
Obs.	152	1,371	1,371	180	1,401	1,401
Controls	No	No	Yes	No	No	Yes
County FE	No	No	Yes	No	No	Yes
Specification	[-5; +5]	3rd Ord. Poly.	3rd Ord. Poly.	[-5; +5]	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on referendum turnout in rural localities, by: the share of migrants abroad (Panel A) below median in columns (1)-(3) and above median columns (4)-(6); the share who report vote buying (Panel B) below median in columns (1)-(3) and above median columns (4)-(6); whether G was in office before 2012, in columns (1)-(3), or not, in columns (4)-(6). Estimates from simple linear regression in a small interval around the cutoff in columns (1) and (4). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (2)-(3) and (5)-(6). The measure of perceived vote prevalence is computed based on the Romanian Electoral Surveys 2011 questions on people's observations or perceived incidence of vote buying. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Rural races heterogeneity: different parties' races

	Turnout		
	(1)	(2)	(3)
	OLS	Poly. Approx.	Full sample
	Panel A: Races between G and P		
G wins	0.028	0.007	0.052***
	(0.022)	(0.027)	(0.016)
Obs.	161	1,359	1,359
	Panel B: Races between G and non-P		
G wins	0.044	0.081***	0.058**
	(0.031)	(0.028)	(0.025)
Obs.	82	757	757
	Panel C: Races between G and L-C		
G wins	0.100*	0.144**	0.039
	(0.049)	(0.057)	(0.030)
Obs.	39	322	322
Controls	No	No	Yes
County FE	No	No	Yes
Specification	[-5; +5]	3rd Order Poly.	3rd Order Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on referendum turnout in rural localities in different close races: between G and P (Panel A), between G and other parties except P (Panel B) and between G and the *liberals/conservatives* (Panel C). Estimates from simple linear regression in a small interval around the cutoff in column (1). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (2)-(3). Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Rural Alignment and referendum turnout. Robustness to race structure

	Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	Local LR	Poly.	Approx.	Full sample
	Panel A: Two-candidates races in rural localities					
G wins	0.125*	0.106*	0.062	0.100	0.127	0.018
	(0.052)	(0.049)	(0.069)	(0.091)	(0.081)	(0.093)
Obs.	6	9	63	150	150	150
	Panel B: Two- and three-candidates races in rural localities					
G wins	0.017	-0.016	0.056	0.043	0.088**	0.104**
	(0.068)	(0.057)	(0.059)	(0.053)	(0.040)	(0.047)
Obs. 18	34	230	604	604	604	
	Panel C: All G races in rural localities					
G wins	0.048**	0.034*	0.038*	0.044**	0.058***	0.068***
	(0.023)	(0.017)	0.038*	(0.021)	(0.015)	(0.015)
Obs.	120	247	934	2,301	2,301	2,301
County FE	No	No	No	No	No	Yes
Controls	No	No	No	No	Yes	Yes
Specification	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on referendum turnout in rural localities, in different races: two-candidate races (Panel A); three-candidate races (Panel B); nearly all G races (including those where G came third, Panel C). Estimates from simple linear regression in a small interval around the cutoff in columns (1)-(2). Estimates from local linear regression using the Calonico et al. (2014) robust bias-corrected confidence intervals in column (3). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (4)-(6). The controls in columns 5-6 and 11-12 include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Alignment and referendum turnout. Alternative polynomials

	Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
	Poly. Approx. Full sample					
	Panel A: All localities					
G wins	0.055***	0.066***	0.016	0.050***	0.025	0.041**
	(0.011)	(0.008)	(0.017)	(0.010)	(0.020)	(0.016)
Obs.	2,386	2,386	2,386	2,386	2,386	2,386
	Panel B: Rural localities					
G wins	0.058***	0.072***	0.028	0.062***	0.044*	0.050***
	(0.011)	(0.008)	(0.017)	(0.010)	(0.024)	(0.017)
Obs.	2,116	2,116	2,116	2,116	2,116	2,116
County FE	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes
Specification	1st Ord. Poly.	1st Ord. Poly.	2nd Ord. Poly.	2nd Ord. Poly.	4th Ord. Poly.	4th Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on referendum turnout in rural localities, using alternative polynomials orders. All localities (Panel A) and rural localities (Panel B). Estimates using polynomial approximations of first order in columns (1)-(2), second order in columns (3)-(4) and fourth order in columns (5)-(6), without controls, with controls and with county fixed effects. The controls include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Alignment, referendum and parliamentary outcomes. Alternative treatment

	Referendum and Parliamentary Outcomes					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	Local LR	Poly.	Approx.	Full sample
	Panel A: Referendum Turnout rural					
G or L-C wins	0.021	0.021	0.011	0.009	0.035**	0.057***
	(0.028)	(0.020)	(0.022)	(0.026)	(0.016)	(0.014)
Obs.	114	238	1,002	1,961	1,961	1,961
	Panel B: Parliamentary Turnout rural					
G or L-C wins	0.001	0.008	0.002	0.003	0.013	0.020*
	(0.014)	(0.012)	(0.015)	(0.014)	(0.011)	(0.011)
Obs.	114	238	855	1,961	1,961	1,961
	Panel C: Senate G Vote shares rural					
G or L-C wins	0.017	0.013	-0.008	-0.015	-0.005	0.008
	(0.021)	(0.015)	(0.018)	(0.018)	(0.015)	(0.014)
Obs.	114	238	825	1,961	1,961	1,961
County FE	No	No	No	No	No	Yes
Controls	No	No	No	No	Yes	Yes
Specification	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G, including independently running *liberals/conservatives*) alignment on referendum (Panel A) and parliamentary outcomes (Panels B and C) in rural localities. Estimates from simple linear regression in a small interval around the cutoff in columns (1)-(2). Estimates from local linear regression using the Calonico et al. (2014) robust bias-corrected confidence intervals in column (3). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (4)-(6). The controls in columns (5)-(6) and (11)-(12) include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Alignment and locality revenues in 2012-2013 (rural). RD Estimates

	Locality Fiscal Revenues					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Poly.	Approx.	OLS	Poly.	Approx.
			Full sample			Full sample
	Panel A: Total per capita income					
	2012			2013		
G wins	166.063	85.920	108.883	301.383***	228.842	222.262*
	(111.149)	(135.762)	(110.568)	(105.005)	(141.873)	(129.179)
Obs.	243	2,116	2,116	243	2,116	2,116
	Panel B: Local tax per capita revenue					
	2012			2013		
G wins	43.751*	38.782	10.930	60.135*	27.611	-9.310
	(25.823)	(29.835)	(26.446)	(32.027)	(35.162)	(31.781)
Obs.	243	2,116	2,116	243	2,116	2,116
	Panel C: Per capita road subsidies					
	2012			2013		
G wins	4.478	0.665	2.533	2.962*	1.059	2.158
	(2.762)	(3.291)	(3.680)	(1.634)	(2.004)	(1.955)
Obs.	243	2,116	2,116	243	2,116	2,116
	Panel D: Per capita other subsidies					
	2012			2013		
G wins	32.176	46.623	32.947	38.521*	25.499	15.901
	(19.290)	(28.030)	(24.470)	(19.388)	(26.597)	(22.926)
Obs.	243	2,116	2,116	243	2,116	2,116
Controls	No	No	Yes	No	No	Yes
County FE	No	No	Yes	No	No	Yes
Specification	[-5; +5]	3rd Ord. Poly.	3rd Ord. Poly.	[-5; +5]	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *governing coalition* (G) alignment on local revenues and government transfers after the referendum in rural localities. Revenues from 2012 in columns (1)-(3) and from 2013 in columns (4)-(6). Estimates from simple linear regression in a small interval around the cutoff in columns (1)-(2). Estimates from local linear regression using the Calonico et al. (2014) robust bias-corrected confidence intervals in column (3). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (4)-(6). The controls in columns (5)-(6) and (11)-(12) include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12: Alignment and referendum turnout impact on parliament elections outcomes. RD and IV estimates.

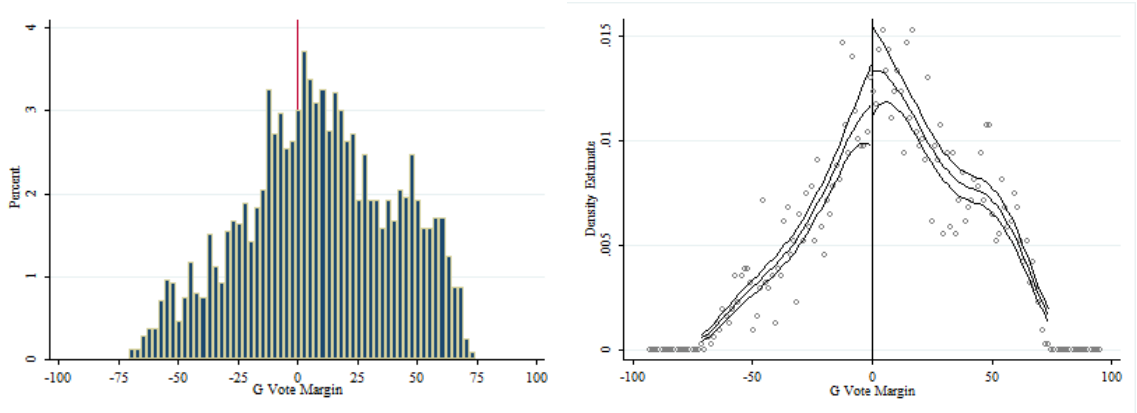
	Electoral Outcomes					
	(1)	(2)	(3)	(4)	(5)	(6)
	Reduced form RD				2SLS	
	Poly. Approx. Full sample				2SLS	
	Panel A: Referendum Turnout			Parliamentary Turnout		
G wins	0.023***	0.014	0.011			
	(0.005)	(0.010)	(0.011)			
Referendum turnout				0.308***	0.331***	0.273**
				(0.068)	(0.068)	(0.111)
Obs.	2,116	2,116	2,116	2,116	2,116	2,116
R-squared	0.474	0.475	0.475	0.544	0.542	0.546
	Panel B: Referendum Turnout			Senate G Vote Share		
G wins	0.040***	0.016	0.003			
	(0.008)	(0.012)	(0.017)			
Referendum turnout				0.579***	0.577***	0.403***
				(0.088)	(0.092)	(0.122)
Obs.	2,116	2,116	2,116	2,116	2,116	2,116
R-squared	0.483	0.485	0.486	0.544	0.545	0.565
	Panel C: Referendum Turnout			Lower Chamber G Vote Share		
G wins	0.037***	0.019	0.015			
	(0.008)	(0.012)	(0.018)			
Referendum turnout				0.535***	0.525***	0.377***
				(0.089)	(0.095)	(0.126)
Obs.	2,116	2,116	2,116	2,116	2,116	2,116
R-squared	0.480	0.482	0.482	0.549	0.550	0.560
First stage F	-	-	-	38.680	29.820	17.290
Controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Specification	1st Ord. Poly.	2nd Ord. Poly.	3rd Ord. Poly.	1st Ord. Poly.	2nd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays OLS estimates of the impact of G alignment on parliamentary outcomes (columns 1-3) and two-stage least squares estimates of the impact of referendum turnout on parliamentary elections turnout (columns 4-6) in rural localities. All estimations use polynomial approximations based on the full sample, with varying polynomial orders (first order in columns 1 and 4, second order in columns 2 and 5, third order in columns 3 and 6). All regressions include locality controls and county fixed effects. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

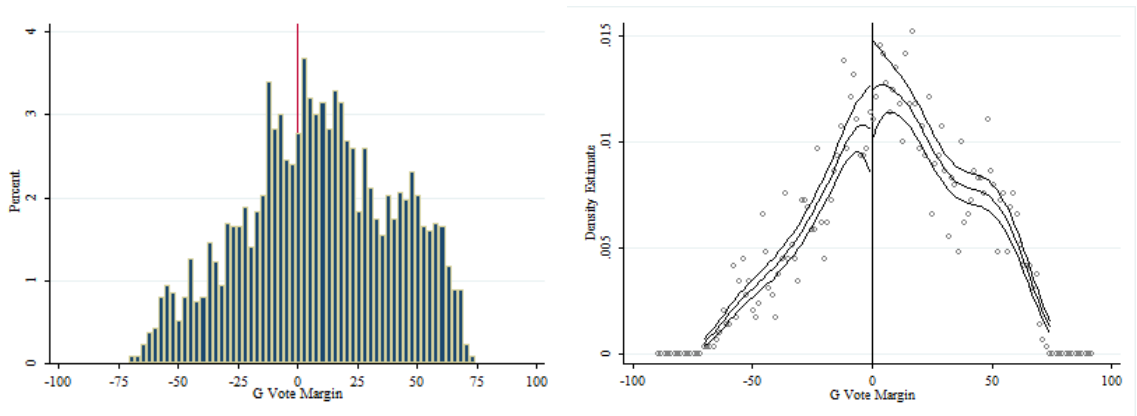
A Appendix

Figure A1 G Vote Margin Distributions and McCrary Density Tests

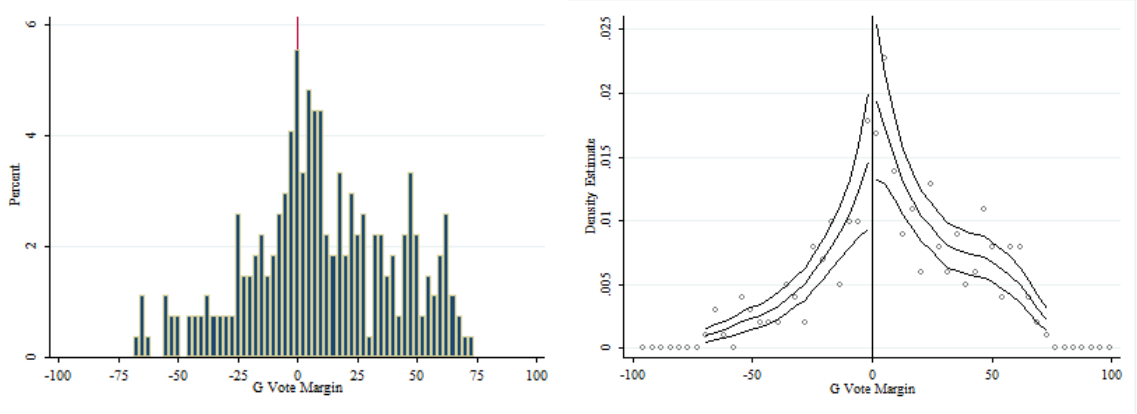
(a) All localities



(b) Rural localities



(c) Urban localities



Note: The figure displays the histograms (left) and McCrary density plots (right) for the running variable (Victory Margin) in races against G. All localities: discontinuity estimate 0.128, standard error 0.124. Rural localities: discontinuity estimate 0.177, standard error 0.146. Urban localities: discontinuity estimate 0.260, standard error 0.277.

Table A1: Descriptive statistics. Races within 5% victory margin vs. two-candidate races.

	Races against G		
	5% margin	**2-cand**	**p-value**
Panel A: Outcomes			
Turnout	0.543	0.519	0.161
Share YES	0.867	0.850	**0.003**
Turnout Parliamentary Elections	0.433	0.450	**0.102**
G Vote Share in Senate	0.612	0.612	0.993
Panel B: Covariates			
No. candidates in local elections	5.336	2.000	**0.001**
Turnout Local Elections	44.554	47.181	**0.000**
Sum % first 2 candidates	79.199	1.000	**0.009**
Log Population	8.170	7.769	**0.000**
Share adult population	0.770	0.771	0.915
Share over 65	0.206	0.206	0.955
Share males	0.496	0.497	0.390
Share high education	0.050	0.038	**0.004**
Share high school	0.163	0.143	**0.001**
Share Romanians	0.883	0.879	0.784
Unemployment rate	0.052	0.049	0.510
Per capita revenue	1,405.645	1,594.426	**0.069**
Per capita own revenue	473.531	490.976	0.582
Per capita roads funds	12.749	10.798	0.706
Per capita subsidies	130.420	101.823	0.299
Per capita expenditures	1,317.394	1,429.516	0.183
Per capita expenditures education	369.123	333.459	**0.031**
Per capita health expenditures	10.294	11.179	0.771
Per capita public expenditures	148.956	193.998	**0.078**
N max	280.000	152.000	

Note: The table displays the comparison in means in outcome variables (Panel A) and locality characteristics (Panel B) between the sample of close races (under 5 percentage points victory margin) and two-candidate races sample, for races against G. The p-values for the differences in means are reported (significant differences marked in bold).

Table A2: Validity Tests of Covariates Discontinuity. Turnout predicted by pretreatment covariates. RD estimates

	Predicted Turnout 1			**Predicted Turnout 2**			**Predicted Turnout 3**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	Poly.	Approx.	OLS	Poly.	Approx.	OLS	Poly.	Approx.
				Panel A: All localities					
G wins	-0.003	-0.018	-0.009	-0.002	-0.015	-0.006	-0.002	-0.015	-0.008
	(0.011)	(0.012)	(0.010)	(0.011)	(0.012)	(0.010)	(0.011)	(0.014)	(0.011)
Obs.	289	2,386	2,386	289	2,386	2,386	289	2,386	2,386
				Panel B: Rural localities					
G wins	-0.004	-0.015	-0.008	-0.002	-0.012	-0.004	-0.002	-0.012	-0.007
	(0.013)	(0.015)	(0.012)	(0.013)	(0.015)	(0.012)	(0.013)	(0.016)	(0.013)
Obs.	243	2,116	2,116	243	2,116	2,116	243	2,116	2,116
				Panel C: Urban localities					
G wins	-0.009	-0.026	-0.021	-0.012	-0.026	-0.024	-0.009	-0.024	-0.024
	(0.021)	(0.024)	(0.028)	(0.019)	(0.022)	(0.027)	(0.019)	(0.023)	(0.027)
Obs.	46	270	270	46	270	270	46	270	270
County FE	No	No	Yes	No	No	Yes	No	No	Yes
Specification	[-5; +5]	3rd Order Poly.	3rd Order Poly.	[-5; +5]	3rd Order Poly.	3rd Order Poly.	[-5; +5]	3rd Order Poly.	3rd Order Poly.

Notes: The table displays RD tests of discontinuity in pretreatment characteristics around the G victory threshold. Instead of testing individual characteristics, I predict referendum turnout with subsets of all covariates: in columns (1)-(3) demographic characteristics (population, age, gender, education) and labor market characteristics (unemployment rate), fiscal covariates are added in columns (4)-(6), and local elections characteristics (number of mayoral candidates and turnout at 2012 local elections) are further included in columns (7)-(9). Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3: The continuity of observable characteristics at the victory threshold.

	Rural localities		
Dependent variable			
	(1)	(2)	(3)
Log Population	-0.017	0.031	0.047
	(0.074)	(0.087)	(0.077)
Share adult population	-0.001	-0.003	0.000
	(0.006)	(0.007)	(0.006)
Share over 65	-0.006	-0.009	-0.002
	(0.008)	(0.009)	(0.008)
Share males	-0.000	-0.000	-0.001
	(0.001)	(0.002)	(0.002)
Share high education	0.006**	0.005	0.005*
	(0.003)	(0.003)	(0.003)
Share high school	0.008	0.008	0.007
	(0.006)	(0.007)	(0.006)
Share Romanians	-0.003	-0.022	-0.014
	(0.019)	(0.023)	(0.021)
Unemployment rate	-0.002	-0.004	-0.006
	(0.005)	(0.006)	(0.005)
Per capita revenue	62.148	-41.792	-97.802
	(132.769)	(131.854)	(119.491)
Per capita own revenue	37.214	32.446	22.186
	(32.706)	(35.631)	(30.034)
Per capita roads funds	-4.939	-4.708	-1.727
	(6.482)	(4.130)	(3.737)
Per capita subsidies	15.674	32.145	23.867
	(42.222)	(38.168)	(38.181)
Per capita expenditures	8.868	-102.439	-153.092
	(109.500)	(113.434)	(101.119)
Per capita expenditures education	-7.938	-17.967	-24.555
	(21.392)	(20.310)	(18.783)
Per capita health expenditures	-4.422	-6.910*	-7.269*
	(3.576)	(4.056)	(3.980)
Per capita public expenditures	0.827	2.182	-7.197
	(26.317)	(35.064)	(36.894)
No. candidates in local elections	-0.040	0.050	-0.008
	(0.216)	(0.230)	(0.228)
Turnout referendum 2007	0.776	-0.749	-1.088
	(1.392)	(1.708)	(1.416)
Sum % votes first 2 candidates	0.487	0.432	0.774
	(1.665)	(1.417)	(1.487)
County FE	No	No	Yes
Specification	[-5; +5]	3rd poly.	3rd poly.

Notes: The table displays RD estimates from regressions with covariates as the dependent variable and G wins as the main independent variable, exclusively in rural areas.

Table A4: President's party (P) alignment and referendum outcomes. RD Estimates

	Turnout						**Share YES**					
	(1) OLS	(2) OLS	(3) Local LR	(4) Poly.	(5) Approx.	(6) Full sample	(7) OLS	(8) OLS	(9) Local LR	(10) Poly.	(11) Approx.	(12) Full sample
					Panel A: All localities							
P wins	-0.022	-0.032*	-0.011	-0.019	-0.043**	-0.055***	0.010	0.005	0.010	0.003	0.001	-0.004
	(0.030)	(0.019)	(0.025)	(0.023)	(0.017)	(0.016)	(0.009)	(0.006)	(0.008)	(0.007)	(0.007)	(0.007)
Obs.	114	233	732	1,770	1,770	1,770	114	233	782	1,770	1,770	1,770
					Panel B: Rural localities							
P wins	-0.020	-0.037	-0.018	-0.021	-0.046**	-0.065***	0.011	0.005	0.010	0.005	0.003	-0.003
	(0.038)	(0.022)	(0.026)	(0.027)	(0.019)	(0.017) (0.010)		(0.006)	(0.009)	(0.008)	(0.008)	(0.008)
Obs.	98	201	704	1,585	1,585	1,585	98	201	740	1,585	1,585	1,585
					Panel C: Urban localities							
P wins	-0.013	0.004	0.054	0.003	-0.045	0.018	0.001	0.004	-0.006	-0.012	-0.013	-0.016
	(0.063)	(0.047)	(0.078)	(0.062)	(0.055)	(0.050)	(0.013)	(0.008)	(0.013)	(0.012)	(0.010)	(0.012)
Obs.	16	32	77	185	185	185	16	32	102	185	185	185
County FE	No	No	No	No	No	Yes	No	No	No	No	No	Yes
Controls	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Specification	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.	[-2.5; +2.5]	[-5; +5]	CCT Opt. h	3rd Ord. Poly.	3rd Ord. Poly.	3rd Ord. Poly.

Notes: The table displays RD estimates of the effect of *the president's party* (P) alignment on referendum turnout (columns 1-6) and the "YES" vote share (columns 7-12). Results are based on all localities (Panel A), rural (Panel B) and urban (Panel C). Estimates from simple linear regression in a small interval around the cutoff in columns (1)-(2) and (7)-(8). Estimates from local linear regression using the Calonico et al. (2014) robust bias-corrected confidence intervals in columns (3) and (9). Estimates using polynomial approximations on the full sample, without controls, with controls and with county fixed effects in columns (4)-(6) and (10)-(12). The controls in columns 5-6 and 11-12 include demographic, labor market, fiscal and electoral characteristics of the localities. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A5: The impact of local alignment and referendum turnout on parliament elections turnout. OLS estimates

	Outcomes in Parliamentary Elections		
	(1)	(2)	(3)
	Panel A: Turnout		
G wins	0.028***	-0.016***	-0.001
	(0.005)	(0.005)	(0.005)
Referendum turnout		0.365***	0.266***
		(0.028)	(0.031)
Obs.	2,116	2,116	2,116
R-squared	0.019	0.386	0.546
	Panel B: G Vote Share Senate		
G wins	0.103***	0.058***	0.051***
	(0.010)	(0.008)	(0.006)
Referendum turnout		0.374***	0.404***
		(0.059)	(0.043)
Obs.	2,116	2,116	2,116
R-squared	0.141	0.350	0.553
	Panel C: G Vote Share Lower Chamber		
G wins	0.101***	0.054***	0.050***
	(0.010)	(0.008)	(0.007)
Referendum turnout		0.392***	0.411***
		(0.058)	(0.048)
Obs.	2,116	2,116	2,116
R-squared	0.129	0.346	0.548
Controls	No	No	Yes
County FE	No	No	Yes

Notes: The table displays OLS estimates of the effect of G alignment and referendum turnout on parliamentary outcomes. Standard Errors clustered at county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Previous doctoral theses in the Department of Economics, Gothenburg

Avhandlingar publicerade innan serien Ekonomiska Studier startades
(Theses published before the series Ekonomiska Studier was started):

Östman, Hugo (1911), Norrlands ekonomiska utveckling
Moritz, Marcus (1911), Den svenska tobaksindustrien
Sundbom, I. (1933), Prisbildning och ändamålsenlighet
Gerhard, I. (1948), Problem rörande Sveriges utrikeshandel 1936/38
Hegeland, Hugo (1951), The Quantity Theory of Money
Mattsson, Bengt (1970), Cost-Benefit analys
Rosengren, Björn (1975), Valutareglering och nationell ekonomisk politik
Hjalmarsson, Lennart (1975), Studies in a Dynamic Theory of Production and its Applications
Örtendahl, Per-Anders (1975), Substitutionsaspekter på produktionsprocessen vid massaframställning
Anderson, Arne M. (1976), Produktion, kapacitet och kostnader vid ett helautomatiskt emballageglasbruk
Ohlsson, Olle (1976), Substitution och odelbarheter i produktionsprocessen vid massaframställning
Gunnarsson, Jan (1976), Produktionssystem och tätortshierarki – om sambandet mellan rumslig och ekonomisk struktur
Köstner, Evert (1976), Optimal allokering av tid mellan utbildning och arbete
Wigren, Rune (1976), Analys av regionala effektivitetsskillnader inom industribranscher
Wästlund, Jan (1976), Skattning och analys av regionala effektivitetsskillnader inom industribranscher
Flöjstad, Gunnar (1976), Studies in Distortions, Trade and Allocation Problems
Sandelin, Bo (1977), Prisutveckling och kapitalvinster på bostadsfastigheter
Dahlberg, Lars (1977), Empirical Studies in Public Planning
Lönnroth, Johan (1977), Marxism som matematisk ekonomi
Johansson, Börje (1978), Contributions to Sequential Analysis of Oligopolistic Competition

Ekonomiska Studier, utgivna av Nationalekonomiska institutionen vid Göteborgs Universitet. Nr 1 och 4 var inte doktorsavhandlingar. (The contributions to the department series ’Ekonomiska Studier’ where no. 1 and 4 were no doctoral theses):

2. **Ambjörn, Erik** (1959), Svenskt importberoende 1926-1956: en ekonomisk-statistisk kartläggning med kommentarer
3. **Landgren, K-G.** (1960), Den ”Nya ekonomien” i Sverige: J.M. Keynes, E. Wigfors och utecklingen 1927-39
5. **Bigsten, Arne** (1979), Regional Inequality and Development: A Case Study of Kenya
6. **Andersson, Lars** (1979), Statens styrning av de kommunala budgetarnas struktur (Central Government Influence on the Structure of the Municipal Budget)
7. **Gustafsson, Björn** (1979), Inkomst- och uppväxtförhållanden (Income and Family Background)

8. **Granholm, Arne** (1981), Interregional Planning Models for the Allocation of Private and Public Investments
9. **Lundborg, Per** (1982), Trade Policy and Development: Income Distributional Effects in the Less Developed Countries of the US and EEC Policies for Agricultural Commodities
10. **Juås, Birgitta** (1982), Värdering av risken för personskador. En jämförande studie av implicita och explicita värden. (Valuation of Personal Injuries. A comparison of Explicit and Implicit Values)
11. **Bergendahl, Per-Anders** (1982), Energi och ekonomi - tillämpningar av input-output analys (Energy and the Economy - Applications of Input-Output Analysis)
12. **Blomström, Magnus** (1983), Foreign Investment, Technical Efficiency and Structural Change - Evidence from the Mexican Manufacturing Industry
13. **Larsson, Lars-Göran** (1983), Comparative Statics on the Basis of Optimization Methods
14. **Persson, Håkan** (1983), Theory and Applications of Multisectoral Growth Models
15. **Sterner, Thomas** (1986), Energy Use in Mexican Industry.
16. **Flood, Lennart** (1986), On the Application of Time Use and Expenditure Allocation Models.
17. **Schuller, Bernd-Joachim** (1986), Ekonomi och kriminalitet - en empirisk undersökning av brottsligheten i Sverige (Economics of crime - an empirical analysis of crime in Sweden)
18. **Walfridson, Bo** (1987), Dynamic Models of Factor Demand. An Application to Swedish Industry.
19. **Stålhammar, Nils-Olov** (1987), Strukturomvandling, företagsbeteende och förväntningsbildning inom den svenska tillverkningsindustrin (Structural Change, Firm Behaviour and Expectation Formation in Swedish Manufactury)
20. **Anxo, Dominique** (1988), Sysselsättningseffekter av en allmän arbetstidsför-kortning (Employment effects of a general shortage of the working time)
21. **Mbelle, Ammon** (1988), Foreign Exchange and Industrial Development: A Study of Tanzania.
22. **Ongaro, Wilfred** (1988), Adoption of New Farming Technology: A Case Study of Maize Production in Western Kenya.
23. **Zejan, Mario** (1988), Studies in the Behavior of Swedish Multinationals.
24. **Görling, Anders** (1988), Ekonomisk tillväxt och miljö. Förorenings-struktur och ekonomiska effekter av olika miljövårdsprogram. (Economic Growth and Environment. Pollution Structure and Economic Effects of Some Environmental Programs).
25. **Aguilar, Renato** (1988), Efficiency in Production: Theory and an Application on Kenyan Smallholders.
26. **Kayizzi-Mugerwa, Steve** (1988), External Shocks and Adjustment in Zambia.

27. **Bornmalm-Jardelöw, Gunilla** (1988), Högre utbildning och arbetsmarknad (Higher Education and the Labour Market)
28. **Tansini, Ruben** (1989), Technology Transfer: Dairy Industries in Sweden and Uruguay.
29. **Andersson, Irene** (1989), Familjebeskattning, konsumtion och arbetsutbud - En ekonometrisk analys av löne- och inkomstelasticiteter samt policysimuleringar för svenska hushåll (Family Taxation, Consumption and Labour Supply - An Econometric Analysis of Wage and Income Elasticities and Policy Simulations for Swedish Households)
30. **Henrekson, Magnus** (1990), An Economic Analysis of Swedish Government Expenditure
31. **Sjöö, Boo** (1990), Monetary Policy in a Continuous Time Dynamic Model for Sweden
32. **Rosén, Åsa** (1991), Contributions to the Theory of Labour Contracts.
33. **Loureiro, Joao M. de Matos** (1992), Foreign Exchange Intervention, Sterilization and Credibility in the EMS: An Empirical Study
34. **Irandoust, Manuchehr** (1993), Essays on the Behavior and Performance of the Car Industry
35. **Tasiran, Ali Cevat** (1993), Wage and Income Effects on the Timing and Spacing of Births in Sweden and the United States
36. **Milopoulos, Christos** (1993), Investment Behaviour under Uncertainty: An Econometric Analysis of Swedish Panel Data
37. **Andersson, Per-Åke** (1993), Labour Market Structure in a Controlled Economy: The Case of Zambia
38. **Storrie, Donald W.** (1993), The Anatomy of a Large Swedish Plant Closure
39. **Semboja, Haji Hatibu Haji** (1993), Energy and Development in Kenya
40. **Makonnen, Negatu** (1993), Labor Supply and the Distribution of Economic Well-Being: A Case Study of Lesotho
41. **Julin, Eva** (1993), Structural Change in Rural Kenya
42. **Durevall, Dick** (1993), Essays on Chronic Inflation: The Brazilian Experience
43. **Veiderpass, Ann** (1993), Swedish Retail Electricity Distribution: A Non-Parametric Approach to Efficiency and Productivity Change
44. **Odeck, James** (1993), Measuring Productivity Growth and Efficiency with Data Envelopment Analysis: An Application on the Norwegian Road Sector
45. **Mwenda, Abraham** (1993), Credit Rationing and Investment Behaviour under Market Imperfections: Evidence from Commercial Agriculture in Zambia
46. **Mlambo, Kupukile** (1993), Total Factor Productivity Growth: An Empirical Analysis of Zimbabwe's Manufacturing Sector Based on Factor Demand Modelling
47. **Ndung'u, Njuguna** (1993), Dynamics of the Inflationary Process in Kenya
48. **Modén, Karl-Markus** (1993), Tax Incentives of Corporate Mergers and Foreign Direct Investments
49. **Franzén, Mikael** (1994), Gasoline Demand - A Comparison of Models
50. **Heshmati, Almas** (1994), Estimating Technical Efficiency, Productivity Growth And Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture

51. **Salas, Osvaldo** (1994), Efficiency and Productivity Change: A Micro Data Case Study of the Colombian Cement Industry
52. **Bjurek, Hans** (1994), Essays on Efficiency and Productivity Change with Applications to Public Service Production
53. **Cabezas Vega, Luis** (1994), Factor Substitution, Capacity Utilization and Total Factor Productivity Growth in the Peruvian Manufacturing Industry
54. **Katz, Katarina** (1994), Gender Differentiation and Discrimination. A Study of Soviet Wages
55. **Asal, Maher** (1995), Real Exchange Rate Determination and the Adjustment Process: An Empirical Study in the Cases of Sweden and Egypt
56. **Kjulin, Urban** (1995), Economic Perspectives on Child Care
57. **Andersson, Göran** (1995), Volatility Forecasting and Efficiency of the Swedish Call Options Market
58. **Forteza, Alvaro** (1996), Credibility, Inflation and Incentive Distortions in the Welfare State
59. **Locking, Håkan** (1996), Essays on Swedish Wage Formation
60. **Välilä, Timo** (1996), Essays on the Credibility of Central Bank Independence
61. **Yilma, Mulugeta** (1996), Measuring Smallholder Efficiency: Ugandan Coffee and Food-Crop Production
62. **Mabugu, Ramos E.** (1996), Tax Policy Analysis in Zimbabwe Applying General Equilibrium Models
63. **Johansson, Olof** (1996), Welfare, Externalities, and Taxation; Theory and Some Road Transport Applications.
64. **Chitiga, Margaret** (1996), Computable General Equilibrium Analysis of Income Distribution Policies in Zimbabwe
65. **Leander, Per** (1996), Foreign Exchange Market Behavior Expectations and Chaos
66. **Hansen, Jörgen** (1997), Essays on Earnings and Labor Supply
67. **Cotfas, Mihai** (1997), Essays on Productivity and Efficiency in the Romanian Cement Industry
68. **Horgby, Per-Johan** (1997), Essays on Sharing, Management and Evaluation of Health Risks
69. **Nafar, Nosratollah** (1997), Efficiency and Productivity in Iranian Manufacturing Industries
70. **Zheng, Jinghai** (1997), Essays on Industrial Structure, Technical Change, Employment Adjustment, and Technical Efficiency
71. **Isaksson, Anders** (1997), Essays on Financial Liberalisation in Developing Countries: Capital mobility, price stability, and savings
72. **Gerdin, Anders** (1997), On Productivity and Growth in Kenya, 1964-94
73. **Sharifi, Alimorad** (1998), The Electricity Supply Industry in Iran: Organization, performance and future development
74. **Zamanian, Max** (1997), Methods for Mutual Fund Portfolio Evaluation: An application to the Swedish market
75. **Manda, Damiano Kulundu** (1997), Labour Supply, Returns to Education, and the Effect of Firm Size on Wages: The case of Kenya
76. **Holmén, Martin** (1998), Essays on Corporate Acquisitions and Stock Market Introductions

77. **Pan, Kelvin** (1998), Essays on Enforcement in Money and Banking
78. **Rogat, Jorge** (1998), The Value of Improved Air Quality in Santiago de Chile
79. **Peterson, Stefan** (1998), Essays on Large Shareholders and Corporate Control
80. **Belhaj, Mohammed** (1998), Energy, Transportation and Urban Environment in Africa: The Case of Rabat-Salé, Morocco
81. **Mekonnen, Alemu** (1998), Rural Energy and Afforestation: Case Studies from Ethiopia
82. **Johansson, Anders** (1998), Empirical Essays on Financial and Real Investment Behavior
83. **Köhlin, Gunnar** (1998), The Value of Social Forestry in Orissa, India
84. **Levin, Jörgen** (1998), Structural Adjustment and Poverty: The Case of Kenya
85. **Ncube, Mkhululi** (1998), Analysis of Employment Behaviour in Zimbabwe
86. **Mwansa, Ladslous** (1998), Determinants of Inflation in Zambia
87. **Agnarsson, Sveinn** (1998), Of Men and Machines: Essays in Applied Labour and Production Economics
88. **Kadenge, Phineas** (1998), Essays on Macroeconomic Adjustment in Zimbabwe: Inflation, Money Demand, and the Real Exchange Rate
89. **Nyman, Håkan** (1998), An Economic Analysis of Lone Motherhood in Sweden
90. **Carlsson, Fredrik** (1999), Essays on Externalities and Transport
91. **Johansson, Mats** (1999), Empirical Studies of Income Distribution
92. **Alemu, Tekie** (1999), Land Tenure and Soil Conservation: Evidence from Ethiopia
93. **Lundvall, Karl** (1999), Essays on Manufacturing Production in a Developing Economy: Kenya 1992-94
94. **Zhang, Jianhua** (1999), Essays on Emerging Market Finance
95. **Mlima, Aziz Ponary** (1999), Four Essays on Efficiency and Productivity in Swedish Banking
96. **Davidsen, Björn-Ivar** (2000), Bidrag til den økonomisk-metodologiske tenkningen (Contributions to the Economic Methodological Thinking)
97. **Ericson, Peter** (2000), Essays on Labor Supply
98. **Söderbom, Måns** (2000), Investment in African Manufacturing: A Microeconomic Analysis
99. **Höglund, Lena** (2000), Essays on Environmental Regulation with Applications to Sweden
100. **Olsson, Ola** (2000), Perspectives on Knowledge and Growth
101. **Meuller, Lars** (2000), Essays on Money and Credit
102. **Österberg, Torun** (2000), Economic Perspectives on Immigrants and Intergenerational Transmissions
103. **Kalinda Mkenda, Beatrice** (2001), Essays on Purchasing Power Parity, RealExchange Rate, and Optimum Currency Areas
104. **Nerhagen, Lena** (2001), Travel Demand and Value of Time - Towards an Understanding of Individuals Choice Behavior

105. **Mkenda, Adolf** (2001), Fishery Resources and Welfare in Rural Zanzibar
106. **Eggert, Håkan** (2001), Essays on Fisheries Economics
107. **Andrén, Daniela** (2001), Work, Sickness, Earnings, and Early Exits from the Labor Market. An Empirical Analysis Using Swedish Longitudinal Data
108. **Nivorozhkin, Eugene** (2001), Essays on Capital Structure
109. **Hammar, Henrik** (2001), Essays on Policy Instruments: Applications to Smoking and the Environment
110. **Nannyonjo, Justine** (2002), Financial Sector Reforms in Uganda (1990-2000): Interest Rate Spreads, Market Structure, Bank Performance and Monetary Policy
111. **Wu, Hong** (2002), Essays on Insurance Economics
112. **Linde-Rahr, Martin** (2002), Household Economics of Agriculture and Forestry in Rural Vienam
113. **Maneschiöld, Per-Ola** (2002), Essays on Exchange Rates and Central Bank Credibility
114. **Andrén, Thomas** (2002), Essays on Training, Welfare and Labor Supply
115. **Granér, Mats** (2002), Essays on Trade and Productivity: Case Studies of Manufacturing in Chile and Kenya
116. **Jaldell, Henrik** (2002), Essays on the Performance of Fire and Rescue Services
117. **Alpizar, Francisco, R.** (2002), Essays on Environmental Policy-Making in Developing Countries: Applications to Costa Rica
118. **Wahlberg, Roger** (2002), Essays on Discrimination, Welfare and Labor Supply
119. **Piculescu, Violeta** (2002), Studies on the Post-Communist Transition
120. **Pylkkänen, Elina** (2003), Studies on Household Labor Supply and Home Production
121. **Löfgren, Åsa** (2003), Environmental Taxation – Empirical and Theoretical Applications
122. **Ivaschenko, Oleksiy** (2003), Essays on Poverty, Income Inequality and Health in Transition Economies
123. **Lundström, Susanna** (2003), On Institutions, Economic Growth and the Environment
124. **Wambugu, Anthony** (2003), Essays on Earnings and Human Capital in Kenya
125. **Adler, Johan** (2003), Aspects of Macroeconomic Saving
126. **Erlandsson, Mattias** (2003), On Monetary Integration and Macroeconomic Policy
127. **Brink, Anna** (2003), On the Political Economy of Municipality Break-Ups
128. **Ljungwall, Christer** (2003), Essays on China's Economic Performance During the Reform Period
129. **Chifamba, Ronald** (2003), Analysis of Mining Investments in Zimbabwe
130. **Muchapondwa, Edwin** (2003), The Economics of Community-Based Wildlife Conservation in Zimbabwe
131. **Hammes, Klaus** (2003), Essays on Capital Structure and Trade Financing
132. **Abou-Ali, Hala** (2003), Water and Health in Egypt: An Empirical Analysis
133. **Simatele, Munacinga** (2004), Financial Sector Reforms and Monetary Policy in Zambia
134. **Tezic, Kerem** (2004), Essays on Immigrants' Economic Integration
135. INSTÄLLD

136. **Gjirja, Matilda** (2004), Efficiency and Productivity in Swedish Banking
137. **Andersson, Jessica** (2004), Welfare Environment and Tourism in Developing Countries
138. **Chen, Yinghong** (2004), Essays on Voting Power, Corporate Governance and Capital Structure
139. **Yesuf, Mahmud** (2004), Risk, Time and Land Management under Market Imperfections: Applications to Ethiopia
140. **Kateregga, Eseza** (2005), Essays on the Infestation of Lake Victoria by the Water Hyacinth
141. **Edvardsen, Dag Fjeld** (2004), Four Essays on the Measurement of Productive Efficiency
142. **Lidén, Erik** (2005), Essays on Information and Conflicts of Interest in Stock Recommendations
143. **Dieden, Sten** (2005), Income Generation in the African and Coloured Population – Three Essays on the Origins of Household Incomes in South Africa
144. **Eliasson, Marcus** (2005), Individual and Family Consequences of Involuntary Job Loss
145. **Mahmud, Minhaj** (2005), Measuring Trust and the Value of Statistical Lives: Evidence from Bangladesh
146. **Lokina, Razack Bakari** (2005), Efficiency, Risk and Regulation Compliance: Applications to Lake Victoria Fisheries in Tanzania
147. **Jussila Hammes, Johanna** (2005), Essays on the Political Economy of Land Use Change
148. **Nyangena, Wilfred** (2006), Essays on Soil Conservation, Social Capital and Technology Adoption
149. **Nivorozhkin, Anton** (2006), Essays on Unemployment Duration and Programme Evaluation
150. **Sandén, Klas** (2006), Essays on the Skill Premium
151. **Deng, Daniel** (2006), Three Essays on Electricity Spot and Financial Derivative Prices at the Nordic Power Exchange
152. **Gebreeyesus, Mulu** (2006), Essays on Firm Turnover, Growth, and Investment Behavior in Ethiopian Manufacturing
153. **Islam, Nizamul Md.** (2006), Essays on Labor Supply and Poverty: A Microeconometric Application
154. **Kjaer, Mats** (2006), Pricing of Some Path-Dependent Options on Equities and Commodities
155. **Shimeles, Abebe** (2006), Essays on Poverty, Risk and Consumption Dynamics in Ethiopia
156. **Larsson, Jan** (2006), Four Essays on Technology, Productivity and Environment
157. **Congdon Fors, Heather** (2006), Essays in Institutional and Development Economics
158. **Akpalu, Wisdom** (2006), Essays on Economics of Natural Resource Management and Experiments
159. **Daruvala, Dinky** (2006), Experimental Studies on Risk, Inequality and Relative Standing
160. **García, Jorge** (2007), Essays on Asymmetric Information and Environmental Regulation through Disclosure

161. **Bezabih, Mintewab** (2007), Essays on Land Lease Markets, Productivity, Biodiversity, and Environmental Variability
162. **Visser, Martine** (2007), Fairness, Reciprocity and Inequality: Experimental Evidence from South Africa
163. **Holm, Louise** (2007), A Non-Stationary Perspective on the European and Swedish Business Cycle
164. **Herbertsson, Alexander** (2007), Pricing Portfolio Credit Derivatives
165. **Johansson, Anders C.** (2007), Essays in Empirical Finance: Volatility, Interdependencies, and Risk in Emerging Markets
166. **Ibáñez Díaz, Marcela** (2007), Social Dilemmas: The Role of Incentives, Norms and Institutions
167. **Ekbom, Anders** (2007), Economic Analysis of Soil Capital, Land Use and Agricultural Production in Kenya
168. **Sjöberg, Pål** (2007), Essays on Performance and Growth in Swedish Banking
169. **Palma Aguirre, Grisha Alexis** (2008), Explaining Earnings and Income Inequality in Chile
170. **Akay, Alpaslan** (2008), Essays on Microeconometrics and Immigrant Assimilation
171. **Carlsson, Evert** (2008), After Work – Investing for Retirement
172. **Munshi, Farzana** (2008), Essays on Globalization and Occupational Wages
173. **Tsakas, Elias** (2008), Essays on Epistemology and Evolutionary Game Theory
174. **Erlandzon, Karl** (2008), Retirement Planning: Portfolio Choice for Long-Term Investors
175. **Lampi, Elina** (2008), Individual Preferences, Choices, and Risk Perceptions – Survey Based Evidence
176. **Mitrut, Andreea** (2008), Four Essays on Interhousehold Transfers and Institutions in Post-Communist Romania
177. **Hansson, Gustav** (2008), Essays on Social Distance, Institutions, and Economic Growth
178. **Zikhali, Precious** (2008), Land Reform, Trust and Natural Resource Management in Africa
179. **Tengstam, Sven** (2008), Essays on Smallholder Diversification, Industry Location, Debt Relief, and Disability and Utility
180. **Boman, Anders** (2009), Geographic Labour Mobility – Causes and Consequences
181. **Qin, Ping** (2009), Risk, Relative Standing and Property Rights: Rural Household Decision-Making in China
182. **Wei, Jiegen** (2009), Essays in Climate Change and Forest Management
183. **Belu, Constantin** (2009), Essays on Efficiency Measurement and Corporate Social Responsibility
184. **Ahlerup, Pelle** (2009), Essays on Conflict, Institutions, and Ethnic Diversity
185. **Quiroga, Miguel** (2009), Microeconomic Policy for Development: Essays on Trade and Environment, Poverty and Education
186. **Zerfu, Daniel** (2010), Essays on Institutions and Economic Outcomes
187. **Wollbrant, Conny** (2010), Self-Control and Altruism
188. **Villegas Palacio, Clara** (2010), Formal and Informal Regulations: Enforcement and Compliance

189. **Maican, Florin** (2010), Essays in Industry Dynamics on Imperfectly Competitive Markets
190. **Jakobsson, Niklas** (2010), Laws, Attitudes and Public Policy
191. **Manescu, Cristiana** (2010), Economic Implications of Corporate Social Responsibility and Responsible Investments
192. **He, Haoran** (2010), Environmental and Behavioral Economics – Applications to China
193. **Andersson, Fredrik W.** (2011), Essays on Social Comparison
194. **Isaksson, Ann-Sofie** (2011), Essays on Institutions, Inequality and Development
195. **Pham, Khanh Nam** (2011), Prosocial Behavior, Social Interaction and Development: Experimental Evidence from Vietnam
196. **Lindskog, Annika** (2011), Essays on Economic Behaviour: HIV/AIDS, Schooling, and Inequality
197. **Kotsadam, Andreas** (2011), Gender, Work, and Attitudes
198. **Alem, Yonas** (2011), Essays on Shocks, Welfare, and Poverty Dynamics: Microeconometric Evidence from Ethiopia
199. **Köksal-Ayhan, Miyase Yesim** (2011), Parallel Trade, Reference Pricing and Competition in the Pharmaceutical Market: Theory and Evidence
200. **Vondolia, Godwin Kofi** (2011), Essays on Natural Resource Economics
201. **Widerberg, Anna** (2011), Essays on Energy and Climate Policy – Green Certificates, Emissions Trading and Electricity Prices
202. **Siba, Eyerusalem** (2011), Essays on Industrial Development and Political Economy of Africa
203. **Orth, Matilda** (2012), Entry, Competition and Productivity in Retail
204. **Nerman, Måns** (2012), Essays on Development: Household Income, Education, and Female Participation and Representation
205. **Wicks, Rick** (2012), The Place of Conventional Economics in a World with Communities and Social Goods
206. **Sato, Yoshihiro** (2012), Dynamic Investment Models, Employment Generation and Productivity – Evidence from Swedish Data
207. **Valsecchi, Michele** (2012), Essays in Political Economy of Development
208. **Teklewold Belayneh, Hailemariam** (2012), Essays on the Economics of Sustainable Agricultural Technologies in Ethiopia
209. **Wagura Ndiritu, Simon** (2013), Essays on Gender Issues, Food Security, and Technology Adoption in East Africa
210. **Ruist, Joakim** (2013), Immigration, Work, and Welfare
211. **Nordén, Anna** (2013), Essays on Behavioral Economics and Policies for Provision of Ecosystem Services
212. **Yang, Xiaojun** (2013), Household Decision Making, Time Preferences, and Positional Concern: Experimental Evidence from Rural China
213. **Bonilla Londoño, Jorge Alexander** (2013), Essays on the Economics of Air Quality Control
214. **Mohlin, Kristina** (2013), Essays on Environmental Taxation and Climate Policy
215. **Medhin, Haileselassie** (2013), The Poor and Their Neighbors: Essays on Behavioral and Experimental Economics

216. **Andersson, Lisa** (2013), Essays on Development and Experimental Economics: Migration, Discrimination and Positional Concerns
217. **Weng, Qian** (2014), Essays on Team Cooperation and Firm Performance
218. **Zhang, Xiao-Bing** (2015), Cooperation and Paradoxes in Climate Economics
219. **Jaime Torres, Monica Marcela** (2015) Essays on Behavioral Economics and Policy Design
220. **Bejenariu, Simona** (2015) Determinants of Health Capital at Birth: Evidence from Policy Interventions
221. **Nguyen, Van Diem** (2015) Essays on Takeovers and Executive Compensation
222. **Tolonen, Anja** (2015) Mining Booms in Africa and Local Welfare Effects: Labor Markets, Women's Empowerment and Criminality
223. **Hassen, Sied** (2015) On the Adoption and Dis-adoption of Household Energy and Farm Technologies
224. **Moursli, Mohamed-Reda** (2015) Corporate Governance and the Design of Board of Directors
225. **Borcan, Oana** (2015) Economic Determinants and Consequences of Political Institutions