ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
UNIVERSITY OF GOTHENBURG
248

Essays on Public Policy in the Informal Sector Context

Carolin Sjöholm

UNIVERSITY OF
GOTHENBURG

ISBN 978-91-88199-55-3 (printed)
ISBN 978-91-88199-56-0 (pdf)
ISSN 1651-4289 (printed)
ISSN 1651-4297 (online)

Printed in Sweden,
Gothenburg University 2021

Till Henry och Ingrid

1 Acknowledgments

This thesis would not have been possible without the many wonderful people who have helped me throughout this process.

First, I want to thank my supervisors Måns Söderbom and Annika Lindskog. Your knowledge, encouragement and insightful comments have guided me through the long process of writing this thesis. Måns, your expertise and intuitive feeling for economics and econometrics has been a true inspiration. Thank you for being my mentor and for challenging me to become a better researcher. Annika, you have an eye for detail that makes you a great researcher. Thank you for always asking the hard questions and for the countless hours your have dedicated to my work. Your efforts have challenged me to push this work further. Thank you both for everything you have taught me. It has been a pleasure working with you both. Finally, I would be remiss if I did not also thank you for being understanding. Life has its ups and downs, regardless of the demands of a PhD. During this process, I have enjoyed both the best and the worst moments in my personal life. Without your patience, guidance and support I would not have been able to finish.

My PhD journey would not have been the same without all my wonderful colleagues and cohort friends. Thank you Josephine, Lisa, Hanna, Martin, Laura, Mikael, Tensay, Vivi, Verena, Simon, Andrea, and Yashoda, for all the late night dinners in L1 (that seems like a lifetime ago), and all the support and good times we shared. I would not have made it through those first years without all of you. I would also like to thank the "younger" phd cohorts for including me in their community when I returned from parental leave. Thank you Melissa, Simon, Tewodros, Eyoual, Sebastian, Louise, Tamś, Maks, and Lina for always being supportive.

Lisa, thank you for all the morning and afternoon coffee breaks, and our discussions about life and econometrics. Our friendship began in L1, but it will be a life-long journey. Josephine, you have become like a sister to me. Thank you for teaching me about policy, for laughing and crying with me, and for spoiling Henry and Ingrid with the love of an aunt. You have a special place in our hearts. Anja, you have encouraged me to believe in myself, to stop whining, and pushed me to continually try harder. Thank you for the comradery of shared research and conference trips, and for inviting me to Barnard. Most of all, thank you for being my friend.

I am greatly indebted to all my colleagues at the Department of Economics for providing a great academic and collegial environment. It has been a pleasure to work with each of you. Thank you Joe Vecci for being a good friend and for always taking the time to listen to my concerns. I truly appreciated your willingness to help me to solve whatever obstacles I confronted. A heartfelt thank you to Eva Ranehill for sharing her job market experiences with me, but more so for the understanding and offerings of help during moments when combining academic work and motherhood was challenging. It meant so much. A special thank you to Anna Bindler and Gustav Kjellson for their insightful comments during my final seminars, I am very grateful. Gustav, I especially enjoyed our interesting discussions on adverse selection and health economics. Thank you also to the development group for all the invaluable seminar discussions. I am indebted to Elizabeth Földi, Mona Jönefors, Ann-Christin Räätäri Nyström, and Maria Siirak, for all the administrative support they provided in helping me sort out my endless questions related to parental leave. Emma, thank you for supporting my teaching assistant duties while I was in the US and for all the conversations in your office. Selma, thank you for always keeping me updated on the latest Netflix series, and Carmen, thank you for letting us stay in your home.

My research on Rwandan health policy would not have been possible without the help and support from Andinet Woldemichael at the African Development Bank. I am grateful for all your good advice and insightful guidance related to public policy and econometrics. Thank you for being a mentor and a friend. It has been my pleasure to work with you, I hope we will have the opportunity to meet soon.

Empirical research would be impossible without the knowledge and expertise of local institutions and policy makers. I am indebted to everyone at the Ministry of Health in Rwanda. Andrew Muhire, Dr. Pascal Kayobotsi, and Emmanuel Ntawuyirusha, thank you for your wonderful reception at the Ministry of Health in Kigali and for sharing you expert knowledge on the HMIS and PBF data. I am humbled by your generosity and your tireless dedication to developing the Rwandan health sector.

The Department of Social Sciences and the Collage of Business at Michigan Technological University have been my home away from home during my PhD. To all of you at Michigan Tech, thank you for the welcoming atmosphere, great research conversations and for all the cross-country skiing. Michigan is my second home and you have helped me make each trip there a pleasure.

A PhD is about so much more than hard work and thesis writing. Linda, thank you for the weekends of adventure with the kids, they were wonderful moments that gave me the energy to continue. Emma, our Åhus walks have meant more to me than you can know. Amanda, I am grateful for your unconditional friendship, for always looking after me and for offering your help when I needed it the most. We have come a long way since our days at Handels working on our master's thesis. You inspire me every day.

Above all, I owe my deepest debt of gratitude to my family. To my mother, for her endless love and encouragement. Sebastian and Rebecca, Ann-Kristin, Johanna, and now Anna, your love, patience and unconditional support have meant everything to me. Thank you all for helping me to achieve my goal and for always having my back. Julie, Amy, Nash, and Pax, you have welcomed me into your family with open arms and I thank you. Dad, you inspired me to start this PhD, to always do my best and to be proud of myself. I know that you would have been so proud of this thesis. I miss you every day.

Finally, I would like to thank my wonderful husband Sam. You have been my biggest supporter during this journey. Thank you for helping to carry my load, for always listening, encouraging me, and challenging me to see the bigger picture. You are the love of my life and I owe you everything.

Henry and Ingrid, I love you more than words can express. This thesis is for you.

Gothenburg, May 2021
Carolin Sjöholm

Contents

Introduction

A majority of all workers in the world are informally employed. Approximately 2 billion workers, or 60% of the world's employed population, earn their livelihoods in the informal sector. Workers in the informal sector often face higher risk of poverty and lower productivity compared to formal workers (International Labour Office, 2018). One reason is that they often lack access to social protection, which makes them vulnerable to adverse shocks such as sickness or income loss. Social protection refers to policies and programs designed to reduce and prevent poverty and vulnerability throughout the life cycle and often include a mix of social insurance and social assistance programs (International Labour Office, 2017). In the absence of these safety nets, adverse shocks risk pushing households deeper into poverty or maintain them in a poverty trap.

During the last decades, large efforts have been made in many developing countries to expand social protection to the informal sector and achieve universal coverage of such programs. Despite wide agreement regarding the importance of social protection as a key factor for inclusive growth, this human right is still not fulfilled for most people in the world. The 1948 Universal Declaration of Human Rights recognizes that everyone has the right to social security (Article 22). Although social protection policies are seen as key elements in national development strategies in most countries, it has been estimated that approximately 4 billion people, almost 55% of the world's population, have access to no or inadequate social protection (International Labour Office, 2017). A majority of this group is represented by households in the informal sector.

Households in the informal sector face substantial idiosyncratic and common risk, resulting in high income variability (Townsend, 1994). For a large share of these households who live on a day to day basis, adverse shocks such as health and employment shocks, could throw families into poverty and have long lasting effects for generations. For example, in 2015 approximately 930 million people incurred catastrophic health expenditures, defined as out-of-pocket health spending exceeding 10% of household consumption, pushing 89.7 million people into extreme poverty (World Health Organization & World Bank, 2019). As a result, informal labor is increasingly recognized as an obstacle to eradicate poverty in many developing countries and a major challenge for achieving the Sustainable Development Goals.

In the absence of social protection, the burden to protect households from idiosyncratic and common shocks is placed on the families and communities themselves. To deal with the effects of adverse shocks, households rely on strategies such as informal borrowing, asset sale and decreased education expenditures (Leive & Xu, 2008; Mitra et al., 2016; Heltberg & Lund, 2009; Islam & Maitra, 2012). Additionally, uninsured risk compels households to diversify income and to engage in low-risk and low-return production activities (Cole et al., 2017; Dercon & Christiaensen, 2011) in order to smooth consumption. These activities hamper the ability of households to grow their incomes and escape poverty (Binswanger & Rosenzweig, 1993). As a result, households are kept in persistent poverty. Despite informal insurance arrangements and strategies income fluctuations often remain high, suggesting that informal income-smoothing mechanisms are inadequate and leave households with uninsured risk (Townsend, 1995; P. Gertler & Gruber, 2002).

Social protection can reduce the cost of coping strategies and enhance the capacity of families and communities to absorb the negative impacts of shocks. For example, cash transfers have been shown to have positive and sustained effects on child education and health (Aizer et al., 2016; Baird et al., 2019), household investment in durable goods and savings (Haushofer & Shapiro, 2016), and productive investments (Handa et al., 2018; Bastagli et al., 2016; P. J. Gertler et al., 2012). Furthermore, in the presence of idiosyncratic shocks, households with health insurance were more likely to invest in schooling for girls, livestock and durable goods compared to uninsured households (Liu, 2016). There is an international consensus on the importance of social protection as a key policy tool for implementing the U.N. Sustainable Development Goals and to ensure inclusive development where no one is left behind. Social protection is essential to achieving a number of the SGDs such as eradicating poverty for all everywhere (SDG 1), ending hunger (SDG 2), and contributing to gender equity and women's empowerment (SDG 5). Furthermore, by increasing access to affordable healthcare, social protection can contribute to achieving universal health care (Target 3.8) and good health and well-being for all (SDG 3). As a result, Target 1.3 explicitly calls on countries to implement nationally appropriate social protection systems to end poverty by 2030 (United Nations, 2015).

Despite a global agreement on the importance of social protection, questions regarding how to best implement and expand effective and sustainable universal programs are still unanswered. Countries often combine contributory social insurance schemes with non-contributory social assistance programs in order to achieve a universal coverage. On the one hand, non-contributory programs include universal and means-tested social assistance programs that are key to ensuring a basic level of social protection for all residents, i.e. a social protection floor (Behrendt & Nguyen, 2018). While universal programs are effective in reaching the poorest and most vulnerable households, they also cover many households that are not in need of social protection. In the face of limited fiscal capacities, programs targeted to the poor might offer a more cost-effective option. However, means-tested programs rely on costly mechanisms to identify the poorest households (Aryeetey et al., 2012) with low levels of accuracy. This often leads to under-coverage and errors of exclusion (Brown et al., 2016), resulting in a trade-off between coverage and effectiveness.

On the other hand, contributory social insurance schemes tend to provide more insurance coverage and a higher level of protection than social assistance programs. However, social insurance schemes might be inaccessible for the poorest households that often lack contributory capacity (Behrendt & Nguyen, 2018). In order to make enrollment equitable, governments can subsidize enrollment. However, subsidies have the potential of being regressive if contributions remain too high for the most vulnerable households, preventing them from enrollment despite government subsidies (Kalisa et al., 2016). Additionally, take-up of the social protection programs might be hindered by factors such as lack of information (Hossain, 2011), high transaction costs (Capuno et al., 2016), and low quality of services. These barriers must be defined and targeted by well designed policies and interventions.

Ultimately, the potential capacity of social protection programs to address risk and vulnerability, by contributing to increased productivity and resilience among households in the informal sector, represents another important factor that is likely to predict take-up of the program and willingness to contribute to enrollment. This is largely contextual. If benefits are not aligned with the need and priorities of households, they may be reluctant to contribute. The design of efficient policies is likely to be particularly challenging for the informal sector that represents a complex and all but homogeneous sector of the labor force in most developing country contexts. Increased knowledge regarding the impacts of social protection programs can improve the predictability and the efficiency of public policy.

Summary Thesis

My dissertation consists of three independent empirical papers on public policy in the informal sector context. The aim of the thesis is to contribute to increase knowledge regarding the efficiency of common policy tools in contributing to universal coverage of social protection programs and to increase resilience among households in the informal sector. To do this, I use quasi-experimental evaluation methods in combination with detailed administrative data to investigate the impact of policies on the take-up and quality of social insurance programs in the informal sector, as well as their impact on the productivity of economic activities. The overall contribution of my thesis is to provide evidence related to the interaction between public policy and household decision making, focusing primarily on economic performance and health. On the one hand, the design and implementation of policy interventions will determine the effects of such policies on the lives and livelihoods of households in the informal sector. On the other hand, household preferences and decision making will determine the success of implementation and take-up of these policies. In the first chapter I evaluate the importance of premium subsidies as a policy tool to achieve universal coverage of community-based health insurance (CBHI) in Rwanda, while considering the financial self-sustainability of the insurance scheme. In the second paper I explore the impact of a national social protection program in Mexico that addresses the burden of unpaid housework on women, when evaluating its impact on female entrepreneurship.Finally, in the third paper I investigate disparities in the quality of health services provided within the community-based health insurance scheme in Rwanda.

In chapter 1, "The Price Sensitivity of Demand for Health Insurance: Evidence from Community Based Health Insurance in Rwanda", I use the introduction of a new premium subsidy scheme to estimate the price sensitivity of the demand for community-based health insurance in Rwanda. The results indicate that the demand for health insurance is price sensitive but not elastic, suggesting that the demand for health insurance varies little in relation to the variation in price. Furthermore, my findings suggest that the price sensitivity varies among socioeconomic groups. I use the estimated price sensitivity to predict insurance coverage and the financial self-sustainability in relation to a number of plausible subsidy schemes. As a direct result of the inelastic insurance demand, take-up does not vary much across premium schemes. However, the heterogeneity in price sensitivity indicates that premium subsidies will affect the composition of enrolled individuals. To estimate the financial self-sustainability of the subsidy schemes, I match administrative data on the cost of providing the insurance scheme to estimate how much of the total insurer costs are covered by premiums for each alternative subsidy scheme. This allows me to control for the potential effects of adverse selection on patient costs. I find a positive slope of the cost curve, which is consistent with adverse selection, although the estimations suggest that the financial implications of this are limited. Overall, the results suggest that premium subsidies might represent an expensive policy tool for reaching universal heath coverage, one of the key targets in the Sustainable Development Goals.

In chapter 2, "The Role of Childcare in Firm Performance: Evidence from Female Entrepreneurship in Mexico", I study the impact of a national daycare program on the performance of female-run microenterprises. *Estancias Infantiles para Apoyar a Madres Trabajadoras* offered subsidized childcare to children younger than four years old, whose mother was working in the informal sector. I explore the variation in availability of the program in a difference-in-difference design, and compare outcomes for women with children just below and above the eligibility threshold for the program. Furthermore, I explore the roll-out of the childcare program in a triple-difference design with treatment intensity that varies across municipalities and over time. I find no evidence that the program was associated with changes in female entrepreneurship and business performance, proxied by number of workhours, physical capital, or the likelihood of operating the business from home, having an employee, or applying for a credit. This paper is one of the first papers to study the importance of childcare obligations as a barrier for business performance among female-run microenterprises by evaluating the impact of a nationwide social policy program on female business performance.

The third chapter, "Variation in the Quality of Primary Healthcare: Evidence from Rural and Urban health services in Rwanda", I document disparities in the quality of healthcare between rural and urban primary health clinics in Rwanda, and to what extent structural and contextual factors can explain such disparities. I use administrative data on 12 quality indicators from each health facility to measure the quality of healthcare. This data was collected through direct observations and chart examinations, as performed by professional health staff during regular quality inspections at the health clinics. I use the quality indicators to construct two quality scores for each health facility—one general score that includes health services, administration and management, as well as laboratory work, and a second patient centered score that focuses on the quality of patient services. The findings show a significant difference in quality between rural and urban health clinics, where rural clinics underperformed in relation to urban clinics in all dimensions. Furthermore, the results indicate that differences in such factors only explain a small part of the disparities in health quality. This paper contributes to the discussion on how to erase inequality in health services within developing countries by providing evidence suggesting that variation in structural inputs is unlikely to erase such disparities.

References

Aizer, A., Eli, S., Ferrie, J., & Lleras-Muney, A. (2016). The long-run impact of cash transfers to poor families. *American Economic Review*, *106*(4), 935–71.

Aryeetey, G. C., Jehu-Appiah, C., Spaan, E., Agyepong, I., & Baltussen, R. (2012). Costs, equity, efficiency and feasibility of identifying the poor in ghana's national health insurance scheme: empirical analysis of various strategies. *Tropical Medicine & International Health*, *17*(1), 43–51.

Baird, S., McIntosh, C., & Özler, B. (2019). When the money runs out: Do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics*, *140*, 169–85.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., & Pellerano, L. (2016). *Cash transfers: What does the evidence say? A rigorous review of impacts and the role of design and implementation features* (Tech. Rep.). London: Overseas Development Institute.

Behrendt, C., & Nguyen, Q. A. (2018). *Innovative approaches for ensuring universal social protection for the future of work* (Future of Work Research Paper Series No. 1). Geneva: International Labour Organization.

Binswanger, H., & Rosenzweig, M. (1993). Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal*, *103*(416), 56–78.

Brown, C., Ravallion, M., & Van de Walle, D. (2016). *A poor means test? econometric targeting in africa* (Working Paper No. 22919). Cambridge, MA, USA: National Bureau of Economic Research.

Capuno, J. J., Kraft, A. D., Quimbo, S., Tan Jr, C. R., & Wagstaff, A. (2016). Effects of price, information, and transactions cost interventions to raise voluntary enrollment in a social health insurance scheme: A randomized experiment in the Philippines. *Health economics*, *25*(6), 650–62.

Cole, S., Giné, X., & Vickery, J. (2017). How does risk management influence production decisions? Evidence from a field experiment. *The Review of Financial Studies*, *30*(6), 1935–70.

Dercon, S., & Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from ethiopia. *Journal of development economics*, *96*(2), 159–73.

Gertler, P., & Gruber, J. (2002). Insuring consumption against illness. *American economic review*, *92*(1), 51–70.

Gertler, P. J., Martinez, S. W., & Rubio-Codina, M. (2012). Investing cash transfers to raise long-term living standards. *American Economic Journal: Applied Economics*, *4*(1), 164–92.

Handa, S., Natali, L., Seidenfeld, D., Tembo, G., Davis, B., & Team, Z. C. T. E. S. (2018). Can unconditional cash transfers raise long-term living standards? Evidence from Zambia. *Journal of Development Economics*, *133*, 42–65.

Haushofer, J., & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: experimental evidence from kenya. *The Quarterly Journal of Economics*, *131*(4), 1973–2042.

Heltberg, R., & Lund, N. (2009). Shocks, coping, and outcomes for pakistan's poor: health risks predominate. *The Journal of Development Studies*, *45*(6), 889–910.

Hossain, Z. (2011). *Extreme poor adivasis and the problem of accessing social safety nets* (Shiree Working Paper No. 4). Dhaka, Bangladesh: Shiree, Extreme Poverty Research Group.

International Labour Office. (2017). *World Social Protection Report 2017–2019: Universal social protection to achieve the Sustainable Development Goals* (World Social Protection Report). Geneva: International Labour Organization.

International Labour Office. (2018). *Women and men in the informal economy: A statistical picture* (Tech. Rep.). Geneva: International Labour Organization.

Islam, A., & Maitra, P. (2012). Health shocks and consumption smoothing in rural households: Does microcredit have a role to play? *Journal of development economics*, *97*(2), 232–43.

Kalisa, I., Musange, S., Collins, D., Saya, U., Kunda, T., & Parfait, U. (2016). *The development of community-based health insurance in rwanda: Experiences and lessons* (Tech. Rep.). Kigali: University of Rwanda College of Medicine and Health Sciences, School of Public Health.

Leive, A., & Xu, K. (2008). Coping with out-of-pocket health payments: Empirical evidence from 15 african countries. *Bulletin of the World Health Organization*, *86*, 849–56C.

Liu, K. (2016). Insuring against health shocks: Health insurance and household choices. *Journal of health economics*, *46*, 16–32.

Mitra, S., Palmer, M., Mont, D., & Groce, N. (2016). Can households cope with health shocks in vietnam? *Health economics*, *25*(7), 888–907.

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica: Journal of the Econometric Society*, 539–91.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *Journal of Economic perspectives*, *9*(3), 83–102.

United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development* (A/RES/70/1). New York: United Nations. Retrieved from https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf (Accessed April 23 2021)

World Health Organization, & World Bank. (2019). *Global monitoring report on financial protection in health 2019.* Retrieved from https://apps.who.int/iris/bitstream/handle/10665/331748/9789240003958-eng.pdf (Accessed on May 2 2021)

I

The Price Sensitivity of Demand for Health Insurance

Evidence from Community Based Health Insurance in Rwanda

Carolin Sjöholm*

Abstract

This study estimates the price sensitivity of the demand for health insurance, exploiting the variation in insurance premiums created by the implementation of a new premium subsidy scheme for community- based health insurance in Rwanda. I use the estimated price elasticity to predict the impact of a number of plausible premium subsidy schemes on two policy- relevant outcomes: insurance coverage and financial sustainability. I find that the demand for health insurance is inelastic, although the price sensitivity varies among different socioeconomic groups. The results suggest that premium subsidies have only a modest effect on the take-up of insurance compared with nonsubsidized premiums, but they affect the composition of individuals enrolled in the insurance. To simulate the financial sustainability of the insurance scheme, measured as the share of total insurer costs covered by insurance premiums, I combine the price elasticity estimates with unique data on insurer costs, enabling me to account for adverse selection. I estimate a positive slope of the average cost curve, consistent with adverse selection. These results indicate that premium subsidies might not represent a financially sustainable policy tool for achieving universal healthcare.

JEL classification: I13, I18, D12, H51, H55
Keywords: community-based health insurance, adverse selection, price sensitivity

*University of Gothenburg, carolin.sjoholm@economics.gu.se. I would like to thank my advisors Måns Söderbom and Annika Lindskog for their invaluable input. This paper has also benefited greatly from comments from Andinet Woldemichael, Anja Benshaul-Tolonen, Anna Bindler, Joe Vecci, and Josephine Gakii. I thank participants at the CSAE Conference (2019), the Pacific Conference for Development Economics (2019), the Midwest International Economic Development Conference (2019), and the Workshop on Public Health and Development at the University of Gothenburg (2018). I am also grateful to Dr. Pascal Kayobotsi, Andrew Muhire, and Emmanuel Ntawuyirusha for sharing their knowledge and expertise regarding the CBHI scheme in Rwanda.

1 Introduction

In 2015 the UN General Assembly included universal health coverage as part of the overall commitment to the Sustainable Development Goals (United Nations, 2015). Community-based health insurance (CBHI) has been adopted by many developing countries as a financing mechanism to reach this goal by pooling health risks and resources at the community level.[1] So far enrollment in CBHI has often been low, particularly among poor households (Gnawali et al., 2009; Yilma et al., 2015; Parmar et al., 2014). In order to increase enrollment levels, many countries have implemented premium subsidies.[2] However, premium subsidies are costly and contribute to a lack of self-financing of the insurance schemes as premium revenues cover only a small share of the patient costs. In addition to impacts on insurance enrollment, premium subsidies might affect the type of individuals who enroll, resulting in an association between premiums and the insurer costs. Previous research from developing country contexts suggests that premium subsidies could exacerbate the effects of adverse selection (Parmar et al., 2012), which would negatively affect the financial sustainability of the CBHI scheme.

The main contribution of this paper is to study the impact of premium subsidies on policy relevant outcomes such as insurance coverage and the financial self-sustainability of the CBHI scheme. I use the introduction of a new premium scheme as a quasi experiment to estimate the price sensitivity of demand for health insurance. This is done in the context of Rwanda, a low-middle income country in Africa. Next, I use the estimated price sensitivity to predict enrollment levels, and subsequently premium revenue, for a number of plausible premium subsidy schemes. In order to evaluate the financial sustainability associated with the different subsidy schemes, I use unique data on the total insurer costs related to the CBHI to consider the potential effects of adverse selection on the cost of providing the health insurance. To the best of my knowledge, this is the first attempt to estimate the consequences of premium subsidies on the financial sustainability of health insurance in a developing country context, considering the financial impact of adverse selection.

In 2011 the government of Rwanda replaced a uniform subsidy scheme, in which all individuals paid the same premium, with a targeted premium subsidy that was directed

[1] India (Aggarwal, 2010), Uganda (Basaza et al., 2008), Burkina Faso (Fink et al., 2013)

[2] Mexico (Bosch et al., 2012), Vietnam (Wagstaff et al., 2016), and Ghana (Asuming et al., 2017), Burkina Faso (Parmar et al., 2012) offers targeted premium subsidies to poor households.

to households with low socioeconomic status. The aim of the targeted subsidy was to increase access to healthcare among poor households and to improve the financial sustainability of the CBHI scheme (Kalisa et al., 2016). The categorization of households into subsidy groups was based on Ubudehe, a classification system developed by the Rwandan government to categorize all households according to socioeconomic status. This resulted in a stratified premium scheme in which households categorized as having low socioeconomic status received fully subsidized premiums whereas relatively wealthier households were subject to a price increase. Exploiting the variation in premium costs created by the policy reform, I estimate the price elasticity of insurance demand using a linear probability model with individual fixed effects. Knowledge of the price sensitivity of demand can inform policy makers regarding the effciency of premium subsidies as a policy tool to promote universal health insurance.

I use panel data from the Rwandan Integrated Household Living Conditions Survey (EICV) in 2010/11 and 2014 to estimate the price sensitivity of the demand for health insurance. The results indicate that the demand for insurance is sensitive to price change but it is not price elastic.[3] An increase of the premium costs by 1,000 Rwandan franc (RwF, corresponding to approximately USD 1.1), is associated with a decrease in the likelihood of enrolling by 10.9 percentage points (ppt) (15% at the mean). This implies an average elasticity of –0.18, indicating that the change in demand is small in relation to the price change. The estimated elasticity is considerably lower than both the elasticity of the demand for the National Health Insurance Scheme in Ghana estimated by Asuming et al. (2017) and the price elasticity of demand for preventive health products such as bed nets and deworming medicine (Dupas, 2011; Kremer & Miguel, 2007; Cohen & Dupas, 2010). The effect of changes to premium costs is heterogeneous between different subgroups of households. Individuals living in poor households or households headed by women have a higher price elasticity than individuals in non-poor or male-headed households. As a result of the heterogeneity in price sensitivity, the composition of beneficiaries varies among different subsidy schemes.

The association between household socioeconomic status and insurance premiums, caused by the introduction of the stratified premium scheme, suggests that endogeneity may be a concern for the interpretation of my results. I demonstrate the robustness of my results to omitted variable bias using the Oster (2019) test. Furthermore, I estimate the price sensitivity using samples that are balanced on observable characteristics.

[3]The demand is considered inelastic if the price elasticity is $< |1|$; that is, a given percentage change in the premium cost will cause a smaller percentage change in the demand for insurance.

Comparing individuals with increasingly similar characteristics decreases the concerns that omitted time-varying factors are driving the price sensitivity estimates.

Overall, the results indicate that government subsidy strategies will have a limited effect on insurance coverage. This is a direct effect of the inelastic demand. For example, I find that a decrease in the overall premium costs from RwF 3,000 to RwF 1,000 (USD 3.4 to USD 1.1) would increase take-up from 66% to 77%. Additionally, a subsidy scheme that offers completely subsidized premiums for young children under six years old corresponds to a predicted take-up of 67%. Overall, the simulations indicate that the average insurance coverage remains relatively constant for different subsidy schemes.

Financial sustainability is calculated as the share of insurer cost that is covered by household premiums. I simulate the financial coverage related to the different pricing strategies by calculating the share of insurer costs covered by premium revenue. In addition to considering the potential effects on enrollment levels, this forces me to further consider the association between premium levels and the cost of providing health insurance. In the presence of selection, changes to the insurance premiums will affect the cost of providing the insurance as the composition of insurance beneficiaries changes in response to the changes in premium costs. As the premium costs increase, so does the cost of providing insurance. Following the analysis presented by Einav et al. (2010), I use unique administrative data on the total costs of providing CBHI in Rwanda and provide evidence of a positive association between insurer costs and premium costs by estimating the average cost curve for administrative sections.[4] A positive slope of the average cost curve indicates that the average insurer cost among enrolled households in a section increases as the average premium level increases, consistent with adverse selection.

I use the association between patient costs and insurance premiums to calculate the financial sustainability in relation to alternative premium schemes. The simulations indicate that the financial coverage of alternative premium subsidies differs depending on whether the insurance market is adversely selected. In the absence of selection, the range of financial coverage levels is wider, between 0.28 and 0.85. In the absence of adverse selection, insurer costs are constant among the different subsidy schemes and variation in the financial coverage is driven by changes in enrollment. Considering the adverse selection scenario, the financial coverage reaches levels between 0.35 and 0.80

[4]A section is an administrative unit for the CBHI scheme that approximately represents the catchment area of a health center

for the majority of subsidy schemes; that is, household premiums cover approximately 35%–80% of the insurer costs. Not surprisingly, the difference between the level of financial coverage in the selection and the levels in no-selection scenarios increases as the premium levels deviate from the mean cost. Differences between average insurer costs in scenarios with and without adverse selection are indicative of the financial implications of selection. This is important knowledge that can inform policymakers on how adverse selection translates into future costs faced by the insurer.

This study makes two contributions to the literature. First, it adds to a relatively small and recent body of literature that seeks to evaluate the role of premium subsidies in the take-up of health insurance in a developing country context (Thornton et al., 2010; Asuming et al., 2017; Capuno et al., 2016; Wagstaff et al., 2016).[5] This literature primarily relies on experimental variation in premium costs to identify the effects of short-term premium subsidy interventions on insurance take-up. To date, the empirical evidence is inconclusive. While some studies find no evidence that premium subsidies represent an efficient policy tool to increase take-up (Capuno et al., 2016; Wagstaff et al., 2016),[6] others find positive impacts on enrollment (Thornton et al., 2010; Asuming et al., 2017).[7] My study contributes to this literature by providing evidence from a nationwide policy intervention that resulted in a considerable and indefinite price change. This is important since previous research argues that onetime external subsidies alone are often insufficient to encourage the take-up of health products (Kremer & Miguel, 2007). Furthermore, this study evaluates the demand of a popular insurance scheme with a high enrollment rate. During 2011, 67% of the target population were enrolled in the insurance scheme. This is a high number compared with enrollment rates in other countries such as Burkina Faso, at 6%; Ghana, at 38% (Chemouni, 2018); and

[5] Another type of insurance that has received much attention in the literature is index-based crop insurance. Evidence from this literature indicates that demand for insurance is price sensitive, but that the insurance has low take-up rates at actually fair prices (Cole et al., 2013; Karlan et al., 2014; Mobarak & Rosenzweig, 2014).

[6] Capuno et al. (2016) find that a 50% premium subsidy in combination with increased access to information regarding the insurance led to a 3% increase in demand among informal worker households in the Philippines. A 25% premium subsidy contributed to an increase in enrollment by 3.5 ppt in Vietnam (Wagstaff et al., 2016).

[7] Asuming et al. find that households that received a premium subsidy were 38 ppt more likely to enroll in the national health insurance scheme when receiving a subsidy that covered 1/3 of the price. Furthermore, when premiums were fully subsidized, enrollment increased from 27% to 75%, indicating that the demand is price sensitive. Levine and colleagues (2016) find that a premium subsidy of 80% contributes to an increase in enrollment in the SKY social health insurance in Cambodia. In contrast to previous studies, these results indicate that the demand is price elastic (–1.1). The Cambodian study deviates from the other papers by evaluating an insurance scheme that targets rural populations.

Laos, at 2% (Alkenbrack et al., 2013). In the context of very low enrollment rates, alternative factors such as low quality of care, limited access to care, or limited information about the insurance are alternative factors that might represent first-order barriers to insurance enrollment.

Second, this paper adds to a an emerging literature on adverse selection in the developing country context,[8] by estimating how insurer costs vary as the insurance premiums change. Using this method, the slope of the cost curve provides a test for selection in the CBHI market (Einav et al., 2010). This method has been frequently used in developed countries (Bundorf et al., 2012; Einav et al., 2010), but this paper is one of the first to provide evidence from the developing country context. Fischer and colleagues (2018) find strong evidence of adverse selection for individual insurance policies by using experimental variation in insurance price to identify the cost curve for hospitalization insurance in Pakistan. Hospitalization insurance is a specialized health insurance that insures households against severe health shocks that require hospital care. This study differs from that of Fischer and colleagues (2018) because it evaluates an insurance scheme that covers services at all service levels in the Rwandan healthcare system, including preventive healthcare. This is a first step in using administrative cost data on insurer costs to estimate the financial implications of selection in health markets in the developing country context.

This paper is organized as follows: Section 2 provides a description of the study context and the CBHI scheme in Rwanda. Section 3 describes the data. The empirical strategy is presented in Section 4, followed by the results and a sensitivity analysis in Section 5. Section 6 provides the results from the financial self-sustainability analysis, and Section 7 concludes.

[8] Earlier literature from the developing country context has primarily used the correlation between ex-ante individual health risk and the likelihood of enrollment (Wang et al., 2006; Zhang & Wang, 2008), as well as the positive correlation test measuring the correlation between insurance coverage and individual risk (Chiappori & Salanie, 2000), to identify adverse selection. The results are mixed: while some studies find evidence of adverse selection (Wang et al., 2006; Zhang & Wang, 2008; Lammers et al., 2010), others do not (Nguyen & Knowles, 2010; Banerjee et al., 2014). Importantly, the literature is often limited to evaluating the relationship between baseline health risk and insurance take-up, few studies consider the financial implications of adverse selection due to a lack of data on the costs of insurance schemes.

2 Community-Based Health Insurance

Community-based health insurance (CBHI) was introduced by the Rwandan government in 1999 as a result of limited utilization and ability to pay for healthcare services among large segments of the population. The main objectives of the CBHI scheme are to provide equal access to healthcare services and to prevent people from catastrophic healthcare costs by pooling resources in district and national risk pools (Kalisa et al., 2016).

CBHI was initially introduced as a pilot project in 3 of the countries 30 districts —Kabgayi, Kabuyare and Byumba—covering 52 health centers and 3 hospitals. During the following years, similar insurance schemes were introduced in other districts throughout the country. In 2006, a national policy was implemented that standardized the different regional schemes and developed a national health insurance scheme. The national scheme is centrally managed by the ministry of health, which is responsible for the overall policy development of the CBHI. At the same time, the CBHI continues to be a highly decentralized insurance scheme that is coordinated in 30 administrative districts. Each CBHI district is a legal body with branches –CBHI sections–at all health centers in its geographic area. Each CBHI section represents approximately the catchment area of a health center (Kalisa et al., 2016).

Enrollment in CBHI has increased drastically during the last decade. Appendix Figure A1 shows national CBHI enrollment levels during the period 2003–15. Insurance coverage increased sharply and peaked around 2010 with enrollment levels around 90%. In the following years, and in conjunction with the introduction of the new premium scheme, enrollment levels have been volatile and decreasing. Despite the recent development, insurance enrollment is high compared with that in other countries with similar insurance schemes: the National Health Insurance Scheme in Ghana reached a coverage level of 40% in 2014 (Wang et al., 2017); the CBHI in Ethiopia, 8%; a national health insurance scheme in Nigeria, 3% (Chemouni, 2018); and Vietnam Social Security, 42% (Lagomarsino et al., 2012).

One reason for the relatively high enrollment levels in Rwanda could be explained by a strong policy focus on improved accessibility and quality of healthcare implemented in part by the introduction of a performance-based financing scheme (Ministry of Health, 2012). During the last decades, Rwanda has recorded an impressive improvement in public health outcomes (World Bank, 2020),[9] which has been accompanied by a large increase in resources through CBHI insurance schemes, resource mobilization, and external funds. Between 1998 and 2010 health expenditure increased from USD 10 to USD 40 per capita (Ministry of Health, 2015).

Figure 1 shows the distribution of health centers covered by the insurance scheme. The facilities are distributed all over the country, with a high concentration in the capital of Kigali. Each health center has a CBHI health section, which results in 100% geographic coverage. All public health facilities are covered by the CBHI (Kalisa et al., 2016).

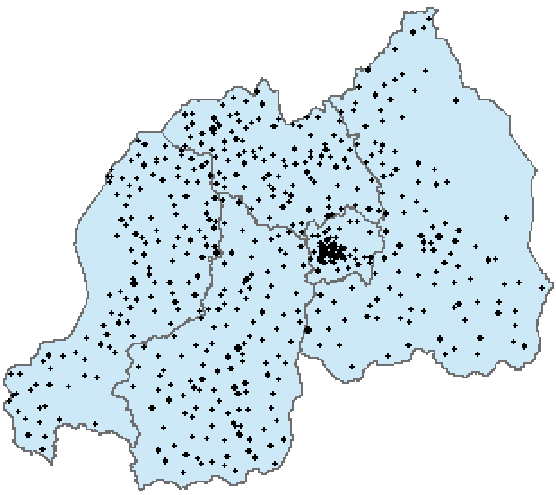


Figure 1: Distribution of CBHI health centers in Rwanda

From a financial point of view, the CBHI scheme can be described as a local insurance scheme that pools funds at the national and district levels.The CBHI scheme is financed mainly through member premiums, which represent approximately 66% of the total budget Kalisa et al. (2016). These monetary contributions are received at the community level and used to reimburse health centers for services provided. Approxi-

[9]Between 1996 and 2018, life expectancy almost doubled, from 35 to 69 years, and the under-five mortality rate dropped from 196 to 35 per 1,000 births, and between 2000 and 2018, the maternal mortality rate dropped from 1,160 to 248 per 100,000 live births (World Bank, 2020).

mately 55% of the premium revenues remain in the CBHI section, while the remaining premium contributions are channeled to fund the district risk pools in order to cover hospital claims. Approximately 10% of the funds directed to the district risk pools are forwarded to a national risk fund that covers services at the referral hospitals. The Rwandan government represents the second-largest source of funding (14%), primarily covering the contributions of indigent members. The global fund covers approximately 10% of the total budget, and patient copayments and contributions from other health insurance schemes in the country cover the remaining costs. Since 2008 the Rwandan government has mandated that other insurance companies provide 1% of their income to the CBHI (Kalisa et al., 2016).

In addition to the premium, members pay a flat copayment of RwF 200 each time they visit a health center, as well as a copayment of 10% of the total hospital bill. The new premium schedule was meant to increase the financial sustainability by increasing premium revenues and reducing dependence on external subsidies (Kalisa et al., 2016).

CBHI beneficiaries are entitled to predefined packages of healthcare services known as the minimum package of activities (MPA) and the complementary package of ctivities (CPA). The MPA covers preventive, promotional and curative health services provided at the health centers, whereas the CPA includes services provided at the national hospitals. The service packages are defined by the Ministry of Health. Beneficiaries can access healthcare at public health facilities at all levels of the public healthcare delivery system: health centers, district hospitals and referral hospitals. However, the insurance does not cover healthcare at private health facilities (Kalisa et al., 2016).

In conjunction with the standardization of the CBHI scheme in 2006, a uniform premium system was developed and introduced. This system required members to pay an individual annual premium of RwF 1,000 (USD 1.1). The premium level was set to cover the cost of health services provided at the health centers, but did not cover costs associated with secondary level care such as at district and national hospitals (de la Sante, 2004). As a result, the premium costs paid by insurance beneficiaries were significantly subsidized through a uniform premium subsidy scheme. Furthermore, the Rwandan government offered full premium subsidies to indigent households. Despite this targeted subsidy, the premium structure was considered strongly regressive and exclusive of the poor (Kalisa et al., 2016).

To promote equal access to healthcare, in 2011, the flat-rate premium was replaced by a stratified premium system based on targeted subsidies to households with low socioeconomic status. Besides increasing the equity of the insurance, the targeted subsidies were part of an overall policy change that aimed to strengthen the administrative structure of the insurance scheme. Another aim was to increase the financial self-sustainability of the insurance. The new subsidy scheme was based on *Ubudehe categories*, a socioeconomic classification system developed by the Ministry of Local Government.

The Ubudehe system was first introduced by the Rwandan government in 2001, well before the introduction of the targeted premium subsidy. This is important because it verifies that the classification of households into Ubudehe groups was not developed with the aim of determining the premium costs for the CBHI. Households using a community-based targeting process, with each community divided into cells, small groups of approximately 10 households each. Each cell has a supervisor who is responsible for keeping track of and updating the categorization of the households in the cell. To control the reporting of Ubudehe categorization, the complete list of households is revised yearly by the whole village on one of the village work days (Umuganda) that are mandatory for all residents. The categorization is revised yearly by the Ministry of Health at the national level (Kayobotsi, 2019).

Under the Ubudehe system, the population was divided into six categories reflective of socioeconomic status. The system considers a wide range of socioeconomic factors including household nutrition, financial and nonfinancial assets, access to property, household livelihood and production capacity, with households in category 1 classified as living in abject poverty and those in category 6 classified as money rich (MINECOFIN, 2002). (For further description of the Ubudehe classification, see appendix table A1) Importantly, although the Ubudehe classification system is correlated with the national poverty measure, based on household consumption levels, this measure does not perfectly predict household Ubudehe classification. Appendix table A2 shows the relation between Ubudehe categorization and poverty classification according to the national poverty line. The majority of both poor and nonpoor households were categorized as Ubudehe group 3. However, a larger share of poor households were placed in lower Ubudehe categories relative to the share of nonpoor households in the higher categories. Importantly, both poor and nonpoor households are present in all Ubudehe categories.

As shown in table 1, households in premium category—1 that is, the lowest two Ubudehe groups, 1 and 2—were subject to an annual individual premium of RwF 2,000

(USD 2.2) according to the new premium scheme implemented in 2011. However, this premium is completely subsidized by the government. Ubudehe groups 3 and 4 fell into CBHI premium category 2, paying a premium of RwF 3,000 (USD 3.4) per individual and year. Households in the two highest Ubudehe groups, 5 and 6, were placed in premium category 3, with a premium of RwF 7,000 (USD 7.9) per year. All members of a household were subject to the same premium level. On average, the policy change resulted in an increase in premium costs among beneficiary households. While the number of beneficiaries who received fully subsidized premiums increased, households in the two highest premium categories faced a relatively sharp price increase following the policy change. According to the summary statistics presented in appendix table A3, the yearly consumption of a household was approximately RwF 218,426 (USD 221). The average household included 5 household members, so a household in premium category 2 would pay RwF 15,000 to enroll all family members in the CBHI scheme, and therefore the premium cost would represent approximately 7% of their yearly consumption. Importantly, beneficiaries of the CBHI scheme are enrolled on an individual basis, paying individual premiums to enroll.

Table 1: Ubudehe and Premium Categories

Ubudehe	(1) CBHI premium	(2) Premium before 2011 (RwF)	(3) Premium after 2011 (RwF)
Ubudehe 1 (abject poverty) Ubudehe 2 (very poor)	Category 1	0 or 1,000	0
Ubudehe 3 (poor) Ubudehe 4 (resourceful poor)	Category 2	0 or 1,000	3,000
Ubudehe 5 (food rich) Ubudehe 6 (money rich)	Category 3	1,000	7,000

Notes: Column 1 shows the CBHI premium categories based on the Ubudehe groups in the leftmost column. Column 2 presents the premium scheme before the policy change in 2011, and column 3 shows the premium scheme after the policy change. RwF 1,000 is equivalent to approximately USD 1.1, RwF 3,000 to USD 3.4, and RwF 7,000 to USD 7.9.

3 Data

This analysis is based primarily on data from the Rwandan Integrated Household Living Conditions Survey (EICV), a household survey representative at the national level, using two survey rounds conducted in 2010–11 (EICV3) and 2014 (EICV4). A subsample of the households surveyed in 2010–11 were tracked and interviewed again in 2014, resulting in a panel with 2108 households.

Data for the EICV3 were collected during a one-year period between October 2010 through October 2011, both before and after July 1, 2011, when the new CBHI premium scheme was implemented. The baseline data have been adjusted to include only those households interviewed before the policy change. Furthermore, for the purpose of this analysis, 188 households that had at least one family member who reported being enrolled in another health insurance scheme were dropped. I restrict the sample to include only households where age and sex are consistent for household members between the rounds.[10] The above data adjustments resulted in a final data set with 937 households and 3806 individuals. The data include household demographics, socioeconomic characteristics, wealth, employment and health conditions.

Appendix table A3 presents summary statistics for all households in the sample. The results suggest that approximately 40% of the sample had access to piped water and 78% to improved sanitation such as toilet or latrines with slab. Almost half of the individuals in the sample stated that they worked and 87% lived in a rural household. Approximately half were younger than 20 years old, and around 13% were older than 50 years. Women made up 53% of the sample, and households had on average five members. Nearly 42% of the sample lived in households that were categorized as poor based on the national poverty line. A large majority of households, 90%, were classified in Ubudehe groups 2 and 3. Additionally, approximately 9% of the sample belonged to Ubudehe group 4 and 1% to group 1. Ubudehe group 5 represented less than 1% of the sample, and no observations were categorized in group 6. Individuals in the two highest Ubudehe categories were likely to be enrolled in private insurance schemes.

[10]I dropped 117 households due to error in coding. Given changes in individual characteristics, such as sex and age, the data suggested that individual identifiers were used for different people and household sin the follow-up data survey in 2014.

3.1 Indigent households

The EICV data do not contain information regarding household insurance premiums. To define the premium cost for each household after the introduction of the new premium scheme, I use information on household Ubudehe category. This allows me to define the premium cost of enrolled households as well as the premiums that unenrolled households would have been subject to if they enrolled in the CBHI.

Before the introduction of the stratified premium scheme, households paid a uniform premium to enroll in the CBHI scheme, independent of their Ubudehe category. According to administrative documents, however, the premium for indigent households was already completely subsidized before the policy change. I lack information that allows me to identify the subsidized households. However, administrative documents indicate that approximately 11%–16% of the poorest households were considered destitute (Kalisa et al., 2016; Lu et al., 2012; Kalk et al., 2010).

I use a number of alternative strategies to identify the indigent households base on household Ubudehe category and a national poverty measure based on household consumption. The definitions are presented in table 2. First, households that were categorized in Ubudehe group 1 or 2 and simultaneously were defined as extremely poor by the national poverty measure are defined as indigent. Second, I use a definition based exclusively on household consumption level, defining households with a consumption level below (i) the 10th percentile or (ii) the 16th percentile as indigent.

Importantly, information about household Ubudehe category is available only in the post-treatment data (2014). I use two strategies to define household Ubudehe category prior to the policy change. First, I make the assumption that the Ubudehe category is constant during the study period (premium Ubudehe). Second, I use a number of household characteristics to predict the likelihood of being categorized in Ubudehe group 1 or 2 (predicted Ubudehe) in the post-treatment data. Appendix table A4 presents the estimated correlation between these households characteristics and the likelihood of being categorized in Ubudehe group 1 or 2 in 2014. Households with a predicted likelihood that exceeds the 75th percentile are categorized in Ubudehe group 1 or 2.

By using a number of different definitions of the households that were exempted from insurance premiums in 2010–11, I show that the estimated price elasticity is robust to the definition of these households. The results are presented in the section 5.

Table 2: Alternative definitions of completely subsidized, indigent households in 2010–11

Ubudehe	Households categorized as Ubudehe 1 or 2 (in 2014) and defined as extremely poor according to national poverty measure
Predicted Ubudehe	Households predicted as Ubudehe category 1 or 2 and defined as extremely poor according to national poverty measure
16th percentile	16 percent of households with the lowest consumption
10th percentile	10 percent of households with the lowest consumption

Note: Households categorized as indigent received subsidized insurance premiums before the policy change in 2011.

4 Empirical Strategy

The introduction of the new premium scheme implied that households in different Ubudehe categories were subject to different price changes as a result of the policy change. I use the price variation to identify the price sensitivity of the demand for CBHI in Rwanda. Table 3 provides a detailed description of the price variation caused by the policy change as well as the corresponding variation in insurance enrollment. The results are given separately for the alternative definitions of indigent households.

Table 3 presents the full range of price variation in the data. According to the stratified premium scheme, households that paid a premium of RwF 1,000 in the baseline were subject to (i) a full premium subsidy after the policy change (row 1), (ii) a price increase by RwF 2,000 if they ended up in premium category 2 after the premium change (row 3), or (iii) an increase by RwF 6,000 if they were classified as premium category 3 (row 5). Households that were exempted from paying an insurance premium before the policy change either ended up in premium category 1 after the policy change and continued to get fully subsidized premiums (row 2) or were classified as category 2 and faced an increase in the premium cost by RwF 3,000 (row 4).

In addition to the different premium changes, table 3 shows related variations in insurance enrollment as well as the number of individuals affected by each price change (N). The results are included for each alternative definition of indigent households. As expected, the results show that the number of households that were subject to each price change depends on the definition of the insurance premium prior to the policy change. The majority of individuals, 73%–65%, endured a premium increase by RwF 2,000 as a result of the new policy scheme. Additionally, approximately 17%–22% were subject to a price decrease by RwF 1,000, whereas 4%–10% of the sample received full premium subsidies both before and after the premium change and were consequently not affected by the new premium scheme. This group is smaller when indigent households are defined in relation to the 10th percentile of household consumption. Finally, given the lack of coherence between the Ubudehe categorization and the poverty measure based on household consumption, households that were defined as destitute in 2010 could face premium costs of RwF 3,000 after the policy change.[11] Less than 1% of the sample endured a price increase of RwF 6,000.

Overall, the results indicate that the distribution of price variations caused by the implementation of the new premium scheme is robust to different definitions of the premium scheme in the pre-treatment period. Additionally, table 3 describes changes in insurance enrollment associated with each premium change. The association between insurance premiums and enrollment is stable across the different definitions of indigent households, but varies among premium groups. The results suggest that there is a negative association between changes in insurance premiums and enrollment within almost all premium groups. There is an increase in insurance enrollment within the group of households that were not affected by the premium change (Δ premium $= 0$). The positive variation in insurance enrollment within this premium group is explained by variables other than the cost of insurance premium, such as for example information campaigns or public efforts to decrease obstacles for enrollment among these households. This analysis focuses on the importance of variation in premium costs to explain variation in insurance enrollment.

[11] This price variation is dropped as a result of the construction of the premium scheme in the pre-treatment period when indigent households are defined based on the assumption of constant Ubudehe categories (Ubudehe, columns 1 and 2).

Table 3: Variation in premium costs and enrollment levels before and after policy change for different definitions of indigent households

	Ubudehe		Pred. Ubudehe		16th pctl.		10th pctl.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ premium	Δ enrollm.	N (%)	Δ enrollm.	N (%)	Δ enrollm.	N (%)	Δ enrollm.	N (%)
1. -1000	0.175	659 (17)	0.169	738 (19)	0.149	731 (19)	0.174	837 (22)
2. 0	0.226	340 (9)	0.257	261 (10)	0.310	268 (7)	0.284	162 (4)
3. 2000	-0.120	2797 (73)	-0.117	2617 (69)	-0.105	2456 (65)	-0.112	2576 (69)
4. 3000	—	—	-0.156	180 (5)	-0.226	341 (9)	-0.213	221 (6)
5. 6000	-0.20	10 (0)	-0.20	10 (0)	-0.20	10 (0)	-0.20	10 (0)

*Notes:*The table describes the price variation used to identify the price elasticity in the analysis and the corresponding variation in insurance enrollment. The results are presented separately for each alternative definition of the premium scheme prior to the policy change (indigent households).

I plot the within-group variation in insurance premiums and enrollment in appendix figure A2 for each separate premium group, as described in table 3. Despite low variation in premium costs, the figures suggest that the relationship between the within-group variations in premium and enrollment is approximately linear and negative. Given the distribution of the observations in combination with the limited price variation in the data, I estimate the price sensitivity using a linear estimation model.

Importantly, the different plots suggest that the distribution of observation is similar for the alternative definitions of indigent households. The results indicate that the estimated price elasticity is expected to be robust to the different definitions of indigent households. In the following analysis, I use the definition of indigent households based on constant Ubudehe categories as the preferred definition of indigent households, but provide price sensitivity estimates for all alternative definitions in appendix table A5.

I use the variation in insurance premiums caused by the introduction of the new premium scheme to identify the effect of a price change on insurance enrollment. My baseline specification to estimate this effect is presented in equation (1). I estimate equation (1) using a linear probability model (LPM) with individual fixed effects (FE):

$$\Pr(CBHI_{ijt} = 1 | X_{ijt}, \mu_i, p_{jt}^{CBHI}, \gamma_t) = \beta_1 p_{jt}^{CBHI} + X_{ijt}\beta_2 + \gamma_t + \mu_i + \epsilon_{it} \tag{1}$$

where i indexes individual, j household, and t time periods. $CBHI_{ijkt}$ indicates individual i's insurance status in time period t. The treatment variable, indicated by p_{jt}^{CBHI}, measures the premium level of each household in time t. X_{ijt} is a vector of individual and household time-varying factors that are potentially correlated with the outcome, such as age, labor and health status, household consumption, access to water and sanitation services, and travel time to closest clinic. μ_i are individual fixed effects, controlling for individual heterogeneity across individuals that are constant over time. Importantly, individual fixed effects control for time constant differences in individual health risk and underlying health status which are likely to be correlated with insurance status, as well as premium costs. The fixed effects also control for time-invariant differences among individuals belonging to different Ubudehe categories. Year fixed effects, γ_t, captures aggregate changes in insurance enrollment over time.

In all specifications, standard errors are clustered at the household level. By doing this, I allow for correlation in the error term for individuals within the same household. I choose this level of clustering because individuals within a household are exposed to a number of factors such as household composition and culture, as well as underlying health conditions that are likely to affect the individual decision to enroll in health insurance.

4.1 Threats to identification

The difficulty of evaluating the effect of insurance premiums on enrollment lies in the fact that the changes in premium schemes are often endogenously determined, that is, the premiums are often likely to be correlated with observable or unobservable characteristics of the population. As described earlier, the introduction of the stratified premium scheme for the CBHI in Rwanda was not random across individuals but depended on household socioeconomic status. Given the structure of the new premium scheme, households with a relatively higher socioeconomic status endured a larger price increase than households with lower socioeconomic status. At the same time, the detailed classification of household socioeconomic status makes it possible to identify households that are relatively similar in terms of a measure of socioeconomic status that is more complex than measures that are based on household income. In this context, the Ubudehe system allows me to identify households, in consecutive Ubudehe categories, that were potentially relatively similar in relation to a number of characteristics but were subject to significantly different price changes.

In order to produce unbiased estimates, the fixed effects estimator is based on the assumption that all unobservable factors that might simultaneously affect insurance enrollment and premium category are time-invariant. The individual fixed effects (μ_i) reduce the concern that differences in individual characteristics among premium groups drive the estimated price sensitivity by controlling for all variables that are consistent over time, such as preferences, risk aversion, and underlying health characteristics. However, time-varying unobservable characteristics remain a concern and a potential source of endogeneity. In this section, I discuss potential channels through which these variables potentially could confound the estimated price elasticity presented in next section.

The relationship between socioeconomic status and health is well documented in the economics literature, suggesting a positive association between health and socioeconomic status (Cutler et al., 2008). The individual fixed effects control for all differences in health status among premium categories that are constant over time. However, previous research also indicates that individuals with low socioeconomic status are more likely to suffer from health shocks than those with relatively higher socioeconomic status (Currie & Hyson, 1999; Currie & Stabile, 2002). As a result, systematic differences in health status among Ubudehe categories could imply that the effects of an adverse health shock, such as a malaria outbreak, disproportionately would affect individuals in the low premium group (who were subject to a price decrease), increasing their willingness to enroll in health insurance. If this shock coincided with the policy change, this could create an association between premium change and the willingness to enroll in health insurance even on the absence of price sensitivity. This would result in an upward bias of the estimated price sensitivity, leading to an upper bound of the price sensitivity. To control for potential differences in health changes among premium categories, I control for variation in health status before and after the policy change by including an indicator for whether an individual was sick during the previous two weeks as well as an indicator for disability.

Another potential concern related to differences in socioeconomic status among the different premium groups are public health interventions that exclusively target poor households in the low Ubudehe categories (groups 1 and 2). If the timing of such policy intervention coincided with the price change of the CBHI, and if the intervention resulted in improved health status among targeted households, this could result in a downward bias of the price sensitivity estimates. Households that received fully subsidized premiums after the policy change would simultaneously decrease their willingness to pay for insurance as a result of improved health status and lower expected health costs. Using the reverse logic, any simultaneous government intervention that contributed to improve incomes among households in the two lower Ubudehe categories would result in an upward bias of the price elasticity estimates, as the willingness to pay for insurance would increase as a result of improved incomes.

According to the Rwandan Health Sector Strategic Plan, one policy focus area has been increased access to improved water and sanitation services among poor households in the lower Ubudehe categories (Ministry of Health, 2012). I control for changes in access to water and sanitation services by including indicators for household main water supply and toilet facility in my specification. Importantly, these proxy variables adjust for all potential changes in unobservable factors that are correlated with improved water and sanitation services such as individual health status. I discuss the variation in access to these services in the next section.

Reverse causality represents another possible source of endogeneity. Increased access to health insurance is likely to affect individual health status and labor productivity positively, which in turn would affect household premium costs through improved socioeconomic status and Ubudehe group. However, because of the complexity of the Ubudehe categorization, household Ubudehe status is unlikely to change drastically from one year to another. To change the classification, households would have to show improvement across several wealth-related factors such as main livelihood, nutrition, assets and children's schooling. Despite the possibility that improved health and labor productivity could affect some of the Ubudehe indicators in the short run, other factors, such as access to land and housing, are likely to change only in the medium to long run, that is, not within the time frame of this study period. As a result, it is reasonable to think that the effects of reverse causality are likely to be larger in the medium and long run. Consequently, I consider the risk of reverse causality limited.

Finally, the identification strategy relies on the assumption that the Ubudehe classification process did not allow for elite capture. Elite capture would bias the results upward, leading to an upper bound of the true price sensitivity. As described earlier, the classification of households into Ubudehe categories is the result of a highly participatory process in which all households in a village have to agree unanimously on the classification of each household. This process works as a control function to minimize elite capture and maintain the accuracy of the Ubudehe categorization. The high level of transparency in the classification process contributes to decreasing the likelihood of elite capture as a common phenomenon. Community-based targeting has often been framed as a trade-off between the better information that communities have on the wealth levels of their population and the risk of elite capture in the targeting process. However, previous empirical research indicates that the community-based targeting is not significantly different from other targeting strategies and that elite capture does not affect the accuracy of beneficiary targeting (Alatas et al., 2012). Importantly, the downward bias caused by elite capture would not undermine the results and conclusions in the analysis.

In the next section, I address the plausible concerns related to omitted variable bias using the Oster (2019) method. The Oster test suggests that the results are robust to omitted variable bias. In addition to the Oster analysis, I estimate the price sensitivity using samples that are increasingly balanced on covariates. Balanced samples contribute to decreasing the correlation between the policy change and household socioeconomic status.

4.2 Differences among groups

As discussed in the previous section, the validity of the fixed effects estimator relies on the assumption that there are no confounding effects. This assumption requires that changes in insurance enrollment were not associated with differential changes in time-varying confounders among the different premium categories. One step in verifying the identifying assumption would be to provide evidence of parallel trends in the outcome variable prior to the treatment. However, lack of access to adequate pre-treatment data does not allow for a parallel trends analysis in this study. Instead, in this section I investigate changes in time-varying observables among premium groups, both pre- and post-treatment. Information regarding the balance in observable covarates between the premium groups, over time, could provide information regarding the importance of confounding effects.

Table 4 provides summary statistics of household and individual characteristics for premium categories 1 and 2. Columns 2 and 3 show the balancing tables for households in each premium categories related to a number of time-varying observable predictors. The results indicate that there are significant differences in characteristics between the two premium categories. Overall, the descriptive statistics indicate that individuals living in a households that are classified as premium category 1 (Ubudehe

1 and 2) have lower socioeconomic status than households in category 2 (Ubudehe 3 and 4). These differences in descriptive statistics confirm the accuracy of the Ubudehe classification system as a tool to categorize households according to their socioeconomic status. Individuals in premium category 1 are less likely to have access to piped water or have a flush toilet or latrine with solid slab, they have significantly lower household consumption, and they are less likely to run a nonfarm enterprise, although individuals in category 1 are more likely to be salary workers. There is no difference between the two groups regarding the likelihood of working.

The descriptive statistics in columns 1 and 2 show that there is a significant difference between the groups in health-related factors; that is, households in the lower premium category are more likely to report having experienced a health issue during the last two weeks prior to the survey, or to have a disability. Additionally, households in premium category 1 need to travel 0.38 hours longer on average to reach the nearest hospital. However, there is no significant difference in travel time to the nearest clinic. Importantly, the fixed effects estimation strategy allows for differences in characteristics among individuals in different premium groups, as long as these differences are constant across time and controlled for in the baseline specification.

The last three columns of table 4 display the estimated changes in the observable characteristics over time. Columns 5 and 6 present changes in individual and household characteristics between the periods before and after the introduction of the new policy scheme. Column 7 shows the difference in these changes between the two premium groups—that is, the difference-in-difference estimates. These estimates indicate that the baseline differences between the two groups persist over time for the majority of individual characteristics. Importantly, there is no significant change over time in the variation of individual health measures and consumption levels, factors that have a strong likelihood of simultaneously affecting Ubudehe category and the demand for health insurance. The results provide tentative evidence that differential changes in unmeasured variables between premium categories are not likely to drive the results in the analysis. However, unobservable and time-varying characteristics could still represent a threat to the identification.

I will discuss this further in the following section.

Table 4: Summary statistics for premium categories 1 and 2, baseline 2010—11

		Levels			Change		
Variables	(1) All	(2) Premium category 1	(3) Premium category 2	(4) Diff	(5) Premium category 1	(6) Premium category 2	(7) Diff-in-diff
Health issue	0.180	0.212	0.168	0.043***	0.065	0.069	-0.004
Disability	0.050	0.074	0.041	0.033***	0.008	-0.005	0.013
Piped water	0.321	0.290	0.332	-0.042**	0.120	0.158	0.038**
Sanitation	0.744	0.620	0.789	-0.169***	0.066	0.077	-0.010
Own house	0.924	0.881	0.940	-0.059***	0.021	0.013	0.008
Own land	0.974	0.961	0.978	-0.016***	-0.022	-0.015	-0.006
Work	0.468	0.454	0.472	-0.0178	0.102	0.067	0.035***
Salary worker	0.227	0.268	0.212	0.057***	0.047	0.020	0.027*
Own nonfarm enterprise	0.116	0.092	0.124	-0.032***	0.017	0.005	0.012
Poor	0.451	0.607	0.396	0.211***	-0.052	-0.055	0.002
Consumption HH	212,486	159,826	231,295	-71,468***	7,273.92	13,467	-6,193
Rural	0860	0.882	0.852	0.030**	0.038	0.013	0.025**
Female	0.530	0.557	0.521	0.036*	-	-	-
HH size	5.167	4.795	5.300	-0.504***	0.928	0.872	0.056
Travel time clinic	0.808	0.842	0.796	0.045	-0.107	-0.004	-0.111
Travel time hospital	3.111	3.390	3.012	0.378***	0.177	0.186	-0.009
Age 0–5	0.198	0.177	0.206	-0.028*	-	-	-
Age 6–19	0.344	0.379	0.331	0.048***	-	-	-
Age 20–29	0.142	0.108	0.154	-0.046***	-	-	-
Age 30–39	0.117	0.104	0.121	-0.017	-	-	-
Age 40–49	0.078	0.071	0.080	-0.009	-	-	-
Age 50–65	0.087	0.101	0.081	0.020*	-	-	-
Age > 65	0.038	0.064	0.028	0.036***	-	-	-
Observations	3 796	999	2 797				

Notes: Column 1 shows baseline summary statistics for individuals in premium categories 1 and 2. Columns 2 and 3 present summary statistics for both premium groups separately, whereas column 4 shows the differences between the two groups for each variable. Columns 5 and 6 present the changes in individual and household characteristics between the pre- and post-policy time periods, for each premium group, whereas column 7 presents the difference in changes in characteristics between CBHI categories 1 and 2 —that is, the difference-in-difference estimate. Consumption in RwF 1,000 and travel time in hours. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5 Results

In this section, I estimate the price sensitivity of the demand for health insurance using the linear probability model with individual fixed effects. I also provide evidence of the robustness of the estimated price sensitivity to omitted variables bias by estimating the Oster approach and providing price sensitivity estimates based on balanced samples.

Finally, I use the estimated price sensitivity to predict insurance enrollment related to a number of plausible premium subsidy schemes.

Table 5 presents the results of estimating equation (1) using the complete sample including the full set of price variation. Each column shows the price sensitivity estimate, including different sets of covariates. Column 1 shows the unconditional estimate of the price sensitivity, including only a time indicator that controls for underlying time-varying factors that affect insurance enrollment and are common to all premium groups, and column 2 adds individual fixed effects. The estimated price elasticity more than doubles when I control for individual heterogeneity compared with the unconditional estimate. The results suggest that the individual heterogeneity is positively correlated with the price change, creating an upward bias of the unconditional estimate. This is likely to be explained by a higher underlying enrollment rate among wealthier households, which face higher premium costs. The following columns show that the estimated price sensitivity is robust to the inclusion of a number of covariates: Column 3 controls for individual labor status and includes an indicator that equals one if a household is situated in a rural household. In column 4, I additionally control for changes in individual health status by including an indicator for whether an individual has a disability or was sick during the two weeks prior to the interview. A control for household consumption is added in column 5, and the last column includes controls for any change in access to piped water and sanitation. The results suggest that the estimated price sensitivity is robust to the inclusion of covariates.

The preferred specification in column 6 suggests that an RwF 1,000 increase in the premium level is associated with a 10.9 ppt decrease in the likelihood of being enrolled in the CBHI. This is equivalent to a 15.2% decrease at the mean (0.717). The policy change resulted in an increase in premium levels by 83.7% on average, across the entire sample, resulting in a price elasticity of the demand for health insurance of –0.18. Taken together the results indicate that although health insurance coverage is sensitive to price change, the overall demand is price inelastic; that is, the elasticity is less than 1.

In appendix table A5, I test that the estimated price sensitivity is not driven by a single premium group by stepwise excluding premium groups in the fixed effects analysis. The results indicate that no single group is driving my results and the estimated price elasticity is relatively stable across the different samples (panel A). Furthermore, panels B—D replicate the results in table 5 for alternative assumptions of the premium scheme prior to the policy change as discussed earlier (see table 2 for further details). The results indicate that the estimated price sensitivity is robust to the different definitions of indigent households.

Importantly, although the fixed effects model controls for all time-invariant heterogeneity, there could still be time-varying heterogeneity that the model does not control for. As discussed in previous sections, this could potentially cause biased estimates. In the following section, I provide evidence of the robustness of the results to omitted variables based on the Oster approach and by providing price sensitivity estimates that use a number of balanced samples.

Table 5: Baseline results - price sensitivity of the demand of health insurance

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Premium (RwF 1,000)	-0.0339***	-0.107***	-0.108***	-0.108***	-0.109***	-0.109***
	(0.0105)	(0.0145)	(0.0145)	(0.0145)	(0.0145)	(0.0143)
Observations	7612	7612	7612	7612	7612	7612
R-squared	0.007	0.054	0.061	0.062	0.063	0.066
Number of PID		3806	3806	3806	3806	3806
FE	No	Yes	Yes	Yes	Yes	Yes
Basic covariates	No	No	Yes	Yes	Yes	Yes
Health status	No	No	No	Yes	Yes	Yes
HH wealth	No	No	No	No	Yes	Yes
Water & sanitation	No	No	No	No	No	Yes
Mean insurance enrollment	0.717	0.717	0.717	0.717	0.717	0.717

Notes: The results for the baseline linear probability regression with individual fixed effects corresponding to estimating equation (1) for the complete sample. The dependent variable is a dummy variable indicating enrollment in CBHI. Column 1 presents the unconditional price sensitivity; in column 2, individual fixed effects are added; and in columns 3—6, household and individual controls are added. Controls include individual and household characteristics, such as labor status and rural location, individual health status and household consumption, and access to water and sanitation services. Column 6 presents the preferred estimation strategy. Standard errors are clustered at household level. Standard errors clustered by household are shown in parentheses below the estimated coefficient. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.1 Heterogeneity

This section investigates heterogeneity in the price sensitivity of the demand for health insurance for individuals in different subsamples. Table 6 presents the price sensitivity estimated by age, gender, relation to household head and health risk.[12] Overall, the estimates are similar among groups. An increase in the premium level of RwF 1,000 is associated with an 8.82 ppt decrease in the enrollment among spouses, compared with 10.9 ppt among household heads. This corresponds to approximately 12% and 15%, respectively, at the mean (semi-elasticity).

Table 7 shows heterogeneity in price sensitivity among individuals living in households with different socioeconomic status and demographic composition. The results indicate that price sensitivity varies among households with different socioeconomic status, but that the demand for health insurance is inelastic in all groups.

An increase in premium levels by RwF 1,000 is associated with a decrease in the likelihood of being enrolled in the insurance by 17 ppt among individuals living in poor households, compared with a 9.13 ppt decrease among individuals in nonpoor households. This corresponds to a decrease of 27.3% and 11.6%, respectively in the likelihood of being enrolled at the mean, indicating that the price elasticity among poor households was more than twice as large as among nonpoor households. The results are in line with previous empirical research suggesting that the price elasticity of demand for preventive healthcare varies with socioeconomic status, and is higher among less wealthy and vulnerable households (Dupas, 2014). The estimated price sensitivity does not differ significantly between rural and urban households or between households headed by women and men. Individuals living in households headed by women, however, had a relatively high price elasticity of -0.257, compared with -0.149 among individuals residing in households headed by men.

[12] I estimate the health risk using a linear probability mode: $y_i = \beta_0 + X_i\beta + \epsilon_i$, where y_i is an indicator that equals 1 if an individual visited the clinic during the two weeks prior to the survey, and zero otherwise. X_i is a vector of explanatory variables including age, sex, income, disability, and access to water and sanitation. Individuals with a predicted risk of having been sick that exceeds 0.25 are defined as high risk, and those with a lower predicted likelihood are defined as low risk.

Table 6: Heterogeneity in the price sensitivity - among individuals

Variables	(1) All	(2) Head	(3) Spouse	(4) Children < 5 yrs	(5) Individual > 50 yrs	(6) Female	(7) Male	(8) High risk	(9) Low risk
Premium (RwF 1,000)	-0.109***	-0.109***	-0.0882***	-0.116***	-0.100***	-0.0991***	-0.121***	-0.102***	-0.111***
	(0.0143)	(0.0148)	(0.0203)	(0.0261)	(0.0221)	(0.0146)	(0.0176)	(0.0171)	(0.0157)
Observations	7612	1832	1193	1101	943	4034	3578	1780	5832
R-squared	0.066	0.073	0.065	0.063	0.103	0.059	0.077	0.079	0.069
Number of PID	3806	936	612	756	515	2017	1789	963	2989
Mean	0.717	0.734	0.752	0.690	0.742	0.721	0.712	0.730	0.713
Semi-elasticity	-0.152	-0.149	-0.117	-0.168	-0.135	-0.137	-0.170	-0.140	-0.156
Elasticity	-0.182	-0.182	-0.130	-0.160	-0.187	-0.162	-0.194	-0.176	-0.177

Notes: The results for the baseline linear probability model with individual fixed effects regression, corresponding to estimating equation (1) for the complete sample, including the full set of price variations (column 1), and for different subsamples of individuals (columns 1—9). All results are estimated using the preferred specification (table 5, column 6) and control for individual and household characteristics, such as labor status and rural location, individual health status and household consumption, and access to water and sanitation services. Standard errors are clustered at household level.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7: Heterogeneity in the price sensitivity—among households

Insurance	(1) Poor	(2) Nonpoor	(3) Rural	(4) Urban	(5) Female HH head	(6) Male HH head
Premium (RwF 1,000)	-0.170***	-0.0913***	-0.115***	-0.132**	-0.119***	-0.107***
	(0.0335)	(0.0175)	(0.0156)	(0.0512)	(0.0233)	(0.0198)
Observations	3224	4388	6622	990	1606	6006
R-squared	0.140	0.054	0.064	0.202	0.115	0.057
Number of PID	2,174	2,756	3,518	702	851	3,051
Mean	0.622	0.787	0.709	0.769	0.721	0.716
Semi-elasticity	-0.273	-0.116	-0.162	-0.172	-0.165	-0.149
Elasticity	-0.311	-0.140	-0.192	-0.177	-0.257	-0.164

Notes: The results for the baseline linear probability model with individual fixed effects regression, corresponding to estimating equation (1) for different subsamples of individuals (column 1—6). All results are estimated using the preferred specification (table 5, column 6), and control for individual and household characteristics, such as labor status and rural location, individual health status and household consumption, and access to water and sanitation services. Standard errors are clustered at household level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.2 Sensitivity analysis

In this section, I test the robustness of the price sensitivity estimates to omitted variable bias. First, I use the Oster approach to evaluate robustness to selection on unobserved factors. Second, I test the robustness of the estimated price sensitivity by providing estimates using a number of samples that are relatively more balanced on observable characteristics than the sample used for the principal estimations.

5.2.1 Bounds

The results presented in table 5, columns 2 and 3, show that the estimated price sensitivity remains stable after controlling for set of covariates, suggesting that the estimate is stable to the inclusion of observable characteristics. This has often been interpreted as an indication that omitted variable bias is limited (Oster, 2019). However, this interpretation builds on the assumption that the selection on observable factors is informative about the selection on unobservable characteristics.

Building on the work by Altonji et al. (2005), Oster (2019) proposes a framework that quantifies the robustness of an estimated coefficient to omitted variable bias. The approach uses movements in an estimated coefficient and corresponding R^2 values, when controlling for a set of observable covariates, to identify bounds of the estimated treatment effect. Furthermore, the Oster test provides a measure of how large the selection on unobservables has to be to erase the treatment effect.

The Oster test requires that I make assumptions about the relative degree of selection on observed and unobserved variables, δ, and the R^2_{max}. The R^2_{max} represents the maximum R^2 value that the regression can attain, including all observable and unobservable variables. Due to measurement errors and idiosyncratic variation, Oster (2019) argues that the R^2_{max} in many empirical settings is likely to be lower than 1. I follow Oster (2019) and set the R^2_{max} equal to 1.3*R^2. R^2 is measured for the preferred fixed effects regression that controls for all time-invariant individual and household characteristics, as well as time-varying observables (table 5, column 6, $R^2 = 0.066$).[13] I also provide calculations based on more conservative values of R^2_{max}, assuming values that are two and three times as large as the R^2 of the baseline regression. Additionally, I assume that the selection on observable and unobservable factors is proportional—that is, $\delta = 1$.

Conditional on the assumptions on δ and R^2_{max}, I compute bounds on the price-sensitivity coefficient. The lower bound is represented by the fixed effects estimate (table 5, column 2) that assumes no selection bias from unobservable variables, $\delta = 0$, whereas the upper bound is represented by the bias-adjusted coefficient, assuming that the selection on unobserved time-varying variables is at most as large as the selection on observed and unobserved time-invariant variables, $\delta = 1$. I control for individual fixed effects in all regressions including the baseline estimations, excluding all time-invariant characteristics from the confounding category. I do this because I consider it reasonable to assume that the selection on time-varying observables explains a proportional part of the selection in relation to unobservables ($\delta = 1$), excluding time-invariant characteristics.

Table 8 shows the estimated coefficient bounds. Based on the most conservative R^2_{max} value, the bounding set is estimated to [–0.136, –0.107]. The results suggest that the inclusion of controls increases the negative price sensitivity, moving the coefficient away from β=0. Oster (2019) proposes that the robustness of such coefficient can be tested by evaluating whether the coefficient bounds fall within +/–2.8 standard errors of the controlled coefficient estimate.[14] The results in table 8 show that all

[13]Oster (2019) determined this value based on a sample of randomized papers: 1.3*R^2 allowed for 90% of randomized results to survive.

[14]For bias-adjusted coefficients moving toward zero, $\beta = 0$, a coefficient bound that does not include zero is considered robust.

estimated bounds related to the different R^2_{max} assumptions fall within the defined interval, indicating that the size of the price sensitivity estimated by the baseline fixed effects regression is similar to the bias-adjusted estimate. Importantly, calculations of the price elasticity based on the bias-adjusted price sensitivity, $\beta = 0.136$, show that the estimated price elasticity is robust to potential omitted variable bias and remains negative and inelastic—that is, adjusting the estimated price sensitivity for omitted variable bias would not change the conclusions presented in the analysis.

The second test provided by the Oster approach measures how large the selection on unobservables has to be, in relation to the selection on observable variables, to erase the estimated treatment effect. The results suggest that the selection on unobservables needs to be at least 7.4 times stronger than the selection on observable factors ($\delta = -7.401$) for the estimated price elasticity to switch sign ($\beta = 0$). $\delta < 0$ suggests that if the observable covariates in the baseline regression are positively correlated with the price change, the omitted variables would have to be negatively correlated with the price change in order for the price sensitivity to switch sign. The estimated price sensitivity from table 5 shows that in the main sample, adding in controls actually increased the negative price sensitivity, implying that observable household and individual characteristics are positively correlated with the price of insurance enrollment. As previously discussed, one of the main sources of bias could arise if the Rwandan government were to introduce policy initiatives that aim to improve the health of individuals in the lower Ubudehe categories. This would create a negative correlation between improved individual health status and price change. In this context, selection on observables as well as unobservables result in an underestimated price sensitivity. The degree of selection on unobservables relative to observables that would be necessary to explain away the price sensitivity, indicates that the estimated price sensitivity is robust to omitted variable bias (Oster, 2019).

Table 8: Sensitivity analysis: selection on unobservables

	(1) Uncontrolled	(2) Controlled
Coefficient	-0.107***	-0.109***
	(0.0145)	(0.0143)
R^2	0.054	0.066
A. Bounds on the treatment effect		[-0.1123, -0.107]
(δ=1,R^2_{max}=0.0.086)		
δ for β=0 (R^2_{max}=0.086)		-27.136
δ for β=-0.600 (R^2_{max}=0.086)		3.067
B. Bounds on the treatment effect		[-0.122, -0.107]
(δ=1,R^2_{max}=0.132)		
δ for β=0 (R^2_{max}=0.132)		-12.951
δ for β=-0.600 (R^2_{max}=0.132)		2.787
C. Bounds on the treatment effect		[-0.136, -0.107]
(δ=1,R^2_{max}=0.198)		
δ for β=0 (R^2_{max}=0.198)		-7.401
δ for β=-0.600 (R^2_{max}=0.198)		2.463

Notes: The outcome variable is a dummy variable taking the number 1 if an individual is enrolled in the CBHI scheme, and zero otherwise. The uncontrolled estimation includes a time dummy, as well as individual fixed effects (table 5, column 2). The controlled treatment effect is calculated from the fixed-effects regression as reported in table 5, column 6. Calculations in panel A follow Oster (2017), using R^2_{max}= 1.3*R^2; panel B assumes an R^2_{max} that is twice as large as the value from the baseline fixed effects specification; and panel C assumes an R^2_{max}=0.198, three times as large as the one related to the baseline specification. Furthermore, the calculations are based on the assumption that the selection on unobservables is proportional to the selection on observables, δ=1. The last row in each panel present the δ value associated with a price elasticity larger than -1. All calculations are made using the psacalc Stata code by Oster (2016).

A potentially even more relevant measure is the robustness of the conclusions drawn in the analysis. In the previous section, the baseline fixed-effects estimation strategy indicates that the price elasticity for the demand of CBHI in Rwanda is negative and inelastic. As mentioned earlier, a price elasticity is considered inelastic if the elasticity is less than $|1|$. Using the Oster approach, I test the robustness of the estimated price elasticity to this conclusion by estimating how large the selection on unobservables would need to be to result in an elastic price elasticity ($\beta < -0.600$). The results in table 8 show that the selection on unobservables needs to be at least 2.5 times stronger than the selection on observable, as well as all time-invariant, factors (δ=2.463) for the demand of CBHI to become elastic in relation to the most conservative assumption on R^2_{max}. The results indicate that the conclusion that the demand for CBHI in Rwanda is price inelastic is robust to omitted variable bias (Oster, 2019). In summary, the results suggest that the estimated price elasticity is not driven by omitted variable bias.

5.2.2 Balanced samples

In this section, I aim to limit the correlation between the policy change and household socioeconomic status, and thereby obtain a treatment that is closer to being independent of the background of covariates. I do this by increasing the balance of the sample. This will reduce the potential for bias (Ho et al., 2007).

First, I limit the sample to include only individuals in Ubudehe groups 2 and 3. These two groups received very different insurance premiums after the policy intervention, but had similar socioeconomic status according to the Ubudehe categorization system. Second, I use propensity scores to create a more balanced sample in relation to individual and household characteristics. Based on the sample limited to households in Ubudehe group 2 and 3, I use a logistic regression to predict the likelihood of an individual being classified in Ubudehe group 3—that is, of having received an increased insurance premium. Based on these predictions, I construct one sample including all individuals in the common support. I limit this sample further by including only individuals with predicted likelihoods between the 10th and 90th percentiles, as well as the 20th to 80th percentiles. The adjusted samples ensure that there is overlap in the distribution of covariates for all observations in the sample; that is, the estimations require no extrapolation to cells without common support (Angrist & Pischke, 2008). By constructing the sample based on the likelihood of being in Ubudehe group 3, I ensure that there will always be a few observations in Ubudehe group 2 that can be used to estimate the counterfactual.

Appendix Figure A3 illustrates the distribution of the predicted likelihood of individuals in Ubudehe groups 2 and 3 being categorized in Ubudehe group 3. The distribution shows a significant overlap between the treatment and control groups, resulting in a wide region of common support. The distributions lack common support only at the low extreme of the distribution, showing that some individuals in the control group are substantially different from individuals in the treatment group and consequently have a significantly lower likelihood of being categorized in Ubudehe group 3. Appendix figure A4 suggests a need for further adjustment in order to obtain a balanced sample with similar baseline characteristics.

Appendix table A6 reports the estimates from the logistic regressions. In addition to the household and individual characteristics included in the baseline estimations, these predictions include a number of observable factors, defined by the Rwandan government, that are used to classify households into specific Ubudehe categories: ownership of the house and livestock, and household consumption (appendix table A1). Additionally, the estimations include an indicator for whether the house has improved floor materials, such as wood or concrete. This variable is a proxy for quality of housing, one of the specific factors in the Ubudehe classification process. The results show that the number of household and individual characteristics that are significantly correlated with the Ubudehe classification decrease as the sample gets more restricted. The results suggest that the significant differences between individuals in Ubudehe groups 2 and 3 decrease, indicating that the balance in covariates between the groups improves between the different samples. Furthermore, the predictions confirm that a number of assessment factors of the classification process have been important in defining the likelihood of being defined as Ubudehe group 3. The results also suggest that relatively old and very young individuals (those over 65 years old and younger than 5), as well as salary workers, were less likely to be categorized in Ubudehe group 3 across all specifications. Access to sanitation services and household size are positively associated with the likelihood of being categorized in the higher Ubudehe category.

The results of the estimations using the balanced samples are presented in appendix table A7. The estimates in column 1 are based on a sample restricted to include individuals in Ubudehe groups 2 and 3, excluding the households that received full subsidies prior to the policy change, as explained earlier. Column 2 restricts the sample to include only individuals with a predicted likelihood of being treated within the common support, whereas columns 3 and 4 restrict the sample even further by including those with a propensity score within the common support between the 10th and 90th percenties (column 3) (Crump et al., 2006) and the 20th and 80th percentiles (column 4). The results indicate that the estimated price sensitivity is stable between the different samples, suggesting that the main estimates of price sensitivity are not likely to suffer greatly from bias caused by different distributions in covariates between the different premium categories.

5.3 Predicted take-up levels

In this section, I use the price sensitivity estimates to predict insurance take-up related to a number of counterfactual premium subsidies. I make a linear prediction of estimated take-up levels, taking into account the individual-specific fixed effects. The predictions are calculated based on the coefficients estimated by equation (1) using data from 2014. Table 9 shows the overall predicted take-up level for different premium structures, referring to both uniform premium subsidies and premiums targeted to specific subgroups. Overall, the results indicate that the variation in predicted insurance coverage among different premium subsidy schemes is limited. This is a direct effect of the low price sensitivity.

Column 1 in table 9 presents the predicted average take-up level based on the estimated price sensitivity from the preferred specification, presented in column 6 in table 5. The first row in table 9 shows the predicted insurance coverage in 2014 for the current subsidy scheme. The total take-up level is predicted to approximately 70%, which corresponds to the take-up level found in the EICV data. The succeeding rows present the average enrollment levels for alternative subsidy structures. The results suggest that the overall enrollment rate would reach approximately 82% in a scenario in which all individuals receive fully subsidized premiums. This is consistent with previous literature that indicates that full premium subsidies are not sufficient on their own to induce universal insurance coverage (Finkelstein et al., 2019; Wagstaff et al., 2016; Thornton et al., 2010). A uniform premium cost of RwF 3,000 results in a predicted take-up of 66%.

Table 9: Predicted insurance coverage, 2014

Premium structure	(1) All	(2) poor	(3) Nonpoor	(4) Children age < 6 yrs.	(5) Senior age > 50 yrs.	(6) Low risk	(7) High risk
Actual premium	0.698	0.588	0.775	0.672	0.738	0.695	0.717
0	0.819	0.694	0.857	0.803	0.821	0.822	0.809
1000	0.765	0.635	0.824	0.746	0.779	0.766	0.764
2000	0.710	0.576	0.792	0.689	0.737	0.710	0.718
3000	0.656	0.517	0.760	0.632	0.694	0.655	0.672
Avg. cost (3543)	0.626	0.485	0.742	0.601	0.671	0.625	0.647
Children <6 yrs free)	0.671	0.533	0.769	0.803	0.694	0.675	0.672
Minors (<19 yrs free)	0.732	0.616	0.800	0.803	0.694	0.759	0.673
Poor households free	0.721	0.694	0.760	0.707	0.732	0.723	0.723

Note: The predictions are based on the coefficients resulting from estimating equation (1), presented in table 5, column 6, across different population samples presented in tables 6 and 7. I use data from EICV4 (2014) to predict take-up levels.

The final bottom three rows in table 9 predict take-up levels of targeted subsidies offering full premium subsidies to poor households (based on household consumption level) and to households with children under 19 and 6 years. In all three targeted subsidy strategies, households that are not subsidized face an individual premium of RwF 3,000. Overall, the simulations suggest that changes to the premium subsidy scheme do not show a large impact on coverage levels. This is consistent with the overall low price sensitivity. Furthermore, one plausible subsidy scenario is represented by an actuarial premium scheme in which the insurance premium is set to average insurer costs—that is, RwF 3543 in this context. The premium cost resulting from pricing at average cost is high compared with the other subsidy schemes, resulting in relatively low insurance take-up.

Besides an overall increase in access to healthcare, equal access to healthcare represents another policy priority for governments in relation to universal health coverage. In this context, it is important to consider the impact of premium subsidy schemes on insurance enrollment in specific subgroups. Columns 2—7 in table 9 present the effects of alternative premium subsidies on take-up for a number of subgroups. Considering the heterogeneity in price sensitivity among households with different socioeconomic status presented in tables 6 and 7, it is expected that the predicted take-up of insurance for alternative premium subsidy schemes will vary among different socioeconomic groups. A premium scheme that targets consumption-poor households (poor in this case refers to the national poverty line based on household consumption, not to the Ubudehe category) predicts higher insurance take-up among individuals in this group in comparison with the current premium scheme that targets households with low socioeconomic status. As previously mentioned, the simulations predict that the overall take-up level related to the actual premium subsidy is approximately 70%, whereas take-up among individuals living in consumption-poor households reaches only 59% under the same subsidy scheme (column 2, row 1). Furthermore, the results indicate that a premium subsidy based on a monetary measure of poverty could increase overall coverage levels and, more importantly increase access to health insurance among the consumption-poor. At the same time, the simulations indicate that the coverage level among children will decrease with this subsidy scheme compared with the schemes that specifically target children.

Young children represent another potential group of interest for policymakers. The simulations indicate that targeted subsidies based on age will bring take-up levels to around 80% among young children. Note that take-up levels among poor individuals are low in relation to the age-based subsidy schemes. A premium structure that targets children 5 years and younger is predicted to contribute to low take-up levels of approximately 53% among individuals living in financially poor households. On the other hand, enrollment rates among the youngest remain relatively high when subsidies target the consumption-poor.

There is not much difference in take-up levels between individuals with predicted high health risk and those with low risk. This is important knowledge when considering the governmental goal of universal healthcare coverage. The similarities in take-up levels between high- and low-risk individuals indicate that individual health status is not likely to affect the decision to enroll in the insurance at a given premium cost and that the health status of enrolled individuals is stable across premium subsidy schemes; that is, changes to the premium costs do not cause selection into the insurance scheme based on health status. This reasoning is further supported by the fact that the price sensitivity of the demand for health insurance is similar between the two groups (see table 6, columns 8 and 9). I further evaluate the potential effects of adverse selection in the following section.

Put together, heterogeneity in price sensitivity among different subgroups will contribute to a variation in the composition of enrolled individuals. This is important because it provides the government with some potential to target its efforts and increase take-up within different vulnerable subgroups. This is in line with the current subsidy scheme that targets households with low socioeconomic status.

6 Financial Self-Sustainability

In this section, I will consider the importance of premium subsidies to the financial self-sustainability of the CBHI scheme. Premium subsidies represent an increasingly common policy tool to promote take-up of health insurance. However, heavily subsidized insurance premiums limit the financial self-sustainability of the insurance scheme as premiums cover only a share of total patient costs. Furthermore, the results presented in this analysis suggest that the overall effect of premium subsidies on insurance coverage might be limited. At the same time, heterogeneity in the price elasticity among different subgroups suggests that changes to the premium scheme might affect the composition of beneficiaries, which could result in changes in the demand for healthcare within the insurance scheme. The selection of individuals into the insurance at different premium costs could create an association between insurance premium and patient costs. In a market with adverse selection, changes to insurance premium costs will affect the cost of providing the insurance, as individuals select into the insurance based on their expected demand for healthcare (Einav et al., 2010). As a result, in order to evaluate the importance of premium subsidies on the financial self-sustainability of the CBHI scheme, it is necessary to consider this potential association between insurance premiums and patient costs.

6.1 Insurer costs and selection

To evaluate the importance of a change in insurance premiums to the financial self-sustainability of the CBHI scheme, I consider the potential impacts of adverse selection on the cost of insurance. To do this, I follow an empirical model proposed by Einav and colleagues (2010). This strategy uses the insurer cost curve to identify and quantify adverse selection. In the presence of adverse selection, individuals with the highest expected patient costs are those who have the highest willingness to pay for the insurance, resulting in a marginal cost curve that is increasing in price. As the premium cost increases, relatively healthier individuals leave the insurance, driving up the average insurer cost among beneficiaries.

In this context, rejecting the null hypothesis of a flat marginal cost curve is evidence for selection. Furthermore, the slope of the marginal cost curve of the insurance provides a test for the direction of the selection: a positive slope of the cost curve is evidence of adverse selection (Einav et al., 2010). This empirical strategy has been widely used in the context of high income countries (Einav et al., 2010) and makes it possible to measure the financial implications of adverse selection in health insurance markets.

I estimate the association between insurer costs and insurance premiums using a unique data set provided by the Ministry of Health in Rwanda. The data provides a register of operational costs, as well as the cost of providing medical coverage for all individuals enrolled in the CBHI scheme. The cost data are provided at the sector level, describing the total insurer cost for each sector in the country. A section is an administrative entity that approximates the catchment area of a health clinic. There are 416 sectors in Rwanda distributed in 30 districts (Ministry of Health, 2012). After being adjusted for missing information regarding all or some expenditures, the cost data represent a sample of 295 sectors, representing approximately 71% of the total population of administrative sectors. Appendix table A8 shows that there is no significant difference between missing sectors and the those included in the sample in relation to a number of sector characteristics, indicating that my sample constitutes a representative sample of sectors and the missing information is not likely to affect the validity of the results.

The total insurer costs include a number of expenditures: cost of health consultations and hospitalization, operational costs, and reimbursement to health clinics and to the district and national risk pools in order to cover hospital claims. Table 10 describes the average cost per beneficiary related to a number of expenditures. Reimbursements to health clinics for their services represent the largest expenditures within the CBHI, followed by payments to the district risk pool and operational costs. On average, an individual enrolled in the CBHI scheme generates a cost of RwF 3543 during one year (approximately USD 4). This cost is almost equivalent to the premium cost of RwF 3,000 paid by individuals in premium category 2. The average cost gives an indication of the level of subsidy provided to households in premium category 1, whose premiums are completely subsidized.

Table 10: Insurer costs, avg. 2013–14

Expenditure	N	Mean	Std. dev.	Min	Max
Health consultations	295	884.60	354.41	180.08	2315.54
Hospitalization	295	98.11	86.23	1.14	579.49
Reimbursements	295	1321.11	464.967	393.63	3951.68
District risk pool	295	1103.78	398.81	418.61	2846.68
Operational costs	295	134.86	85.06	24.19	621.87
Total costs	295	3542.45	815.18	1943.59	6839.57

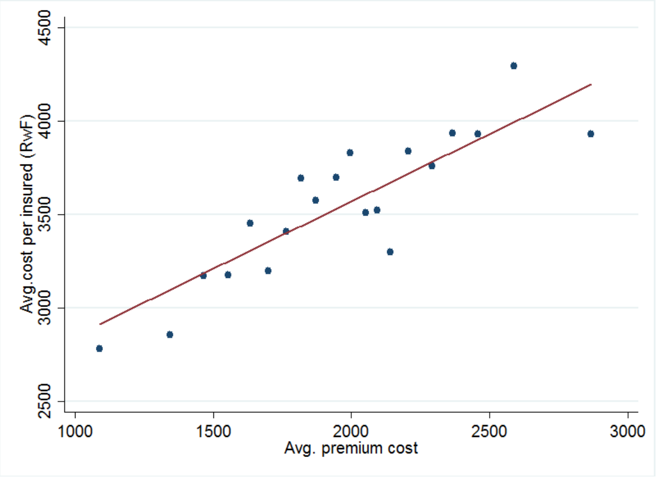
Notes: The total insurer costs includes reimbursement to health centers for health consultations and hospitalizations, as well as other costs related to the care provided at health centers, contributions to the district risk pool, and overall operational costs.

6.2 Estimating the cost curve

To estimate the association between insurer costs and insurance premiums, I use the measure of average total patient costs described in table 10 and a measure of the average insurance premium in each sector. I use the composition of beneficiary households across premium categories in each sector to construct the average premium. Given this construction, the premium is calculated as the average premium cost paid by enrolled individuals in one section. As a result, sections with a higher share of completely subsidized households (Ubudehe groupd 1 and 2) will have a lower average premium in relation to sections with a relatively larger share of households in premium categories 2 and 3.

Figure 2 illustrates how average insurer cost among enrolled individuals covaries with the average premium cost in each section. The figure shows average patient costs by bins of 5% of the average premium cost per section. The positive slope of the curve is consistent with adverse selection (Einav et al., 2010): as premium costs increase, relatively healthier individuals drop out of the insurance, driving up the average insurer cost.

Figure 2: Insurer costs and enrollment



Notes: Average total patient costs for individuals enrolled in the CBHI scheme during 2013–2014, showing the rough correlation between the average insurer costs and the average premium cost in each administrative section. The average costs are calculated as an average of the total costs or medical consultations, hospitalizations, and reimbursement for services at health clinic and district hospitals , as well as administrative costs (presented in table 10). The average premium cost represents the average insurance premium paid by enrolled individuals in each section. The dots represent the average insurer cost per enrolled individual by bins of 5% of the average premium cost per section.

Using the variation in average insurance premium and patient costs across sectors, I estimate the cost curve presented in equation (2). Given the seemingly linear association between insurance premiums and costs shown in figure 2, I estimate a linear model assuming that insurer costs are linear in price. Importantly, the cost equation is estimated based on patient costs among insured individuals. As a result, the insurer costs are affected only by insured individuals who already selected into the insurance scheme. Furthermore, changes in patient costs among sections will therefore not emerge as a result of differences in insurance coverage, since all individuals face the same insurance contract.

For the linear cost curve, the marginal cost curve can be derived by MC'=2*AC'.

$$c_k = \delta_0 + \delta_1 p_{jk}^{CBHI} + \boldsymbol{X}_{jk}\delta_3 + \gamma_k + \sigma_{ijk} \tag{2}$$

where c_k is the average total cost of providing the insurance per sector k, including medical consultations, hospitalizations, medicine and ambulance use, as well as administrative costs. The average cost is calculated per sector. δ_1 measures the association between insurance premium and average insurer costs, and X_k is a vector of sector characteristics. According to Einav et al. (2010), the sign of δ_1 is informative about the presence and nature of selection in the health insurance market. A positive relationship between individual insurer cost and the premium indicates adverse selection as individuals enter and exit the market endogenously as a result of the price change. In this context, individuals with relatively better health status exit the insurance scheme as premium levels increase and exceed their expected healthcare expenditure (Einav et al., 2010). γ_k represents district fixed effects, controlling for systematic differences among administrative districts.

Table 11, shows the results from estimating equation (2) using the cost data described in table 10. Column 1 presents the unconditional estimate of the association between premium and patient costs. In column 2, I add district fixed effect to make sure that the price effect does not pick up only underlying differences among districts. All health sections pool resources in a district risk pool to cover the costs of district hospitals. Consequently, it is likely that the financial setup in different districts could influence the cost structures of the sectors in each district. There are 28 administrative districts in the data with an average of 6 health sectors per unit.

Column 3 adds a number of sector-specific covariates that can explain variation in insurer costs such as the level of access to water and sanitation services, urbanization, age composition, and average household consumption. The results suggest that an increase in premium costs has a positive impact on average insurer costs. An increase of RwF 1,000 in the average premium level is associated with an increase in average insurer costs by RwF 674, approximately 20% at mean. The positive slope of the cost curve is consistent with individuals adversely selecting into the insurance scheme since the average cost increase among beneficiaries as the premium increases. The results are stable to the inclusion of district covariates, whereas district fixed effects show some effect on the estimates.

Table 11: Average insurer costs

	(1)	(2)	(3)
Average insurer costs			
Avg. premium (RwF1,000)	0.720***	0.563***	0.572***
	(0.126)	(0.151)	(0.153)
Observations	295	295	295
R-squared	0.148	0.452	0.460
District FE	No	Yes	Yes

Notes: the table shows the results from estimating equation (2) using the cost data described in table 10. Column 1 provides the unconditional association between average premium and patient costs across sectors, whereas column 2 include district fixed-effects and column 3 additionally controls for sector characteristics. When included, the controls contain share of section households with piped water, access to sanitation, share of households in urban areas, share of children below 5 years old and individuals older than 65, total population size and average household consumption. The estimated correlation between the average premium and the average insurer cost is the slope of the cost curve. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

According to the framework, the identification of adverse selection requires that the variation in premium costs be unrelated to the underlying health status and expenditure of the population in different sectors. However, due to the construction of the average premium cost, there is a possibility that the positive correlation between premium costs and insurer costs mirrors underlying differences in the distribution of households across categories of socioeconomic status. The results show that the average patient cost is higher in sections with a larger share of households with relatively higher socioeconomic status than in to sectors with a larger share of enrolled households that are worse off. This positive slope would appear if individuals, on average, in households with higher socioeconomic status had a greater ability than others to negotiate and demand an increasing number of expensive healthcare treatments, even in the absence of adverse selection. Following the same reasoning, underlying cultural differences in the use of healthcare services between households with high and low socioeconomic status could cause a similar outcome. Both scenarios would result in an upward bias in the association between insurance premiums and insurer costs presented in table 11. Given this reasoning, the effect of changes in insurance premiums is likely to have a more limited effect on patient costs than what is proposed in this analysis. I adjust for this bias in the following analysis of the financial sustainability of the insurance scheme by providing results that assume no presence of adverse selection. The overall results are not undermined by the potential upward-biased estimation of adverse selection.

Differences in operational costs among geographic regions represent another potential cause of bias if individuals in different Ubudehe categories were unevenly distributed between rural and urban areas. Sectors with a higher share of urban residents and clinics could face higher operational costs than rural regions due to relatively higher salaries, rent and materials in urban areas. If high-paying insurance beneficiaries are concentrated in urban sectors, this variation in operational costs could result in an upward bias of the association between insurer costs and premiums. Conversely, health facilities in larger urban areas could face cost advantages as a result of increases in their scale of production. In this context, the economies of scale would bias the cost curve downwards.

I control for potential variation in operational costs between rural and urban areas by controlling for the share of individuals living in urban areas in each CBHI section. Additionally, to further rule out that variation in operational costs is driving the variation in average patient cost, I provide estimates of the cost curve excluding the operational costs of the insurer. The results are presented in appendix table A9 and indicate that the positive association between average premium costs and average insurer costs hold when operational costs are excluded from the analysis and is robust to the alternative definition of the total cost.

In general, any factor that simultaneously predicts patient costs and socioeconomic status, and consequently premium cost, could bias the cost curve. As discussed earlier, the relation between health and socioeconomic status has been well documented in previous literature (Cutler et al., 2008). In this context, sections with a higher share of completely subsidized beneficiaries could have higher average insurer costs primarily due to low health status and not as a result of adverse selection. Importantly, the results in figure 2 suggest that this mechanism is not driving the association between insurance premium and insurer costs. However, I cannot rule out that the association between premiums and health status could lead to an underestimation of the association between premium costs and insurance costs, leading to a lower bound of adverse selection. I control for differences in average health status by including covariates that control for the share of population with access to piped water and sanitation in each sector, as well as the share of elderly (table 11).

In conclusion, the administrative cost data in this analysis do not allow me to further investigate and rule out potential drivers of the correlation between patient and premium costs. Importantly, the underlying reasons for changes in the insurer cost are not crucial for estimations of the financial sustainability of the insurance scheme. The estimated association between insurance premiums and patient costs implies that changes in premium costs are associated with modest increases in patient costs. Whether the positive association between cost and premium is caused by adverse selection or other factors does not change the conclusions regarding the potential role of a correlation between premium subsidies and patient costs in relation to financial sustainability. This should be considered when reasoning about the financial sustainability of the CBHI scheme as well as premium subsidies in general.

6.3 Simulations

In this section, I estimate the financial self-sustainability of the CBHI scheme in relation to the different premium subsidy schemes previously described. Self-sustainability is defined as the share of insurer costs covered by member premiums. I first use the predicted insurance coverage levels presented in table 9 to estimate the premium revenues corresponding to the different subsidy schemes. I then consider the effect of the premium schemes on insurer costs. I use the estimated cost sensitivity in table 11, column 3, to predict the average insurer cost corresponding to each subsidy scheme. By combining the predicted revenue and insurer cost, I estimate the level of self-financing that corresponds to each subsidy scenario.

Table 12 presents the simulated financial self-sustainability of the CBHI in Rwanda related to the different subsidy strategies discussed in the previous section. Column 1 presents the average predicted enrollment for all subsidy schemes presented in table 9, column 1. Column 2 shows the premium revenues that correspond to each subsidy scheme. The premium revenues follow directly from the take-up level and the level of premium subsidies - as premium costs increase, enrollment levels decrease. Consequently, the effect of subsidies on insurer revenue is an empirical question that will depend on the price elasticity of demand for health insurance.

The second part of the table focuses on the implications of changes in premium costs on insurer costs and the financial self-sufficiency of the insurance scheme. Financial self-sustainability is estimated as the share of the total costs covered by the premium incomes. By definition, this measure of financial sustainability directly hinges on the premium income and the average individual insurer cost predicted by the different subsidy schemes. The results are presented for two scenarios. In one scenario, I assume that there is no adverse selection in the CBHI market and estimate the financial coverage level using a constant average patient cost, independent of the premium level.[15] The second scenario predicts the average patient cost related to each subsidy scheme, using the association between patient costs and insurance premiums estimated by equation (2). The estimates are presented in table 11, column 3. I use the estimated price sensitivity to predict the relationship between average premium costs and average patient costs at the health section level. The positive slope of the average cost curve is reflected in the levels of average patient cost related to each premium scheme: as the average premium cost increases, the average insurer cost increases.

The simulations indicate that the financial sustainability of alternative premium subsidies differs depending on whether there is adverse selection. This is expected. In a setting with adverse selection, the financial coverage reaches levels between 0.35–0.80 for the majority of subsidy schemes, meaning that household premiums cover approximately 35%—80% of the insurer costs. However, in the absence of selection, the range of coverage levels for the corresponding subsidy schemes are wider, suggesting that the coverage levels range between 0.28 and 0.85. The wedge between the level of financial coverage in the selection and no-selection scenarios increases as the premium levels deviate from the mean cost.

Overall, the results suggest that the effects of selection are limited but may be important from the insurers' point of view. The results are in line with previous results from developed countries that indicate that the cost of adverse selection (mainly in terms of social welfare) might be relatively negligible (Finkelstein et al., 2019).

Previous studies on adverse selection (Parmar et al., 2012) indicate that targeted subsidy schemes are associated with increased adverse selection. Unfortunately, I lack access to the cost data necessary to perform this analysis. The cost calculations related to the targeted premiums are calculated based on the average cost among all individuals at each premium level. The calculations provide an indication of the patient costs based on sample average but do not consider variation in patient costs among individuals with different socioeconomic and demographic characteristic.

[15] The average cost is calculated as a raw average of the administrative cost data provided by the Rwandan government.

Table 12: Financial sustainability and insurance coverage: alternative subsidy schemes

Premium subsidy scheme	(1) Predicted coverage	(2) Premium payments	No selection (3) Avg.cost	(4) Financial Coverage	Selection (5) Avg.cost	(6) Financial Coverage
Actual premium (RwF 2223)	0.698	$1.37e10$	3543	0.593	3585	0.586
0	0.819	0.00	3543	0.000	2421	0.000
1000	0.765	$6.98e09$	3543	0.282	2993	0.347
2000	0.710	$1.30e10$	3543	0.564	3565	0.562
3000	0.656	$1.79e10$	3543	0.847	4137	0.725
Avg. cost = 3543	0.627	$2.02e10$	3543	1.000	4434	0.799
Children (0–5 yrs free)	0.671	$1.61e10$	3543	0.737	3849	0.679
Minors (<19 yrs free)	0.732	$9.14e09$	3543	0.403	3238	0.447
Poor households free	0.721	$1.22e10$	3543	0.533	3896	0.485

Notes: The leftmost column listss the different premium subsidy schemes that are used in the estimations of sustainable financing. Columns 1 and 2 show the predicted coverage levels and the corresponding premium income generated by each subsidy scheme (see table 9 for further information on the predicted take-up levels). Columns 3–6 present the estimated average patient cost and the resulting financial coverage level related to each scheme. The estimations are provided for context with and without selection. The average patient costs have been predicted using the cost curve estimations given in table 11, column 3.

7 Conclusions

Over the last two decades, governments in many developing countries have taken important measures to achieve universal health coverage. In this study, I examine the potential of premium subsidies as a policy instrument to reach this goal, considering the financial sustainability of insurance schemes. The effect of premium subsidies is directly dependent on the price sensitivity of the demand for health insurance. I study the effects of premium subsidies on the take-up of the CBHI scheme in Rwanda, using a policy change in the insurance premium to identify the effect of a price change on the demand for health insurance. The results suggest that the demand for CBHI is price inelastic. I find that an increase in premium costs by RwF 1,000 (USD 1.1) contributes to an overall decrease in the likelihood of being enrolled by 10.9 ppt (15.2% at the mean). This translates into a price sensitivity of –0.18.

Furthermore, the effect of changes to the premium costs is heterogeneous among different subgroups of households. Individuals living in poor households or households headed by women have a higher price elasticity than individuals in nonpoor or male-headed households. The results suggest that premium subsidies affect the composition of individuals who decide to enroll.

I use the estimated price sensitivity to simulate the take-up level, insurer costs, and financial sustainability of alternative subsidy schemes. Overall, the results indicate that government subsidy strategies will have a limited effect on insurance coverage. This is a direct effect of the inelastic demand. The results suggest that the current premium scheme achieves a relatively high coverage level compared with the other counterfactual subsidy strategies. However, when analyzing the predicted coverage by subgroups, I find that the evidence indicates that there is great variation in enrollment. Although one of the primary aims of the new premium policy in Rwanda was to increase equity in access to healthcare (Kalisa et al., 2016), the results suggest that in comparison with alternative premium structures, the current premium scheme does not achieve high enrollment levels among poor households. According to the simulations, the implementation of a subsidy scheme that targets monetary poor households not only would contribute to an increase in overall insurance coverage but also would increase insurance take-up among vulnerable individuals living in poor households and among the youngest children. As a result, a premium scheme that targets monetary poor households is likely to increase equality in the access to healthcare. A poverty definition that is based on monetary measures differs from the unbudehe classification system that is based on a more integral measure of household socioeconomic status. However, it is important to note that the CBHI scheme primarily targets households in the informal sector. In this context, classification of households according to income could be problematic, implying that the implementation of this premium subsidy policy might not be feasible from a practical point of view, resulting in an inefficient targeting instrument. It is not obvious which method provides a better and more efficient poverty measure.

Another important aim of the Rwandan policy reform was to increase financial sustainability of the insurance scheme by increasing revenue from household premium payments (Kalisa et al., 2016). I simulate the financial coverage related to the different pricing strategies by calculating the share of insurer costs covered by premium revenue. Importantly, in the presence of selection, changes to the insurance premiums will affect the cost of providing the insurance, as the composition of insurance beneficiaries changes in response to the change in premium costs. I use variation in aggregate premium costs among administrative sectors to identify the average insurer cost curve. The results show a strong correlation between premium costs and insurer costs, consistent with adverse selection. The presence of adverse selection is important for financial sustainability of the insurance scheme. As the premium costs increase, so does the cost of providing insurance.

I predict the financial coverage, or financial self-sustainability, levels of subsidy schemes in both a setting that allows for adverse selection and one that assumes no selection. The results indicate that the financial effects of adverse selection are limited in relation to many subsidy schemes. At the same time, selection is likely to contribute to unsustainable insurance schemes: as premium costs increase, so does the cost of providing insurance coverage among beneficiaries. Furthermore, in regard to the aim of reaching universal healthcare coverage, the take-up levels related to relatively high premium costs are far from universal. In the end, a health insurance scheme financed by household premiums is not likely to represent a financially sustainable strategy to reach universal health coverage. From another perspective, it is important to emphasize that as enrollment levels start to increase in markets with adverse selection, the cost of enrolling another individual is smaller relative to the cost of individuals already enrolled in the insurance. This represents an important consideration when governments make decisions on introducing premium subsidies.

The results in this study indicate that insurance premiums are not likely to promote universal insurance coverage without the support of external funds. Dependence on external donors represents a threat to the overall sustainability of the insurance schemes, as external funding might fluctuate as a result of cuts in aid budgets, so in the long run, governments should find other potential sources of funding. However, because of the small tax base in countries with extended informal sections, the limitations of tax-based financing are obvious in the short run. Nevertheless, a number of countries have decided to put legal tax commitments towards financing the expansion of national social health insurance schemes. Premium subsidies can represent a good alternative to expand CBHI in the short run.

The question of how to finance the expansion of health insurance is important for the long-run sustainability of health financing in developing countries. The financial implications of selection in the health market should also be evaluated in the context of public health and welfare. From a government point of view, the expansion of health insurance coverage is likely to improve the health and well-being of the population, factors that should be considered by policymakers when conducting a complete assessment of the value of premium subsidies. From this point of view, increased patient costs could for example imply that poor and vulnerable households with high needs for medical care were able to enroll in the insurance scheme and access healthcare services. Ultimately, it is important to recognize that this study is limited to evaluating the impact of premium subsidies for health insurance on insurance take-up and from a financial perspective.

References

Aggarwal, A. (2010). Impact evaluation of India's Yeshasvini community-based health insurance programme. *Health Economics*, *19*(S1), 5–35.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., & Tobias, J. (2012). Targeting the poor: Evidence from a field experiment in indonesia. *American Economic Review*, *102*(4), 1206–40.

Alkenbrack, S., Jacobs, B., & Lindelow, M. (2013). Achieving universal health coverage through voluntary insurance: What can we learn from the experience of Lao PDR? *BMC Health Services Research*, *13*(1), 521.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, *113*(1), 151–84.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ, USA: Princeton University Press.

Asuming, P. O., Kim, H. B., & Sim, A. (2017). *Long-run consequences of health insurance promotion: Evidence from a field experiment in Ghana* (IZA Discussion Paper No. 11117). Bonn, Germany: IZA Institute of Labor Economics.

Banerjee, A., Duflo, E., & Hornbeck, R. (2014). Bundling health insurance and microfinance in india: There cannot be adverse selection if there is no demand. *American Economic Review*, *104*(5), 291–97.

Basaza, R., Criel, B., & Van der Stuyft, P. (2008). Community health insurance in Uganda: Why does enrolment remain low? A view from beneath. *Health Policy*, *87*(2), 172–184.

Bosch, M., Cobacho, M. B., & Pages, C. (2012). *Taking stock of eight years of implementation of Seguro Popular in Mexico* (Inter-American Development Bank. Mimeo). Washington, DC, USA: Inter-American Development Bank.

Bundorf, M. K., Levin, J., & Mahoney, N. (2012). Pricing and welfare in health plan choice. *American Economic Review*, *102*(7), 3214–48.

Capuno, J. J., Kraft, A. D., Quimbo, S., Tan Jr, C. R., & Wagstaff, A. (2016). Effects of price, information, and transactions cost interventions to raise voluntary enrollment in a social health insurance scheme: A randomized experiment in the philippines. *Health Economics*, *25*(6), 650–62.

Chemouni, B. (2018). The political path to universal health coverage: Power, ideas and community-based health insurance in Rwanda. *World Development*, *106*, 87–98.

Chiappori, P.-A., & Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy*, *108*(1), 56–78.

Cohen, J., & Dupas, P. (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics*, 1–45.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R., & Vickery, J. (2013). Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics*, *5*(1), 104–35.

Crump, R., Hotz, V. J., Imbens, G., & Mitnik, O. (2006). *Moving the goalposts: Addressing limited overlap in the estimation of average treatment effects by changing the estimand* (Technical Working Paper No. 330). Cambridge, MA, USA: National Bureau of Economic Research.

Currie, J., & Hyson, R. (1999). Is the impact of health shocks cushioned by socioeconomic status? The case of low birthweight. *American Economic Review*, *89*(2), 245–250.

Currie, J., & Stabile, M. (2002). *Socioeconomic status and health: Why is the relationship stronger for older children?* (Working Paper No. 9098). Cambridge, MA, USA: National Bureau of Economic Research.

Cutler, D. M., Lleras-Muney, A., & Vogl, T. (2008). *Socioeconomic status and health: Dimensions and mechanisms* (Working Paper No. 14333). Cambridge, MA, USA: National Bureau of Economic Research.

de la Sante, M. (2004). *Politique de developpement des mutuelles de sante au Rwanda* (Policy Document). Kigali, Rwanda: Author.

Dupas, P. (2011). Health behavior in developing countries. *Annual Review of Economics*, *3*(1), 425–49.

Dupas, P. (2014). Global health systems: Pricing and user fees. In A. Culyer (Ed.), *Elsevier encyclopedia of health economics* (Vol. 2, pp. 136–141). San Diego: Elsevier.

Einav, L., Finkelstein, A., & Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *Quarterly Journal of Economics*, *125*(3), 877–921.

Fink, G., Robyn, P. J., Sié, A., & Sauerborn, R. (2013). Does health insurance improve health? evidence from a randomized community-based insurance rollout in rural Burkina Faso. *Journal of Health Economics*, *32*(6), 1043–56.

Finkelstein, A., Hendren, N., & Shepard, M. (2019). Subsidizing health insurance for low-income adults: Evidence from Massachusetts. *American Economic Review*, *109*(4), 1530–67.

Fischer, T., Frölich, M., & Landmann, A. (2018). *Adverse selection in low-income health insurance markets: Evidence from a RCT in Pakistan* (IZA Discussion Paper No. 11751). Bonn, Germany: IZA Institute of Labor Economics.

Gnawali, D. P., Pokhrel, S., Sié, A., Sanon, M., De Allegri, M., Souares, A., … Sauerborn, R. (2009). The effect of community-based health insurance on the utilization of modern health care services: Evidence from burkina faso. *Health Policy*, *90*(2-3), 214–222.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199–236.

Kalisa, I., Musange, S., Collins, D., Saya, U., Kunda, T., & Parfait, U. (2016). *The development of community-based health insurance in rwanda: Experiences and lessons* (Tech. Rep.). Kigali, Rwanda & Medford, MA, USA: University of Rwanda College of Medicine and Health Sciences, School of Public Health and Management Science for Health.

Kalk, A., Groos, N., Karasi, J. C., & Girrbach, E. (2010). Health systems strengthening through insurance subsidies: The GFATM experience in rwanda. *Tropical Medicine and International Health*, *15*(1), 94–7.

Karlan, D., Osei, R., Osei-Akoto, I., & Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *Quarterly Journal of Economics*, *129*(2), 597–652.

Kayobotsi, P. (2019, July). Personal conversation.

Kremer, M., & Miguel, E. (2007). The illusion of sustainability. *Quarterly Journal of Economics*, *122*(3), 1007–65.

Lagomarsino, G., Garabrant, A., Adyas, A., Muga, R., & Otoo, N. (2012). Moving towards universal health coverage: health insurance reforms in nine developing countries in Africa and Asia. *The Lancet*, *380*(9845), 933–93.

Lammers, J., Warmerdam, S., & Ecorys, R. (2010). *Adverse selection in voluntary micro health insurance in nigeria* (Tech. Rep. No. 10-06). Amsterdam, Netherands: Amsterdam Institute for International Development.

Levine, D., Polimeni, R., & Ramage, I. (2016). Insuring health or insuring wealth? An experimental evaluation of health insurance in rural Cambodia. *Journal of Development Economics*, *119*, 1–15.

Lu, C., Chin, B., Lewandowski, J. L., Basinga, P., Hirschhorn, L. R., Hill, K., … Binagwaho, A. (2012). Towards universal health coverage: an evaluation of rwanda mutuelles in its first eight years. *PloS one*, *7*(6).

MINECOFIN. (2002). *Poverty reduction strategy paper* (Tech. Rep.). Rwanda, Kigali: Ministry of Finance and Economic Planning. Retrieved from `https://www.imf.org/external/np/prsp/2002/rwa/01/063102.pdf` (accessed Mars 2021)

Ministry of Health. (2012). *The third health sector strategic plan 2012-2018* (Policy Document). Kigali: Ministry of Health.

Ministry of Health. (2015). *Health financing sustainability policy* (Tech. Rep.). Kigali, Rwanda: Ministry of Health.

Ministry of Health of Rwanda. (2013). *Annual report july 2012 - june 2013* (Annual Report). Kigali: Ministry of Health.

Mobarak, A. M., & Rosenzweig, M. (2014). *Risk, insurance and wages in general equilibrium* (Working Paper No. 19811). Cambridge, MA, USA: National Bureau of Economic Research.

Nguyen, H., & Knowles, J. (2010). Demand for voluntary health insurance in developing countries: The case of Vietnam's school-age children and adolescent student health insurance program. *Social Science & Medicine*, *71*(12), 2074–82.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Parmar, D., De Allegri, M., Savadogo, G., & Sauerborn, R. (2014). Do community-based health insurance schemes fulfil the promise of equity? A study from Burkina Faso. *Health Policy and Planning*, *29*(1), 76–84.

Parmar, D., Souares, A., De Allegri, M., Savadogo, G., & Sauerborn, R. (2012). Adverse selection in a community-based health insurance scheme in rural Africa: Implications for introducing targeted subsidies. *BMC Health Services Research*, *12*(1), 181.

Thornton, R. L., Hatt, L. E., Field, E. M., Islam, M., Solís Diaz, F., & González, M. A. (2010). Social security health insurance for the informal sector in Nicaragua: A randomized evaluation. *Health Economics*, *19*(S1), 181–206.

United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development* (A/RES/70/1). New York: United Nations. Retrieved from `https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf` (Accessed April 23 2021)

Wagstaff, A., Nguyen, H. T. H., Dao, H., & Bales, S. (2016). Encouraging health insurance for the informal sector: A cluster randomized experiment in Vietnam. *Health Economics*, *25*(6), 663–74.

Wang, H., Otoo, N., & Dsane-Selby, L. (2017). *Ghana national health insurance scheme: improving financial sustainability based on expenditure review* [World Bank Studies]. Washington, DC, USA: The World Bank.

Wang, H., Zhang, L., Yip, W., & Hsiao, W. (2006). Adverse selection in a voluntary Rural Mutual Health Care health insurance scheme in China. *Social science & medicine*, *63*(5), 1236–45.

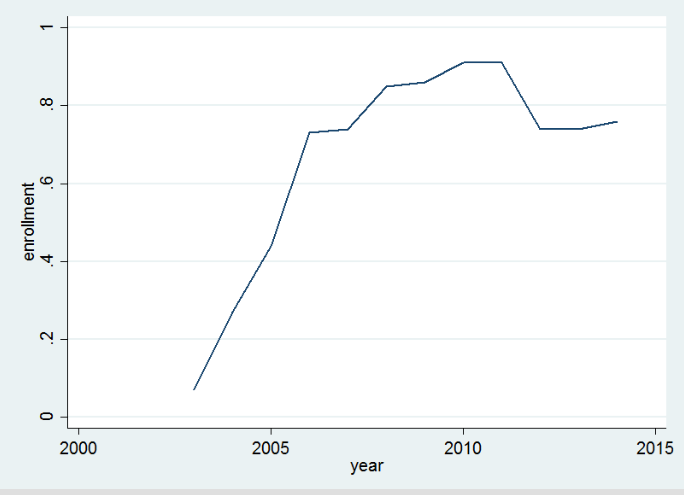
World Bank. (2020). *World development indicators.* Retrieved from `https://data.worldbank.org/indicator` (accessed June 2020)

Yilma, Z., Mebratie, A., Sparrow, R., Dekker, M., Alemu, G., & Bedi, A. S. (2015). Impact of Ethiopia's community based health insurance on household economic welfare. *World Bank Economic Review*, *29*(suppl_1), S164–73.

Zhang, L., & Wang, H. (2008). Dynamic process of adverse selection: Evidence from a subsidized community-based health insurance in rural China. *Social Science & Medicine*, *67*(7), 1173–82.

Appendix

Figure A1: National enrollment levels in CBHI, 2003–15



Sources: Kalisa et al. (2016); Ministry of Health of Rwanda (2013).

Table A1: Ubudehe classification

Group	Characteristics
Category 1 (abject poverty)	Households in this category of the population own no property, live by begging, and are wholly dependent on others. Children are malnourished, and households have no access to medical care.
Category 2 (very poor)	Households in this category have poor housing, live on a poor diet, are able to work a little, but do not own land or livestock.
Category 3 (poor)	Households in this category own a small portion of land and housing, live on their own labor, but have low production capacity and no savings. Their food is not very nutritious, and they often have no access to healthcare.
Category 4 (resourceful poor)	Households in this group share many of the characteristics of the poor. In addition they have small ruminants and children go to primary school. They own some land, cattle, and a bicycle, and have average production capacity.
Category 5 (food rich)	Households in this group own large areas of land, can afford a balanced diet, and live in decent homes. They employ others, own cattle, and can afford university education for their children.
Category 6 (money rich)	Households in this category have money in banks and can receive bank loans; own an above-average house, a car, livestock, and fertile lands; have access to sufficient food; and have permanent employment

Source: Adapted from MINECOFIN (2002)

Table A2: Correlation between Ubudehe categories and consumption poverty, 2010

	(1) Non-poor	(2) Poor
Ubudehe group 1	0.96%	1.75%
Ubudehe group 2	17.82%	33.63%
Ubudehe group 3	68.71%	59.43%
Ubudehe group 4	12.04%	5.20%
Ubudehe group 5	0.48%	0.00%
Ubudehe group 6	–	–

Note: The poverty measure is defined by the Rwandan government and is based on household consumption level.

Table A3: Summary statistics

Variables	(1) N	(2) Mean	(3) Std. Dev	(4) Min	(5) Max
Health issue	7 612	0.213	0.409	0	1
Disability	7 612	0.049	0.216	0	1
Piped water	7 612	0.396	0.489	0	1
Sanitation	7 612	0.781	0.413	0	1
Work	7 612	0.505	0.500	0	1
Salary worker	7 612	0.239	0.427	0	1
Own nonfarm enterprise	7 612	0.120	0.325	0	1
Poor	7 612	0.423	0.494	0	1
Consumption HH	7 612	218 426	183 718	10 951	3 301 187
Own house	7 612	0.931	0.252	0	1
Own land	7 612	0.965	0.183	0	1
Rural	7 612	0.869	0.336	0	1
Age 0–5	7 612	0.144	0.351	0	1
Age 6–19	7 612	0.371	0.483	0	1
Age 20–29	7 612	0.143	0.350	0	1
Age 30–39	7 612	0.127	0.333	0	1
Age 40–49	7 612	0.080	0.272	0	1
Age 50–65	7 612	0.093	0.291	0	1
Age $>$ 65	7 612	0.041	0.199	0	1
Female	7 612	0.529	0.499	0	1
HH size	7 612	5.613	2.279	1	12
Travel time clinic (hours)	7 612	0.792	1.504	0	30
Travel time hospital (hours)	7 612	3.125	2.05	0	12
Ubudehe 1	7 612	0.013	0.489	0	1
Ubudehe 2	7 612	0.249	0.432	0	1
Ubudehe 3	7 612	0.645	0.478	0	1
Ubudehe 4	7 612	0.089	0.285	0	1
Ubudehe 5	7 612	0.003	0.051	0	1
Ubudehe 6	.	.	.	.	.

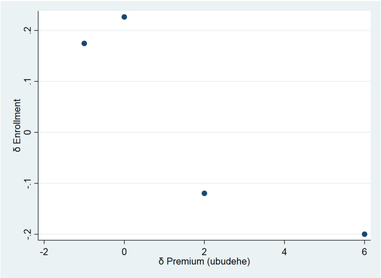
Note: The results present sample means, standard deviation, and maximum and minimum values for each variable.

Table A4: Estimated correlation between household characteristics and Ubudehe groups 1 and 2, 2014

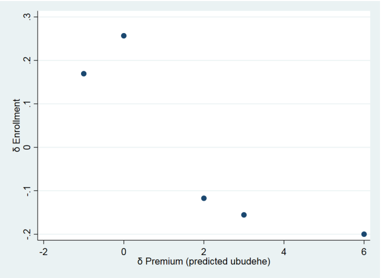
Variables	Ubudehe 1 or 2
Sanitation	-0.174***
	(0.0394)
Piped water	-0.0136
	(0.0363)
Main light source electric	-0.274***
	(0.0776)
Improved flooring	-0.113**
	(0.0450)
Own land	-0.146*
	(0.0847)
Own house	-0.160***
	(0.0601)
Poor	-0.117**
	(0.0580)
Non-poor	-0.172***
	(0.0530)
Observations	3,806
District FE	Yes

The table describes the estimated association between household characteristics and the likelihood of being classified as Ubudehe 1 or 2. Estimations are based on data from 2014, post-treatment, and describe marginal effects at means.

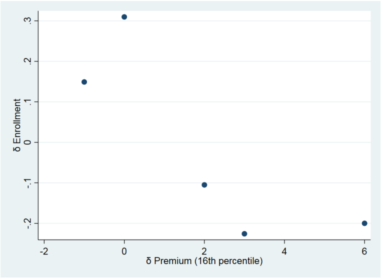
Figure A2: Within-group variation in insurance premium and enrollment within each premium group, before and after the implementation of the new premium scheme



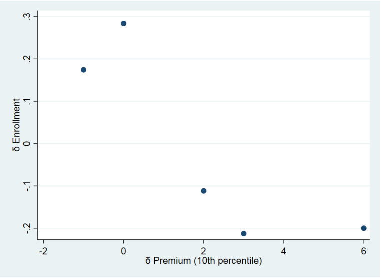
(a) Ubudehe



(b) Ubudehe (predicted)



(c) Consumption 16th percentile



(d) Consumption 10th percentile

*Notes:*Differences in insurance premiums and enrollment levels before and after the policy change within each separate premium group, as presented in table 3. The observations show the relation between changes in premium costs and insurance enrollment within each premium group. The graphs plot the distribution of observations for each alternative definition of the premium scheme prior to the policy change (see tables 2 and 3 for more detailed information).

Table A5: Price sensitivity of the demand of health insurance

Variables	(1) All	(2) Drop δ premium = 0	(3) Drop δ premium = (−1)	(4) Drop δ premium = 3
Panel A: Ubudehe				
Premium (RwF 1,000)	-0.109***	-0.0971***	-0.156***	
	(0.0143)	(0.0140)	(0.0398)	
Observations	7,612	6,932	6,294	
R-squared	0.066	0.062	0.063	
Number of PID	3,806	3,466	3,147	
Panel B: Predicted ubudehe				
Premium (RwF 1,000)	-0.103***	-0.0929***	-0.146***	-0.104***
	(0.0137)	(0.0135)	(0.0399)	(0.0142)
Observations	7,612	7,090	6,136	7,252
R-squared	0.065	0.062	0.063	0.064
Number of PID	3,806	3,545	3,068	3,626
Panel C: 16th percentile				
Premium (RwF 1,000)	-0.101***	-0.0870***	-0.168***	-0.0983***
	(0.0136)	(0.0134)	(0.0361)	(0.0144)
Observations	7,612	7,076	6,150	6,930
R-squared	0.066	0.061	0.074	0.058
Number of PID	3,806	3,538	3,075	3,465
Panel D: 10th percentile				
Premium (RwF 1,000)	-0.101***	-0.0944***	-0.150***	-0.101***
	(0.0136)	(0.0135)	(0.0470)	(0.0141)
Observations	7,612	7,288	5,938	7,170
R-squared	0.068	0.065	0.062	0.063
Number of PID	3,806	3,644	2,969	3,585
FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Note: The results from estimating equation (1) using the complete sample (column 1) and restricted samples that stepwise exclude one premium group at a time (columns 2–4). The dependent variable is a dummy variable indicating enrollment in CBHI. Column 1 presents the preferred estimation strategy shown in table 5, and columns 2–4 estimate the same specification based on the restricted samples. All estimations include individual fixed effects and a full set of individual and household controls, such as labor status and rural location, individual health status and household consumption, and access to water and sanitation services. All estimations are presented separately for the alternative definitions of the premium scheme prior to the policy change (panels A–D). For further details see table 2. Standard errors are clustered at household level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure A3: Predicted likelihood of being categorized in Ubudehe group 3,—by Ubudehe group, before matching

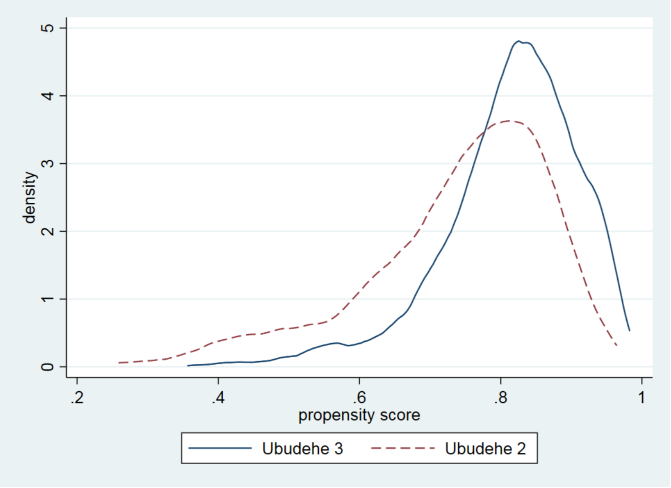


Figure A4: Predicted likelihood of being categorizedin Ubudehe group 3, —by Ubudehe group, common support

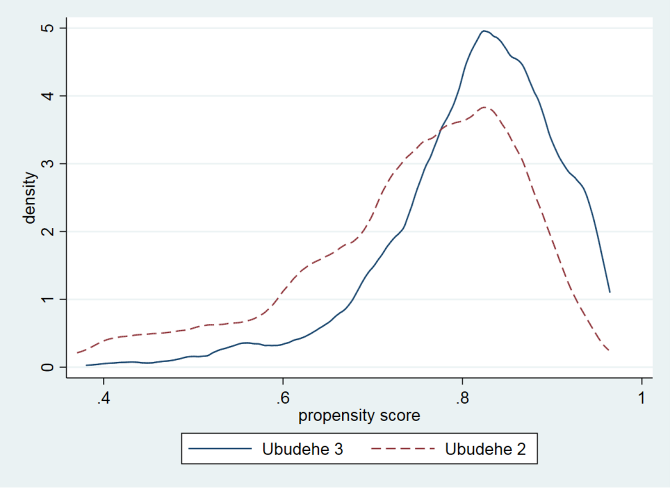


Table A6: Likelihood of being categorized in Ubudehe group 3

Variables	(1) Ubudehe 2 & 3	(2) Common support	(3) $10th$–$90th$ percentile	(4) $20th$–$80th$ percentile
Health status	-0.186	-0.0277	-0.0301	-0.0351
	(0.119)	(0.0178)	(0.0203)	(0.0305)
Disability	-0.265	-0.0351	-0.0465	-0.0287
	(0.208)	(0.0312)	(0.0352)	(0.0492)
Work	0.298	0.0419	0.0454	0.0622
	(0.201)	(0.0300)	(0.0350)	(0.0549)
Salary worker	-0.248*	-0.0368*	-0.0509**	-0.0274
	(0.146)	(0.0218)	(0.0255)	(0.0431)
Own nonfarm	0.0749	0.0148	0.00968	0.0317
	(0.170)	(0.0254)	(0.0298)	(0.0497)
HH size	0.175***	0.0261***	0.0286***	0.0556***
	(0.0274)	(0.00404)	(0.00499)	(0.0102)
Own house	0.909***	0.134***	0.146***	0.207***
	(0.163)	(0.0245)	(0.0280)	(0.0395)
Own land	0.694**	0.0950**	0.0987**	0.108
	(0.287)	(0.0437)	(0.0498)	(0.0675)
HH consumption (RwF1,000)	0.000366	6.66e-05	0.000112	0.000126
	(0.000359)	(5.48e-05)	(6.98e-05)	(0.000121)
Rural	-0.0176	-0.000388	0.0157	0.0176
	(0.147)	(0.0220)	(0.0256)	(0.0384)
Age6–19	-0.340**	-0.0501**	-0.0608***	-0.136***
	(0.135)	(0.0201)	(0.0233)	(0.0398)
Age20–29	0.168	0.0244	0.0288	-0.0446
	(0.233)	(0.0349)	(0.0405)	(0.0679)
Age30–39	-0.0132	0.00106	0.00973	-0.0907
	(0.254)	(0.0381)	(0.0444)	(0.0740)
Age40–49	-0.0988	-0.0174	-0.00161	-0.0814
	(0.269)	(0.0401)	(0.0472)	(0.0770)
Age50–65	-0.349	-0.0501	-0.0517	-0.115*
	(0.245)	(0.0367)	(0.0421)	(0.0633)
Age $>$ 65	-0.895***	-0.128***	-0.142***	-0.222***
	(0.269)	(0.0406)	(0.0462)	(0.0656)
Female	-0.127	-0.0184	-0.0199	-0.0331
	(0.0955)	(0.0142)	(0.0165)	(0.0263)
Piped water	0.0498	0.00750	0.000141	-0.00975
	(0.108)	(0.0161)	(0.0187)	(0.0285)
Toilet	0.483***	0.0709***	0.0663***	0.126***
	(0.103)	(0.0154)	(0.0177)	(0.0281)
Constant	-1.331***	-1.289***	0.310	-1.030
	(0.362)	(0.372)	(0.707)	(1.076)
Observations	3,072	3,031	2,591	1,290

Notes: Predictions of Ibudehe categorization based on household characteristics. Column 1 presents the predicted likelihood of being categorized in Ubudehe group 3 using a sample that includes only households from Ubudehe groups 2 and 3, estimated using a logistic regression model. Column 2 further restricts the sample by including only households in Ubudehe group 2 or 3 that are in the range of common support, whereas columns 3 and 4 are restricted to include only households within the 10th–90th and 20th–80th percentile of common support. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7: Sensitivity analysis - price sensitivity

Variables	(1) Ubudehe 2 & 3	(2) Common support	(3) 10th–90th percentiles	(4) 20th–80th percentiles
Premium (RwF 1,000)	-0.106***	-0.0951***	-0.101***	-0.106***
	(0.0173)	(0.0142)	(0.0150)	(0.0173)
Observations	2,580	6,062	5,182	2,580
R-squared	0.089	0.076	0.076	0.089
Number of PID	1,290	3,031	2,591	1,290

Notes: Table shows the results for the baseline linear probability regression with individual fixed effects corresponding to estimating equation (1) for the restricted samples constructed in appendix table A6. Controls include individual and household characteristics, such as labor status and rural location, individual health status and household consumption, access to water and sanitation services, and distance to nearest clinic. Standard errors clustered by household are shown in parentheses below the estimated coefficient. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Missing values: descriptive statistics on sectors included in and excluded from in the sample due to missing cost information

Sector characteristics	(1) All	(2) Sample	(3) Missing	(4) Diff
Share access piped water	0.479	0.484	0.469	-0.014
Share access sanitation	0.840	0.839	0.842	0.002
Share urbanization	0.159	0.165	0.145	-0.020
Share individuals > 50 yrs.	0.117	0.116	0.121	0.005
Share children < 5 yrs.	0.163	0.162	0.165	0.003
Tot. population of sector	21 872	21 673	22 471	798
Avg. HH consumption	241 475	234 692	257 537	22844
Observations	416	295	121	

Notes: Columns 1–3 provide descriptive statistics for all sectors, those included in the study sample and those excluded from the sample due to missing cost information. Column 4 presents differences between the sample and the missing sectors, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A9: Cost curve: average insurer costs excluding operational costs

	(1)	(2)	(3)
Average insurer costs			
Avg. premium (RwF 1,000)	0.711***	0.559***	0.567***
	(0.126)	(0.151)	(0.154)
Observations	295	295	295
R-squared	0.148	0.439	0.448
District FE	No	Yes	Yes
Sector characteristics	No	No	Yes

Notes: the table shows the results from estimating equation (2) using the cost data that excludes operational costs. Column 1 provides the unconditional association between average premium and patient costs (excluding operational costs) across sectors, column 2 includes district fixed-effects, and column 3 additionally controls for sector characteristics. When included, the controls contain share of section households with piped water, access to sanitation, share of households in urban areas, share of children below 5 years old and individuals older than 65, total population size and average household consumption. The estimated correlation between the average premium and the average insurer cost is the slope of the cost curve. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

II

The Role of Childcare in Firm Performance: Evidence from Female Entrepreneurship in Mexico

Carolin Sjöholm*

Abstract

Microenterprises represent an important source of employment in many developing countries. Earlier literature has documented that female-run microenterprises underperform relative to those run by men on many indicators, although the reasons for this discrepancy in large part remain enigmatic. This paper estimates the importance of childcare obligations as a barrier for female entrepreneurship. I use difference-in-difference and triple-difference designs to study how a federal daycare program affects the performance of female-run microenterprises in Mexico. The program provided childcare services for children under 4 years old whose mothers worked in the informal sector, and varied across time and space. I find no evidence that the program was associated with changes in business performance measured by the likelihood of running a home-based business or having an employee, the number of hours worked, physical capital or the likelihood of applying for a credit. The results are consistent, irrespective of the choice of estimation strategy.

Keywords: female self-employment, childcare, microenterprises, daycare program
JEL classification: H55, J13, J22, J46, J48

*I am thankful to Måns Söderbom and Annika Lindskog for advice and support. This paper has also benefited from comments from Anna Bindler, Anja Benshaul-Tolonen, and Josephine Gakii Gatua.

1 Introduction

Microenterprises account for an important share of employment in many developing countries and women run the majority of the firms (Klapper & Parker, 2011). A wide body of literature has shown that businesses run by women are smaller and less profitable than those run by men (World Bank, 2012; Rijkers & Costa, 2012; Hardy & Kagy, 2018; Bruhn, 2009). Several factors that could restrict female business activities have been studied in literature, but the reasons why female-run firms underperform businesses run by men remains enigmatic. One potential explanation is that women face additional barriers that prevent them from taking advantage of economic opportunities in the market (World Bank, 2012). This paper studies the importance of childcare obligations as a constraint for female entrepreneurship by placing restrictions on female mobility, time use and market opportunities. Although unpaid housework and childcare have often been mentioned as a limiting factor for female entrepreneurship (Duflo, 2012; Jayachandran, 2020; Hardy & Kagy, 2018; Bruhn, 2009; Fitzpatrick & Delecourt, forthcoming.), to my knowledge the current study is one of the first to investigate the importance of childcare obligations as a bottleneck for firm performance among female-run microenterprises in the developing country context.

I use the introduction of a subsidized childcare program in Mexico as a quasi-experiment to estimate the impact of increased access to affordable childcare on firm performance among female-run microenterprises (defined as small businesses with less than six employees). *Estancias Infantiles para Apoyar a Madres Trabajadoras* was introduced in 2007 and offered subsidized childcare to low-income mothers working in the informal sector. According to the program rules, Estancias Infantiles was intended exclusively for children younger than 4 years old (Secretaría de Gobernación, 2007). This created an eligibility cutoff, where mothers with children who were just under 4 years old were eligible for the program, but those with children who had just turned 4 were not. I use this discontinuity in eligibility to estimate the effect of the program on female-owned microenterprises using a difference-in-difference (DD) estimation strategy. Additionally, I take advantage of the sequential rollout of the program across Mexican municipalities, which created a geographic variation in program intensity within and across municipalities over time, in a triple-difference design (DDD). Firm performance is proxied by the likelihood of having an employee, hours worked, capital stock, business location, and demand for credit.

Women typically spend disproportionately more time than men on unpaid housework and childcare (Samman et al., 2016; Potter et al., 2013).[1] Deep-rooted social norms regarding responsibility for housework and childcare limit female entrepreneurship by placing a constraint on time use and subsequently labor market opportunities (Samman et al., 2016; Bruhn, 2009; Jayachandran, 2020; Duflo, 2012; Razavi, 2012). Demand and responsibilities related to childcare and housework require female entrepreneurs to balance business and family obligations, limiting their capacity to separate the two spheres. This interconnection between family and business likely represents an important factor for understanding female entrepreneurship (Jayachandran, 2020; Friedson-Ridenour & Pierotti, 2019).

By limiting female market labor supply, childcare obligations could affect the optimal size of capital and investment among female-run firms. According to standard economic theory, an entrepreneur will invest resources into a business until the interest rate equals the rate of return on the investment. In this context, childcare obligations are likely to further constrain the marginal return to investment by limiting both mobilization and productivity among female entrepreneurs. Previous studies suggest that female-headed firms are more likely to operate from home than male-headed enterprises (Razavi, 2012; Bruhn, 2009). Furthermore, approximately 4 out of 10 female business owners stated that they had to bring their children to work, which resulted in lower profits compared to other female-run businesses with no children present (Fitzpatrick & Delecourt, forthcoming.). By alleviating the interrelation between family obligations and business activities, Estancias Infantiles could lead to increased marginal return to investment by improving entrepreneur mobilization and productivity among female entrepreneurs and, hence returns to investment.

In this study I combine data from the Mexican National Survey of Occupation and Employment (ENOE) and the National Survey of Micro-enterprises (ENAMIN) with administrative data on the rollout of Estancias Infantiles, to study the effect of the program on the performance of female microenterprises in Mexico. ENOE is a national household survey that provides information on self-employed women and their businesses, as well as individual and household characteristics. This makes it possible to match information on business performance with the number and ages of each entrepreneur's children. ENAMIN complements the ENOE survey by providing

[1]There are substantial differences in how men and women spend time on unpaid housework in Mexico. On average, women dedicate 2.5 times more time than men per day to care for household members. Men spend 131 minutes per day on average per day on unpaid housework and childcare, and 478 minutes on paid work. For women, the situation is close to the reverse: they dedicate 331 minutes per day to unpaid work and 236 minutes to market work (OECD, 2014)

information from in-depth interviews with a subsample of all individuals from the ENOE sample who stated that they were self-employed. Finally, I merge the data on self-employed women and their businesses with administrative data on the number of children enrolled in the program per month in each municipality. The data on program enrollment are used to construct a variable of treatment intensity that measures the availability of Estancias Infantiles in each municipality.

Overall, I find no evidence that Estancias Infantiles was associated with business performance, suggesting that mothers do not adjust their entrepreneurship as a consequence of increased access to subsidized daycare services. The results indicate that the estimated treatment effects were statistically insignificant for the majority of firm performance measures, such as the location of firm operations, physical capital and the likelihood of having applied for a loan or paid staff. I find some evidence that the program was associated with an increase in the number of hours worked per week among urban firms. Furthermore, I rule out that in- and outflows of entrepreneurs from the self-employed sector explain the results by showing that Estancias Infantiles was not associated with the decision to become self-employed.

This study relates primarily to two bodies of literature. First, it contributes to a rich literature on the determinants of microenterprise development in the developing country context, which has documented a substantial gender gap in relation to a number of indicators, such as business profits and sales, where female entrepreneurs have often underperformed their male counterparts (Fiala, 2018; Hardy & Kagy, 2018). I add to this literature by providing evidence of the importance of childcare obligations to female entrepreneurship and business performance. Several factors have been proposed by researchers and policymakers as reasons for the lack of growth among female-run enterprises, including a lack of access to financial capital (Karlan & Zinman, 2011), business training (Valdivia, 2015), and saving mechanisms (Dupas & Robinson, 2013). However, while earlier research has suggested that the expansion of microcredit has a positive effect on firm profits when evaluated at the household level (Crépon et al., 2015; Banerjee et al., 2015), recent studies have consistently found that increased access to credit has no effect on profits among female-run microbusinesses (Fafchamps et al., 2011; De Mel et al., 2008; Fiala, 2018). One explanation for this has been that women face social pressure to share their income with the family (Fiala, 2018; Bastian et al., 2018). A number of papers have evaluated the importance of the lack of access to saving mechanisms among female entrepreneurs as a constraint for business growth.

These studies have found some effects of savings on business practices but have failed to find evidence that such improvements have translated into increased profits, sales, and investment (Bastian et al., 2018). Another strand of literature has examined the relevance of business training on firm performance. Few studies find significant effects of managerial training on female business performance and survival (Fiala, 2018; McKenzie & Woodruff, 2013; Bruhn & Zia, 2013; Karlan & Valdivia, 2011; Drexler et al., 2014). Overall, while policy interventions that aim to grow microenterprises, such as increased access to financial capital and business skills training, often had no effect on female entrepreneurship, such policies have frequently been effective among male-run businesses (Fiala, 2018; Bastian et al., 2018; McKenzie & Woodruff, 2013).

Second, this study relates to the literature on the effect of subsidized childcare on female labor force participation, the findings of which are inconclusive. Although some studies found no evidence that formal childcare or preschool programs affected female labor force participation (Manley & Vásquez Lavín, 2013; Medrano, 2009; Havnes & Mogstad, 2011; Fitzpatrick, 2010), numerous papers have found a positive effect in both developed and developing countries (Ángeles et al., 2011; Baker et al., 2008; Lefebvre & Merrigan, 2008; Hallman et al., 2005; Attanasio & Vera-Hernández, 2004; Halim et al., 2019; Berlinski & Galiani, 2007; Martínez & Perticará, 2017). Also the effects of Estancias Infantiles have been studied before. Calderon (2014) showed that a 10% rise in childcare availability increased the probability of working among eligible women by 1.5 percentage points on average.

The identification strategy in my study is closely related to the empirical strategy employed by Calderon (2014), although the current paper primarily relies on the variation in eligibility resulting from the age cutoff create by the program rules. Importantly, my study differs from that of Calderon by studying a different aspect of Estancias Infantiles, focusing on different outcomes. While Calderon investigated the effects of the program on female and male labor force participation, the purpose of this paper is to evaluate the effects of the introduction of Estancias Infantiles on the performance of female-run microenterprises. To my knowledge, this is the first paper to use the introduction of a nation wide daycare program as a quasi-experiment to study the importance of childcare obligations as an obstacle for female entrepreneurship, using observational data. The analysis in this paper was made possible by the unique possibility of combining survey data on firm operations with data on family demographics.

The paper is organized as follows: Section 2 describes Estancias Infantiles and other formal childcare services in Mexico. Section 3 provides the data and summary statistics. Section 4 outlines the empirical framework and discusses threats to identifications and provides 5 presents the results and robustness checks. Section 6 concludes.

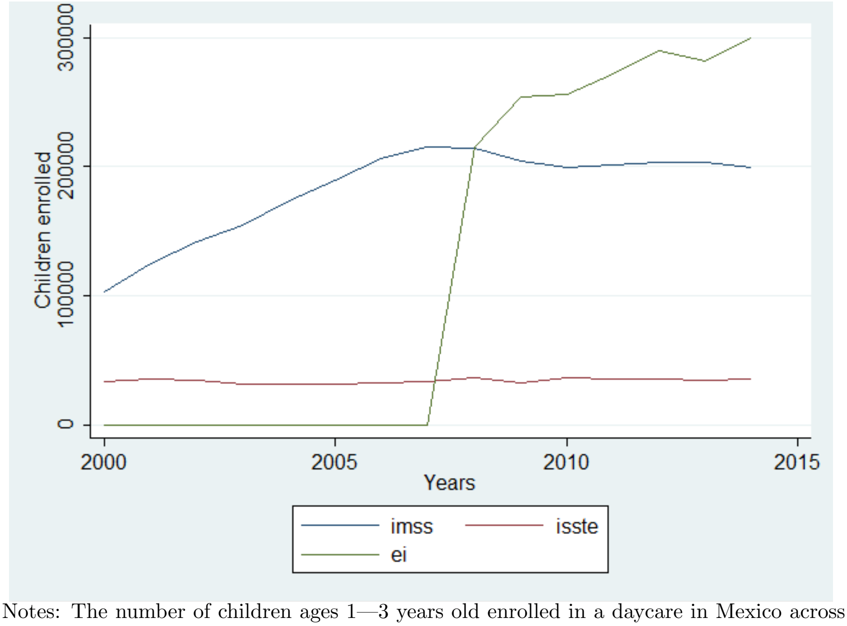
2 Childcare in Mexico

Estancias Infantiles para Apoyar a Madres Trabajadoras is a federal daycare program introduced in Mexico in 2007, with the aim of expanding public daycare to workers in the informal sector. Estancias Infantiles offered subsidized childcare to low income mothers that were working, studying, or looking for a job. Estancias Infantiles was part of the government's strategy to eradicate poverty by diminishing the vulnerability of low-income and single parent households. The program targeted working mothers with children 1 to 3 years old (under 4 years old) who did not have access to social security daycare servicesand lived in households with an income less than six times the minimum wage (Ángeles et al., 2011).

Public daycare has been offered in Mexico since the 1970s, when the Mexican Institute for Social Security (IMSS) introduced daycare for mothers covered by social security who were working in the formal sector. Despite efforts to expand the coverage of the daycare program, IMSS far from satisfied the childcare demand among workers in the formal sector (Staab & Gerhard, 2011). In 2005, the program enrolled approximately 200,000 children, representing 20% of the eligible group. In 2008, the program had become the most important childcare provider for children under 4 years old, representing 84% of all childcare centers and covering 56% of all enrolled children (Staab & Gerhard, 2010). Importantly, IMSS services left out half of the labor force by not targeting informal workers. In addition to IMSS, the Institute for Social Security and Services for Public Employees (ISSSTE) offered daycare services to public employees (Staab & Gerhard, 2011).

Figure 1 shows the development of childcare services in Mexico from 2000 to 2015. Estancias Infantiles saw a dramatic expansion during the first years of operation. In 2014, the program offered childcare services to 300,000 children. During this time, the supply of daycare centers from the formal sector was stable.

Figure 1: Childcare service in Mexico



Notes: The number of children ages 1—3 years old enrolled in a daycare in Mexico across time. Prior to the implementation of Estancias Infantiles (EI), the Mexican Institute for Social Security (IMSS) and the Institute for Social Security and Services for Public Employeed (ISSSTE) provided childcare services for parents who were working in the formal sector.

Estancias Infantiles provided subsidized daycare services that covered up to 90% of the cost of childcare with a maximum amount of 700 Mexican pesos a month per child. Beneficiaries had to pay the remaining part of the childcare cost, resulting in a daycare fee of about 200 pesos per month on average for enrolling a child in Estancias Infantiles (Secretaría de Gobernación, 2007). This fee was low compare with the cost of already existing daycare alternatives. Households in the informal sector did not have access to other public daycare centers operated by the national social security. However, a study of private daycare centers in Mexico City indicated that there was a large variation in costs and services among centers. The monthly cost for children under 1 year ranged from approximately 500 to 5,350 pesos for full-time service, whereas the tuition for children between 2 to 6 years old ranged from 650 to 3,780 pesos per month (PROFECO, 2014).

According to the program rules, any existing childcare center that satisfied the rules of affiliation was eligible to join the Estancias Infantiles daycare network. At the same time, any individual or organization who wanted to establish and operate a childcare center could apply to enter the program. The applications were verified by the responsible government institution and an evaluation was made of the suitability of the applicant and the property intended for the daycare center (Secretaría de Gobernación, 2007). Furthermore, the program offered a governmental grant to those willing to open a childcare center so that the could upgrade their facilities in accordance with the program guidelines. The maximum amount was 61,000 Mexican pesos (approximately USD 2,900). The rules of operation allowed the use of 20% of the funds to be used to cover operational costs during the first two months of operation. The strategy was intended to help the program quickly expand (Secretaría de Gobernación, 2007).

The introduction of Estancias Infantiles increased the total supply of daycare services by adding to the existing childcare centers. Approximately 87% of the childcare centers within the program were newly founded, while the other 13% were existing centers that converted to Estancias Infantil (Ángeles et al., 2011). As a result of the implementation strategy the geographic location of the centers was defined by the supply and demand in the market, without any government intervention. I discuss the implications of the rollout process in the following section.

3 Data

The analysis in this paper is based on data from the Mexican National Survey of Occupation and Employment (ENOE). ENOE is a quarterly survey with a rotating panel design that is representative at the national level. I use one survey round per year between 2005 and 2014 to construct a repeated cross sectional data set.[2] Besides detailed information about the labor conditions for all working-age individuals in a household, the ENOE survey also contains sociodemographic characteristics for all household members. This makes it possible to link information about self-employed women and their businesses with the ages and number of children they had. I refer to the ENOE sample as the national sample.

[2]Data are collected through an extended and a shorter versions of the survey. The extended version is rolled out once every year, in different trimesters, whereas a shorter version is used for the remaining survey rounds. In the baseline sample, I use only data from the extended survey since this version contains better and more specific information on the microenterprises and entrepreneurs included in this analysis.

Every second year, the National Survey on Micro-enterprises (ENAMIN) is conducted. A a random subsample of all self-employed individuals is drawn from the ENOE data and asked to participate in in-depth interviews regarding their businesses and entrepreneurial activities. The survey provides comprehensive information on microenterprises with less than 6 employees (15 for manufacturing firms), such as access to financial credit, physical capital, and sales and profits. Given this sampling strategy, the ENAMIN includes formally registered businesses as well as small and unregistered firms that would have been excluded from business surveys based on administrative records on registered businesses. This is important for this analysis because Estancias Infantiles was directed primarily towards working parents in the informal sector.

I include data from the 2002, 2008, 2010 and 2012 ENAMIN survey rounds. The current version of the survey dates back to 2008. Before that, the survey in 2002 was conducted using a sub-sample of the National Urban Employment Survey (ENEU) and covered only urban areas.[3] To make the surveys comparable over time, I restrict the ENAMIN analysis to urban areas with more than 100,000 inhabitants. Additionally, I limit the analysis to include only survey questions that are comparable over the years.[4] As a result, the pre-treatment analysis is limited to urban areas. I refer to the ENAMIN data as the urban sample.

The ENOE survey stretches back several decades, but was also reconstructed in 2005. The reconstruction of the survey does not affect the main analysis since Estancias Infanties was first introduced in 2007. However, to construct pre-treatment trends I use data from an earlier version of ENOE, the National Employment Survey (ENE), for the period 1995—2004. Similarly, I rely on the ENEU survey from 2000 to 2004 to construct pre-treatment trends, for urban areas.

To identify entrepreneurs living in municipalities that were treated by the program, I link administrative records on program enrollment with the data on firm and household characteristics. The administrative data were provided by the Secretary of Social Development (Sedesol) in Mexico City and contains information regarding the number of enrolled children per month in each municipality and the date when each center began to operate. I construct a measure of program intensity by dividing the number of children enrolled in the program by the total number of children between 1 to 3 years old in each municipality. Data on the number of children are provided by the 2010 Mexican Census (Censo de Población y Vivienda 2010). Importantly, these data

[3]The survey was not conducted between 2002 and 2008.

[4]This has previously been done by BenYishay & Pearlman (2014)

includes all children in the target age group in each municipality, including children whose parents were eligible for childcare services offered by the formal social security system. The share of enrollment in Estancias Infantiles serves as a proxy for availability of the program in each municipality in a specific year.[5]

The analysis in this paper is restricted to self-employed women. For the purpose of this stud, a woman is considered self-employed if she stated that she had worked at least one hour during the past week and did not have a boss. Furthermore, I restrict the sample to include only household heads and spouses of household heads that are between 15 and 65 years old, who stated that they were not covered by social security. I do not restrict the sample according to income level. As previously mentioned, Estancias Infantiles was offered to households with an income below a threshold of six times the minimum wage. In practice the program relied on self-reported income and employment records (Calderon, 2014), which likely made the program available to most self-employed households in the informal sector. According to the descriptive statistics in appendix table A1, the average income level among entrepreneurs in the national sample was approximately 2,900 pesos (USD 146). This is well below the program income threshold of a monthly income of 8,500 Mexican pesos (USD 420), which is the equivalent of six times the minimum wage in the country in 2007.[6] The results indicate that most households in the study sample were not likely to have been affected by the income threshold.

3.1 Eligibility

According to the program rules, entrepreneurs with children 1 to 3 years old were eligible for Estancias Infantiles. I use the program regulation to define treatment. An entrepreneurs was eligible if her youngest child was 1 to 3 years old and she lived in a municipality that had introduced the program. Furthermore, I limit the sample to include only women whose youngest child was under 6 years old, which results in a control group that consists of women whose youngest child was between 4 and 6 years old. (I present results with alternative control groups in the robustness section.) I chose this sample in order to obtain a control group of women with childcare obligations as similar as possible to those of the women in the treatment group. Children

[5]This strategy has previously been used by (Calderon, 2014).

[6]In 2007, the minimum wage was 50.57 pesos per day (INEGI, 2019). During a month with 28 working days, this would result in a minimum income of 1,416 pesos, corresponding to approximately 8,500 pesos.

between 4 and 6 years old were not eligible for Estancias Infantiles, but were not yet attending primary school. They were, however, most likely enrolled in mandatory public preschool, which was held only four hours daily and consequently did not fulfill the daycare needs of a full-time working mother. As a result, working women with children 4 to 6 years old continued to be in need of childcare services.

Given the definition of eligibility, it is possible that women in the treatment group also had older children aged 4—6 years. This would lead to an underestimation of the treatment effect since these women most likely continued to face barriers to economic activities due to childcare obligations for the older siblings, even in the presence of Estancias Infantiles. I do not exclude this group of women from the treatment group since I consider this part of the treatment effect when studying the effect of any childcare program directed towards young children under 4 years old. I do, however, provide estimations based on a sample that excludes women in the treatment group who also have children aged 4—6 years in the section for robustness tests.

Descriptive statistics for women in the treatment and control groups defined by the age-cutoff are presented in appendix tables A2 and A3. The results suggest that there are significant differences between the control and treatment groups. Eligible women were significantly younger and more likely to have at least a secondary education compared with women in the control group, in both the national and the urban samples. Furthermore, the family composition differed between both groups: compared with the control group, entrepreneurs in the treatment group were significantly more likely be married and to live in households with fewer adult household members, but with a larger number of children. This difference is statistically insignificant among urban entrepreneurs. Importantly, there is no significant differences in income between households in the treatment and control groups in the national sample. Furthermore, eligible and ineligible households are as likely to live in rural areas in the national sample.

The significant differences between entrepreneurs in the treatment and control groups could lead to biased estimates if characteristics associated with program eligibility were simultaneously associated with business performance. I further discuss the potential effects of such confounders in the following section.

3.2 Rollout of the program

Appendix figure A1 shows the cumulative number of total childcare centers from the introduction of Estancias Infantiles in 2007 until 2014. The program expanded rapidly during the first years of implementation and subsequently leveled off. During the first year of the program, 6315 centers were opened in 935 municipalities. In December 2014, 1146 municipalities had at least one childcare center operating. This represents approximately half of Mexico's 2456 municipalities.

The availability of Estancias Infantiles varied among and within municipalities, over time. Table 1 shows the rollout of the program across the municipalities in the ENOE data. Row 1 in Panel A presents the total number of municipalities included in the ENOE survey per year, after adjusting the sample to include only female entrepreneurs with children between 1 to 6 years old (the target group for this study). The number of municipalities varies from year to year as a result of the ENOE sampling frame; that is, all municipalities were not surveyed every year.The ENOE survey is constructed to be representative at the level of federal entities, as well as communities of four population sizes.

The baseline sample used in the current study includes only municipalities that introduced Estancias Infantiles directly in 2007. Given this strategy, I exclude municipalities that introduced the program after 2007 or not at all. These municipalities represent the difference between row 1 and row 2 in Panel A.[7] The numbers of municipalities included in the main national sample each year are presented in row 2 (panel A). The results suggest that Estancias Infantiles quickly expanded across municipalities, and that the program was implemented directly in 2007 in 379 of the sample municipalities.

[7] I also estimated the effect of Estancias Infantiles using the full ENOE sample, including all municipalities in the ENOE sample. I found that the results are consistent irrespective of the choice of sample. Results are available upon request. Additionally, the multiple-period DD estimations using all municipalities are presented in the robustness section.

Table 1: Descriptive statistics: intensity and rollout of the Estancias Infantiles program

VARIABLES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Panel A. Municipalities										
All municipalities	455	451	462	468	422	446	422	441	401	379
Baseline sample (treated in 2007)	369	368	379	384	347	358	337	356	328	311
Panel B. Intensity										
Mean	0	0	0.039	0.055	0.065	0.068	0.071	0.076	0.077	0.078
Min	0	0	0.001	0.003	0.002	0.002	0.002	0.003	0.002	0.002
Max	0	0	0.147	0.198	0.212	0.232	0.225	0.333	0.299	0.253

Notes: Panel A presents the number of treated municipalities included in the analysis sample each year. The first row shows the total number of municipalities in ENOE per year, and the second row displays the number of municipalities where Estancias Infantiles was introduced in 2007–the baseline national sample used in this analysis. Panel B displays the average intensity of the program per year among municipalities in the baseline national sample, as well as maximum and minimum program intensity levels.

Although Estancias Infantiles soon became the most important provider of formal childcare in Mexico (Staab & Gerhard, 2010), exposure to the program was limited within municipalities. Panel B of table 1 displays the intensity of the program in each municipality in the baseline sample, measured by the number of enrolled children as a share of the total number of children between 1 to 3 years old in each municipality. Given the rollout process describe earlier, some municipalities expanded the program faster than others, resulting in variation in treatment intensity both across and within municipalities over time. The variation in average program intensity among sample municipalities ranged between approximately 3.9% and 7.8% during the period, increasing with time. At the same time, variation across municipalities at a given point in time ranged between 0.01% and 33.3% in some municipalities. Importantly, all municipalities in the urban sample introduced Estancias Infantiles directly in 2007.

Appendix figure A2 illustrates the geographic variation in treatment exposure across municipalities at two points in time, 2008 and 2012. Estancias Infantiles appears to be relatively evenly distributed across the country, with high concentrations of daycare centers in urban areas.

According to the program rules, anyone who fulfilled the stated requirements was able to start a daycare center. This rollout strategy is likely to have caused a correlation between business performance and exposure to Estancias Infantiles as a result of a higher demand for childcare services in municipalities with high business growth. In appendix table A7, I further examine the rollout of the program across municipalities. Column 1 shows the association between a number of municipality characteristics and the likelihood of being treated by the Estancias Infantiles program, whereas column 2 presents the relation between municipality characteristics and the timing of the introduction of Estancias Infantiles in a specific municipality, conditional on treatment.[8] The results in column 1 suggest that Estancias Infantiles was not randomly introduced across municipalities. The likelihood that the program was introduced in a municipality was negatively correlated with the share of self-employment, but positively correlated with the share of the population with a low income (one minimum wage or lower). This is likely to be a direct effect of the fact that Estancias Infantiles targeted low-income women working in the informal sector. Municipalities with a higher average level of education were more likely to offer Estancias Infantiles, although education level was not associated with the timing of the rollout. Furthermore, the results in column 2 indicates that, controlling for the level of urbanization in a municipality, an increased share of the population working in the tertiary sector was associated with an earlier roll-out of the program. Also, the share of self-employed individuals in a municipality was associated with a later introduction of the program. This was also true for urban municipalities.

Overall, the results suggest that Estancias infantiles was more likely to have been introduced in more dynamic areas with a potentially relatively high demand for daycare services. The potential effects of the rollout of the program on the estimated treatment effects are discussed further in the next section. Importantly, the main estimation strategy relies primarily on the age of the youngest child to identify treatment.

3.3 Microenterprises

The female-run microenterprises are described in appendix tables A5 and A6. The results indicate that women operated primarily their businesses from home or in a specific installation, and were mostly active in the sales or service sector. Appendix table A5 describes enterprises in the national sample. Approximately one out of every four

[8] I measure the timing of the introduction of Estancias Infantiles in a specific municipality, using an index that takes the value 1 January of 2007, 2 in February 2007, and so on. Given the construction of the index, a municipality with a high index number introduced the program at a later stage than did a municipality with a relatively lower number.

enterprises had at least one employee and about 9% of the entrepreneurs employed a paid worker. Among those who did not operate from specific establishment, approximately 31% ran their businesses from home and 13% were walking vendors. About 57% of all microenterprises operated in the sales sector, whereas almost 20% and 19% were active in the service and manufacturing sectors, respectively. Nearly half of the entrepreneurs in the sales sector operate in specific installations, 37% were walking salesmen of miscellaneous items, and 10% sold food by a public road. Approximately 17% of the entrepreneurs in the service sector prepared and served foods in specific plants, and 9% were garment makers. Nearly one out of five were hairdressers. Female entrepreneurs worked 30.5 hours per week, on average.

Appendix table A6 shows that, similarly to the national sample, about half of the urban firms operated in the sales sector, but urban firms were more likely to operate in the service sector, and less likely to operate in the manufacturing sector, than firms in the national sample. The vast majority of the urban firms were individual businesses started by the owner. Approximately 86% of the firm owners used start-up capital when setting up their businesses. The average firm had a physical capital representing a value of 9,700 pesos (approximately USD 490), but half of the entrepreneurs evaluated the physical capital in their firms to be less than 1,000 pesos (USD 50). One out of four urban businesses operated from a plant, and one out of three operated from home. This structure is similar to the results from the national sample. About 17% of the urban firm owners applied for a credit during the last year prior to the survey, and 97% of these also received it. The average loan was 2,100 pesos (USD 100).

The descriptive statistics indicate that the primary reason (52%) for starting a business was to be able to supplement household income. Around 5% of the entrepreneurs stated that theyad started their businesses as a response to good business opportunities, whereas 10% considered that they would receive a higher income being self-employed than from wage employment. Time flexibility has often been identified as one of the main reasons that women become self-employed, so that they can balance housework and market work (Samman et al., 2016). However, only 6% of the entrepreneurs stated that they had started their business in order to receive more flexible working hours.

4 Empirical Strategy

The empirical analysis in this paper builds on the variation in access to Estancias Infantiles caused by program eligibility and rollout. According to the program rules, entrepreneurs with at least one child 1 to 3 years old (under 4 years old) were eligible, whereas families whose youngest child was just over 4 years old were not. The eligibility rules created a variation in access to Estancias Infantiles across entrepreneurs living in the same geographic area. I use this variation to identify the effect of the program on female entrepreneurship and business performance.

My baseline specification to estimate the effect of Estancias Infantiles on female entrepreneurship and business performance is presented in equation (1):

$$Y_{imt} = \beta_0 + \beta_1 A_i + \beta_2 post_t + \beta_3 (A \cdot post)_{it} + X_i \gamma + \delta_t + \alpha_m + \varepsilon_{imt} \tag{1}$$

where i indicates an individual observation; m, municipality; and t, years. Y_{imt} measures a number of different outcomes that proxy firm performance: hours worked during one week, physical capital, and indicators that take the value 1 if (i) the business is located in the entrepreneur's home, (ii) the business has a payed employee, or (iii) the entrepreneur has applied for a financial credit during the last year, and zero otherwise. The likelihood of operating the business from home is a proxy for female mobility, whereas the indicators for having a paid employee and applying for credit, as well as having physical capital, are different measures of firm size. A_i is a dummy variable indicating that the entrepreneur has at least one child 1 to 3 years old, controlling for time-invariant differences in firm performance between the treatment and control groups. The specification also includes year fixed effects, δ_t, which controls for aggregate changes in female entrepreneurship over time, and municipality fixed effects, α_m, which control for time-invariant municipality characteristics that may influence both entrepreneurship and treatment. X_i is a vector of household and individual characteristics covariates, including age, education, marital status, number of adults and children in the household, and household income. *post* is a treatment indicator that takes the value 1 in all time periods after the introduction of Estancias Infantiles in 2007. ε_{imt} is the error term. Standard errors are clustered at the municipality level. The coefficient of interest, β_3, is the DD estimate of the treatment effect. I estimate equation (1) as a linear probability model.

The DD design compares changes in business performance between eligible and ineligible entrepreneurs, before and after the program. An entrepreneur is defined as treated in the DD estimation strategy if she had at least one child 1 to 3 years old and lived in a municipality with at least one daycare center enrolled in Estancias Infantiles. As a result, the DD design estimates treats all women in treated municipalities as eligible if their youngest child is 1 to 3 years old, regardless of whether they received the childcare services; this is the intention-to-treat (ITT) effect.

In addition to the variation in eligibility of Estancias Infantiles caused by the age cutoff, the sequential rollout of the program caused another source of variation in program availability. As described in table 1, the program was gradually rolled-out across municipalities. Some municipalities expanded the program faster than others, creating a variation in treatment exposure over time and among municipalities. As a result, some municipalities had a higher share of eligible children enrolled in the program at a given point in time, than others.

I expand the DD design by incorporating the geographic variation as a second source of variation in treatment exposure. I use the variation in program availability to estimate the difference in business performance between eligible and ineligible entrepreneurs, before and after the introduction of the program (the DD estimate), in municipalities with different levels of program intensity. The DD estimations from high- and low-intensity municipalities can be combined in a tripe-difference model.

I estimate the following DDD specifications:

$$Y_{imt} = \beta_0+\beta_1 A_i+\beta_2 EI_{mt}+\beta_3 post+\beta_4(A_i \cdot post_t)+\beta_5(A_i \cdot \alpha_m)+\beta_6(EI_{mt} \cdot A_i)+X_i\gamma+\delta_t+\alpha_m+\varepsilon_{imt} \tag{2}$$

where again i indicates an individual observation; m, municipality; and t, years. Similar to equatinon (1), equation (2) also includes a variable EI_{mt} that measures program intensity in municipality m and time t. I use both a binary and a continuous measure of treatment intensity. The continuous measure takes a value between zero and 1 that was constructed by dividing the number of children enrolled in the program by the total number of eligible children in each municipality. The binary treatment indicator takes the value 1 if the municipality reached a treatment intensity above the sample mean at ant time during the study period. EI_{mt} is a continuous variable that takes a value between zero and 1, or a treatment indicator that equals 1 if program intensity exceeds the median treatment intensity at any time during the study period. Zero otherwise. The median maximum treatment intensity was 5.5% and 6.7% in the national and urban samples, respectively. β_4, is the DD estimate, whereas β_6 is the DDD estimate that measures the effect of Estancias Infantiles on business performance. Again, I estimate equation (2) using a linear probability model with municipality fixed effects.

4.1 Threats to identification

According to the program rules, Estancias Infantiles targeted children 1 to 3 years old, whose mothers were working in jobs that were not covered by the national social security system, were actively looking for a job, or studying. For entrepreneurs whose youngest child was born just before or after the age cutoff, the eligibility rules were likely to have resulted in a treatment assignment that was as good as random. As discussed earlier, baseline summary statistics reveal that eligible and ineligible entrepreneurs differ significantly in relation to a number of individual and household characteristics (see appendix table A2 for the national sample and appendix table A3 for the urban sample). Among entrepreneurs who live in households with the same income, eligible entrepreneurs are younger, have a higher education level, and are more likely to be married, than those who are ineligible for the program.

The DD design controls for all systematic differences between eligible and ineligible entrepreneurs that do not change over time, including for example differences in cultural expectations and motivations related to childcare obligations and labor force participation between women with young children compared to those with relatively older children. As a result, time-invariant factors are likely to represent an important source of differences in entrepreneurship between eligible and ineligible entrepreneurs, implying that the model control for an important share of unobservables. Additionally, municipality fixed effects control for time-invariant heterogeneity across municipalities such as differences in economic and market opportunities between rural and urban municipalities, which could cause differential trends in business performance between eligible and ineligible entrepreneurs if entrepreneurship among women with young children under 4 years old was limited to urban areas. The time fixed effects control for time-varying, but group invariant, factors such as cultural and economic changes related to female entrepreneurship over time that are similar to all female business owners. Finally, I control for differences in characteristics between treatment and control groups by including a number of covariates in the regression model.

Given the DD design, unobservable and time-varying characteristics remain as a source of omitted variable bias. To produce unbiased estimates, the DD model assumes that in the absence of Estancias Infantiles, differences in business performance between eligible and ineligible entrepreneurs would have been constant over time. This is referred to as the common trends assumption.

I study the pre-treatment trends in outcome levels between eligible and ineligible entrepreneurs to determine whether it is plausible to make the assumption of common trends. Appendix figures A3 and A4 show the trends for the average number of hours worked during a week, as well as the likelihood of having an employee and operating the business from home, in the national and urban samples, respectively. The trend among ineligible entrepreneurs represents the common trend that would have occurred in the absence of treatment—that is, if Estancias Infantiles were never introduced. As mentioned earlier, in 2005 the labor survey was reconstructed. As a result, the survey samples are not completely comparable over time. Despite the reconstruction of the surveys in 2005, I provide pre-trend data between 1995 and 2007 with the understanding that the data are less comparable before and after the changes in the surveys. The restructuring is marked in each figure by a vertical line. The trends appear to be parallel by visual inspection prior to the introduction of Estancias Infantiles, indicating that the assumption of common trends is reasonable. As a result, the assumption that the estimated treatment effect is unbiased becomes more credible.

While these figures provide evidence that supports the common trends assumption in the pre-treatment period, there is no guarantee that the parallel trend would continue in the post-treatment period. A common concern with the DD analysis is that factors unrelated to treatment might affect outcomes differently among eligible entrepreneurs than among ineligible entrepreneurs. If this occurred simultaneously with the introduction of Estancias Infantiles, this could lead to a violation of the common trends assumption by causing differences in the underlying post-treatment trends between the two groups. For example, any government program that aimed to increase business opportunities for women with young children (under 4 years old), such as increased access to microcredits or managerial education, could cause a positive trend in business performance among eligible entrepreneurs even in the absence of Estancias Infantiles. This would result in an upward bias of the DD estimates. Following the same reasoning, any positive economic shock to the economy that coincided with the introduction of Estancais Infantiles, and affected entrepreneurs whose youngest child was under 4 years old more than compared to entrepreneurs whose youngest child was between 4 and 6 years old, would bias the results upward. Though possible, I consider it unlikely that such shocks would have affected eligible and ineligible entrepreneurs differently.

Given the above reasoning, the DD estimation design will produce unbiased estimates assuming that there were no confounding effects, but that it was only the introduction of Estancias Infantiles that changed female business performance. One way of addressing differential trends between eligible and ineligible groups, caused by time-varying confounders, would be to compare the DD estimates in treated municipalities with the same estimate in municipalities where Estancias Infantiles was not implemented, or that had a different level of treatment intensity. This is the DDD estimate. The DDD model controls for potential bias caused by time-varying confounding variables by comparing the difference in outcome variables between eligible and ineligible entrepreneurs, before and after the program introduction, with the corresponding differences in municipalities that offered a different level of program availability. This strategy is based on the assumption that the estimated differences in the control municipalities were exposed to the same confounding factors causing the potential bias of the DD estimates but were not exposed to the treatment of the Estancias Infantiles program to the same extent. This isolates the treatment effect. I provide DDD estimates as a complement to the DD estimates in the next section.

Another potential source of bias arises if the implementation of Estancias Infantiles was correlated with other time-varying determinants of business performance among eligible women. For example, the introduction of Estancias Infantiles might cause eligible, high-ability, and well-educated entrepreneurs from untreated municipalities to move to a treatment municipality in order to access the program. This could contribute to improving the performance of female enterprises and confound the effect of Estancias Infantiles. This would result in an up-ward bias of the treatment effect.

I investigate the correlation between Estancias Infantiles and time-varying, observable, characteristics by estimating the DD model using such characteristics as dependent variables, including age, education level, and marital status. Appendix table A4 provides the results of these covariate-balance regressions. Panel A presents the estimates based on the national sample, whereas Panel B shows the estimates from the urban sample. The results indicate that Estancias Infantiles was not significantly nor materially associated with changes in entrepreneur characteristics in the national sample. For example, the introduction of the program was associated with a decrease in average age among entrepreneurs by 0.026 years, compared with the pre-treatment average age of approximately 32.6 years in the national sample. The association is somewhat higher in the urban sample, although still insignificant both statistically and materially.

Out of the 9 coefficients presented in panel B, 2 are statistically significant at the 10% level. The results suggest that the introduction of Estancias Infantiles was associated with a statistically significant decrease in the number of children in the household (10% at the mean) and an increase in the share of entrepreneurs with a primary education among urban entrepreneurs (4% at the mean). Importantly, the program was not significantly associated with changes in any other education outcome, suggesting that changes in the education level among eligible entrepreneurs is unlikely to drive the result. The association between Estancias Infantiles and the total number of children in the household suggest that the program could have affected fertility. Given the lack absence of association between the program and other entrepreneur characteristics, I consider it unlikely that differential demographic trends among eligible entrepreneurs drive the results discussed in the analysis in this paper.

5 Results

In this section I present the results from the empirical analysis. First, I exploit the variation in access to Estancias Infantiles caused by program eligibility in a difference-in-difference (DD) estimation framework. Furthermore, I examine heterogeneity by education level and business sector. Second, I include the variation in program intensity between and within municipalities, caused by the stepwise rollout of the program, in a triple-difference (DDD) framework. In addition, I demonstrate the robustness of my results to potential bias caused by unobservable confounders by showing that the results are not sensitive to the choice of control groups.

5.1 Difference-in-Difference estimations

Table 2 shows the estimated association between Estancias Infantiles and female entrepreneurship obtained by estimating equation (1). Panel A presents the standard two-period DD estimates based on data from two time periods, the years just before and after the introduction of the program (2006—7), and Panel B displays the results from DD estimations including all time periods between 2005 and 2014. The reduced study period estimates the direct effects of the implementation of the program, whereas the complete study period includes a longer post-treatment period. All results are estimated using the national sample (ENOE).

Overall, I find little evidence that Estancias Infantiles affected female entrepreneurship. The results in both panels A and B reveal a positive but statistically insignificant association between the introduction of the program and the likelihood of operating a business from home (column 2) and hiring a paid worker (column 3). Furthermore, the results in panel A indicate that the number of working hours during a week is insignificant and negatively associated with the program, suggesting that eligible entrepreneurs worked on average 1.4 hours less than ineligible women per week after the introduction of the program (column 1, 5% at the mean). The negative and insignificant association between the program and working hours remains when the effects are followed for a longer time period, although smaller in magnitude.

The results suggest that the standard two-period DD estimates are consistently larger than the multiple-period DD estimates. The two-period DD estimation strategy is limited to measure the direct effects of the program on female business, indicating that the average treatment effect decreases over time.

Table 2: Difference-in-difference (DD) estimations, national sample

Variabes	(1) Hours worked	(2) Operated from home	(3) Had paid worker
Panel A: Two-period DD			
Treatment effect	-1.449	0.011	0.011
	(1.487)	(0.030)	(0.018)
Observations	3,817	3,817	3,817
R-squared	0.019	0.026	0.106
Number of clusters	448	448	448
Mean	31.315	0.316	0.091
Panel B: Multiple-period DD			
Treatment effect	-0.432	0.001	0.005
	(0.894)	(0.016)	(0.010)
Observations	16,437	16,437	16,437
R-squared	0.024	0.022	0.097
Number of clusters	546	546	546
Mean	30.514	0.313	0.087
Year FE	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: All results are based on estimations using the national sample (ENOE). Panel A presents the standard two-period DD estimates, based on two time periods just before and after the introduction of the program (2006–2007). Estimations in panel B extend the two-period DD estimation strategy and include all time-periods between 2005 and 2014. All regressions control for municipality and year fixed effects. Individual and household controls include age, household income, household size, total number of children, and dummies for education level (primary school, secondary school, or post-secondary education) and marital status. Estimations are conditional on being self-employed. The treatment effect is the coefficient of the interaction between an indicator for having a child 3 years or younger and an indicator that takes the value 1 in the post-treatment period, that is, β_3 in the baseline equation (1). Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3 shows the results from estimating equation (1) using the urban sample. Panel A presents the estimated treatment effect using one pre- and one post treatment time period, textendash 2002 and 2008, whereas the estimations in panel B extends the post-treatment time period, to also include survey years 2010 and 2012.

Similarly to the results of my analysis based on the national sample, I find little evidence that Estancias Infantiles was associated with female business performance.

Using both the reduced and complete sample periods, the results reveal insignificant associations between the introduction of Estancias Infantiles and most measures of business performance (column 2–5). The results in column 2 indicate that eligible women were less likely than ineligible entrepreneurs to operate their businesses from home after the introduction of Estancias Infantiles, although the results are insignificant at conventional levels. Column 4 indicates that Estancias Infantiles was associated with a decrease in the likelihood of applying for a credit by 5.7 percentage points (35% at the mean) when estimated using the full number of sample years. The the estimated association between Estancias Infantiles and physical capital of the business is small and imprecisely estimated with large standard errors.

In contrast to the national sample, the urban estimates indicate a positive and statistically significant association between the program and the number of hours worked per week using both the 2-period and multiple-period DD estimation strategies (column 1). In panel A, the introduction of the program was associated with an increase of approximately 6.3 hours dedicated to the business during a week among eligible women, an 18% increase in working hours at the mean. The positive association between Estancias Infantiles and hours worked remains when including the full study period, at a 5% significance level. One plausible explanation for the differences in the effect of Estancias Infantiles on female working hours between the national and urban samples could be that business profitability and performance in urban areas are less likely to be constrained by limited demand. As a result, the financial reward from working additional hours is likely to differ between urban areas and the national average, making urban entrepreneurs more willing to invest additional hours in their business.

Table 3: Difference-in-difference (DD) estimations, urban sample

Variables	(1) Hours worked	(2) Operated from home	(3) Had paid worker	(4) Applied for credit	(5) ln(physical capital)
Panel B: two-period DD					
Treatment effect	6.296**	-0.093	0.020	-0.066	-0.230
	(2.496)	(0.065)	(0.044)	(0.041)	(0.524)
Observations	1,244	1,244	1,244	1,244	1,244
R-squared	0.028	0.023	0.033	0.063	0.070
Number of mun2	80	80	80	80	80
Mean	34.607	0.328	0.078	0.119	6.619
Panel B: multiple-period DD					
Treatment effect	4.244*	-0.051	0.017	-0.057*	0.023
	(2.197)	(0.050)	(0.027)	(0.033)	(0.422)
Observations	2,334	2,334	2,334	2,334	2,334
R-squared	0.034	0.015	0.017	0.043	0.124
Number of clusters	84	84	84	84	84
Mean	31.516	0.323	0.074	0.172	5.463
Year FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Note: All results are based on estimations using the urban sample (ENAMIN). Panel A presents estimates from a standard two-period DD design, including two time periods:—one pre-treatment and one post-treatment (2002 and 2008). Estimations in panel B extend the two-period DD estimation strategy and include all time-periods: 2002 and biannually between 2008 - 2012. All regressions control for municipality and year fixed effects. Individual and household controls include age, household size, total number of children, and dummies for education level (primary school, secondary school or post-secondary education) and marital status. Estimations are conditional on being self-employed. The treatment effect is the coefficient of the interaction between an indicator for having a child 3 years or younger and an indicator that takes the value 1 in the post-treatment period—, that is–β_3 in the baseline equation (1). Standard errors are clustered at municipality level. *** p<0.01, ** p<0.05, * p<0.1.

5.2 Treatment intensity

In this section I expand the DD estimation strategy to also consider the variation in the availability of Estancias Infantiles within and across municipalities, referred to as treatment intensity. In table 4, I start by estimating equation (1) separately for municipalities with low (column 2) and high (column 3) treatment intensity, using the national sample. I define a municipality as high intensity if it reached a treatment intensity that was above the sample median at any time during the study period. Municipalities with a maximum intensity lower than the median are classified as low intensity. Columns 4 and 5 present the results from estimating equation (2), comparing the DD estimates in high-and low-intensity municipalities in a DDD framework. I provide estimates using both a binary (column 4) and continuous (column 5) treatment intensity measure. The continuous intensity variable measures the share of eligible children enrolled in the program at a specific point in time, and the binary measure is defined according to the classification of high- and low-intensity municipalities in columns 2 and 3. The DDD estimates provide a complement to the DD estimation strategy by including the variation in access to the program in the estimation of the treatment effect.

The results suggest that the estimated association between business performance and Estancias Infantiles differs between municipalities with high and low treatment intensity (columns 2–3). Albeit insignificant, the point estimates suggest that the introduction of Estancias Infantiles was associated with a larger decrease in the number of working hours in the low-intensity municipalities than in the high-intensity municipalities. Additionally, I find a positive association between the program and the likelihood of operating the business from home in low-intensity municipalities, which corresponded to a negative association in the high-intensity municipalities. Importantly, most estimated effects are insignificant, with standard errors often nearly twice as large as the estimated treatment effects in both high- and low-intensity municipalities. As a result, size and magnitude of the coefficient should be analyzed interpreted with caution.

The estimated DDD coefficients in column 4 and 5 provide a formal test of the differences in treatment effects between municipalities with different treatment intensities. Overall, the results show that the differences in treatment effects between municipalities with high and low treatment intensity are insignificant. The binary DDD estimates in column 4 suggest that there was negatively and statistically significant association between Estancias Infantiles and the likelihood of running the business from home. This is a direct result of the relation between the estimated treatment effect in municipalities with high and low program intensity discussed earlier. Importantly, the effect appears to be driven by the significant positive estimate in low-intensity municipalities, indicating that other factors but Estancias Infantiles might drive this result. Albeit insignificant, the negative association remains with estimates using a continuous treatment intensity measure.

Importantly, the DDD estimates in column 5 describe the effect of an increase in program intensity by 100%—that is, the effect of enrolling all eligible children in a municipality in Estancias Infantiles. The average treatment effect, however, reached 6.5% during the period. I adjust the coefficients in column 5 to measure the effects of an increase in treatment intensity by the average intensity. The results are presented for each outcome in brackets in column 5. In relation to the likelihood of having a home-based business, the DDD coefficient suggests that treated entrepreneurs are 1.8 percentage points less likely than ineligible entrepreneurs to run the business from home after the introduction of Estancias Infantiles.

The results presented in table 4 indicate that the introduction of Estancias Infantiles was not associated with changes in female entrepreneurship, consistent with the general conclusions of the DD analysis. The continuous DDD estimates are imprecisely estimated with large standard errors. Again, this suggests that the size of the coefficients should be interpreted with caution. Taken together, the different estimation strategies provide strong evidence of a weak association between Estancias Infantiles and female entrepreneurship.

Table 4: DD and DDD estimations including treatment intensity,—national sample

VARIABLES	(1) DD baseline	(2) DD low EI treatment	(3) DD high EI treatment	(4) DDD binary EI treatment	(5) DDD continuous EI treatment
Hours worked					
Child1–3*post	-0.432	-0.968	-0.221		
	(0.894)	(1.512)	(1.086)		
Child1–3*high EI *post				0.747	
				(1.859)	
Child1–3*EI intensity					13.375
					(17.513)
Adjusted DDD coefficient					[0.869]
Observations	16,437	4,755	11,682	16,437	16,437
R-squared	0.024	0.030	0.023	0.026	0.061
Number of clusters	546	273	273	546	546
Operatexd from home					
Child1–3*post	0.001	0.057**	-0.020		
	(0.016)	(0.028)	(0.020)		
Child1–3*high EI *post				-0.077**	
				(0.034)	
Child1–3*EI intensity					-0.279
					(0.327)
Adjusted DDD coefficient					[-0.018]
Observations	16,437	4,755	11,682	16,437	16,437
R-squared	0.022	0.021	0.025	0.024	0.054
Number of clusters	546	273	273	546	546
Had paid worker					
Child1–3*post	0.005	-0.001	0.008		
	(0.010)	(0.014)	(0.013)		
Child1–3*high EI*post				0.009	
				(0.019)	
Child1–3*EI intensity					-0.053
					(0.184)
Adjusted DDD coefficient					[-0.003]
Observations	16,437	4,755	11,682	16,437	16,437
R-squared	0.097	0.092	0.100	0.098	0.124
Number of clusters	546	273	273	546	546
Year FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Note: The results from estimating equation (1) (column 1–3) and equation (2) (column 4–5) using the national sample. Columns 1 and 2 present the estimated effect of Estancias Infantiles (EI) on firm performance among municipalities with a low and high treatment intensity. The DDD estimates in column 4 use a binary treatment indicator and in column 5 a continuous treatment intensity. All regressions control for municipality and year fixed effects, age, household income, household size, total number of children, and dummies for education level and marital status. The coefficients in column 5 describe the effect of an increase in program intensity from zero to 100%, whereas the coefficients in brackets measure the effect of an increase in treatment intensity represented by the sample mean. Estimations are conditional on being self-employed. Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5 replicates the estimations in table 4 using the urban sample. Similarly to the results estimated using the national sample, the results from the urban sample show differences in the estimated treatment effects of Estancias Infantiles between municipalities with high and low program intensity. Overall, the estimated associations between the program and female entrepreneurship are statistically insignificant and support the conclusion that Estancias Infantiles has little or no effect on female business performance. This reasoning is supported by the DD as well as the DDD estimates.

In panel B, the results indicate significant associations between Estancias Infantiles and the number of hours worked, as well as the likelihood of applying for credit. The results suggest that factors other than Estancias Infantiles affected business performance among entrepreneurs with young children, which drove the average treatment effects of the program presented earlier. Again, the estimated DDD coefficients have large standard errors and should be interpreted with caution.

Table 5: DD estimations including treatment intensity - urban sample

Variables	(1) Hours worked	(2) Operated from home	(3) Paid workers	(4) Applied for credit	(5) ln(physical capital)
Panel A: DD Baseline					
Child1–3*post	4.244*	-0.051	0.017	-0.057*	0.023
	(2.197)	(0.050)	(0.027)	(0.033)	(0.422)
Observations	2,334	2,334	2,334	2,334	2,334
R-squared	0.034	0.015	0.017	0.043	0.124
Number of clusters	84	84	84	84	84
Panel B: DD low intensity					
Child1–3*post	4.796	-0.056	-0.052	-0.034	-0.048
	(3.245)	(0.067)	(0.049)	(0.063)	(0.613)
Observations	587	587	587	587	587
R-squared	0.035	0.049	0.044	0.073	0.106
Number of clusters	42	42	42	42	42
Panel C: DD high intensity					
Child1–3*post	4.091	-0.039	0.039	-0.078**	0.014
	(2.780)	(0.067)	(0.034)	(0.038)	(0.547)
Observations	1,747	1,747	1,747	1,747	1,747
R-squared	0.038	0.017	0.017	0.049	0.134
Number of clusters	42	42	42	42	42
Panel D: DDD binary treatment					
Child1–3*high EI*post	-0.705	0.017	0.091	-0.044	0.062
	(4.238)	(0.094)	(0.060)	(0.073)	(0.815)
Observations	2,334	2,334	2,334	2,334	2,334
R-squared	0.037	0.024	0.023	0.054	0.128
Number of clusters	84	84	84	84	84
Panel E: DDD continuous treatment					
Child1–3*EI intensity	-63.825	1.732	0.138	-0.234	-6.419
	(59.045)	(1.486)	(0.490)	(0.662)	(9.332)
Adjusted DDD coefficient	[-5.170]	[0.140]	[0.011]	[-0.019]	[-0.520]
Observations	2,334	2,334	2,334	2,334	2,334
R-squared	0.078	0.067	0.059	0.081	0.186
Number of clusters	84	84	84	84	84
Municipality & year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Mean treatment intensity	0.063	0.063	0.063	0.063	0.063

Notes: The results from estimating equation (1) (panel A–C) and equation (2) (panel D–E) using the urban sample. Panel B and C present the estimated treatment effect among municipalities with low and high treatment intensity, whereas panels D and E give the DDD estimates. The estimates in panel C use a binary treatment indicator and panel E is based on a continuous treatment intensity. All regressions control for municipality fixed effects, year fixed effects, age, household size, total number of children, and dummies for education level and marital status. Additionally, the DDD estimations control for treatment intensity. The coefficients in column 5 describe the effect of an increase in program intensity from zero to 100%, whereas the coefficients in brackets measure the effect of an increase in treatment intensity represented by the sample mean. Estimations are conditional on being self-employed. Clustered standard errors at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

5.3 Local effects

It is possible that some groups could have benefited more than others from Estancias Infantiles. One potential reason could be that childcare obligations constrain business performance differently for women who operate in different business sectors. As a result, the expected efficiency gains from enrolling children in Estancias Infantiles might differ among sectors, causing the expected benefits from the program to vary among entrepreneurs. For example, the expected benefits from formal childcare services might be larger for domestic workers who offer services in clients' homes, than for entrepreneurs who operate small grocery shops in their own homes. Another potential reason for variations in the expected benefits from the program could be that the ability to take advantage of the efficiency gains offered by formal daycare might also vary among entrepreneurs. Recent literature has shown some evidence that managerial education could have a positive effect on firm growth among female-run businesses (Campos et al., 2017). This could imply that the effect of childcare services would potentially be higher among entrepreneurs with a relatively higher education level, and managerial training.

Appendix tables A8 and A9 present the effects of Estancias Infantiles on entrepreneurship for different subgroups in the national and urban samples, using the multiple-period DD estimation strategy. The first and second panels estimate the effects of childcare among entrepreneurs in the service and sales sectors respectively, whereas the last panel describe the estimated effects among women with at least a secondary education. Overall, I find no evidence that Estancias Infantiles affected female entrepreneurship for any of the subgroups in either the national or urban sample. The results suggest a positive and significant association between the program and the number of worked hours among those with a relatively higher education level in the national sample (5% level) and among entrepreneurs in the urban service sector (10% level). The results indicate that the previously estimates positive association between Estancias Infantiles and the number of hours worked in the urban sample was partly driven by entrepreneurs in the service sector. In the national sample, the results suggest that the Estancias Infantiles could have affected the labor supply among entrepreneurs who were initially better off.

5.4 Selection into self-employment

Besides potentially affecting entrepreneurship among female business owners, increased access to subsidized childcare is likely to affect the selection into and out of self-employment. Previous literature suggests that self-employment represents a flexible labor market opportunity that makes it possible for women to balance work and family life (Marshall & Flaig, 2014; Hundley, 2000, 2001; Allen & Curington, 2014; World Bank, 2012; Boden, 1996; Hamilton, 2000). Broad evidence from developing countries indicates that childcare obligations represent one of the main reasons for not taking a job in the formal sector (Cassirer & Addati, 2007). Furthermore, a number of studies find that increased access to formal childcare services has a positive effect on female labor force participation (Baker et al., 2008; Lefebvre & Merrigan, 2008; Hallman et al., 2005; Attanasio & Vera-Hernández, 2004).[9] By alleviating the trade-off between childcare obligations and labor faced by women, formal childcare could affect not only the decision to enter the labor market but also the type of jobs that women do.

The effect of subsidized childcare on female self-employment and wage employment is an important consideration when interpreting the effect of Estancias Infantiles on female entrepreneurship. On one hand, increased availability of childcare could mean that women with relatively low business profitability, who face better conditions as wage employees, would transition from self-employment into a wage job. As a result, only relatively more profitable entrepreneurs would remain in the self-employment sector. On the other hand, increased availability of subsidized childcare could contribute to an increased likelihood that low-productivity entrepreneurs, who previously did not enter self-employment because of low expected profits in in the face of the cost of childcare, would find it profitable to start a business. In both scenarios, any results, or lack thereof, could be due to changes in willingness to start a business across different groups, or changes in the composition of entrepreneurs in the sector and not primarily due to changes in entrepreneurial behavior among existing firm owners.

In table 6, I test for selection into the self-employment sector, estimating equation (1) using the complete sample of women whose youngest child was 1 to 6 years old. The results indicate that Estancias Infantiles was not associated with the average likelihood of being self-employed. The results are similar for estimations both conditional and unconditional on working. In the last column, I use quarterly data from the ENOE

[9] Other studies find small or insignificant effects (Manley & Vásquez Lavín, 2013; Berlinski & Galiani, 2007; Havnes & Mogstad, 2011)

survey to construct a panel based on the last quarter before the introduction of the program, the fourth quarter of 2006, and the consecutive five quarters. The panel makes it possible to follow individual entrepreneurs one year after the introduction of Estancias Infantiles. Column 5 describes the results from estimating equation (1) with individual fixed effects. The fixed effects estimates confirm the previous findings in the table providing small and statistically insignificant association between Estancias Infantiles and the likelihood of being self-employed, indicating that individuals did not exit or enter self-employment as a result of Estancias Infantiles.

The results in table 6 suggest that Estancias Infantiles did not affect the likelihood of being self-employed among women with young children. Furthermore, as discussed earlier in this section, the program could potentially have resulted in movement of entrepreneurs with different individual characteristics and entrepreneurial opportunities in and out of the self-employed sector. In table A4, I estimate the change in individual characteristics among self-employed women before and after the introduction of Estancias Infantiles using both the national sample (palen A) and the urban sample (panel B). The results suggest that Estancias Infantiles was not associated with any significant changes in the composition of self-employed women in the national sample, suggesting that the average treatment effects were not driven by differential demographic trends among eligible and ineligible entrepreneurs. However, in the urban sample, the estimated association between self-employment and individual characteristics shows a significant association between self-employment and the likelihood of having at least primary education and the number of children in the household. In relation to the lack of average treatment effects on female entrepreneurship, the results suggest that an increase in the average education level among entrepreneurs did not translate into changes in business performance.

Table 6: The probability of being self-employed

Variables	(1) DD (2006–7)	(2) DD (2005–14)	(3) DD Low	(4) DD High	(5) DD Panel (2006–7)
Panel A: Unconditional on working					
Treatment effect	-0.006	0.003	0.003	0.003	-0.004
(multiple period DD)	(0.010)	(0.005)	(0.008)	(0.007)	(0.008)
Observations	24,919	110,157	34,081	76,076	51,019
R-squared	0.035	0.036	0.036	0.036	0.002
Number of clusters	516	533	264	269	20,977
Panel B: Conditional on working					
Treatment effect	-0.011	0.000	0.009	-0.002	0.016
(multiple period DD)	(0.023)	(0.012)	(0.020)	(0.015)	(0.016)
Observations	8,488	38,485	11,006	27,479	17,935
R-squared	0.032	0.030	0.028	0.031	0.004
Number of clusters	496	533	264	269	9,507
Year FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	—
Individual FE	No	No	No	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: The results from estimating equation (1) using the complete sample of women whose youngest child was 1 to 6 years old. Panel A presents the estimated likelihood of being self-employed among all women in the sample, and panel B displays the estimated likelihood of self-employment conditional on working. All regressions control for municipality and year fixed effects, as well as household and individual characteristics such as education level, age, marital status, household size, and total number of children. Column 5 gives the results from estimating equation (1), including individual fixed effects, using a panel data that follows women one quarter before the introduction of the program and one year post treatment. Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

5.5 Robustness

To make sure that the estimated association between Estancias Infantiles and female entrepreneurship is not sensitive to the definition of the control group, in this section I estimate the baseline DD equation (1), using alternative definitions of the control group. The results are presented in appendix tables A10 and A11.

Table A10 present the multiple-period DD estimates based on the national sample.

The baseline estimates (table 2, panel B, columns 1–3) are presented in column 1. Column 2 gives the results from estimating the baseline DD equation using a control group that excludes entrepreneurs whose youngest child was 4 years old. This definition eliminates the possibility that children remained in the program for the transition period after they turned 4 years old, which would result in lower-bound estimates. In column 3, I exclude all entrepreneurs in the baseline treatment group who also had children between 4 and 6 years old. As a result, mothers whose youngest children were eligible for the Estancias Infantiles program, but whose older siblings were between 4 and 6 years old (similar to households in the control group), are likely to still face childcare obligations related to their older siblings who were enrolled in preschool. This would underestimate the effect of formal childcare on female entrepreneurship. In column 4 I restrict the sample to entrepreneurs whose youngest child is between 3 and 4 years old, with the aim of increasing the comparability between women in the treatment and control groups. Women in the treatment group with very young children around 1 year old may face very different childcare obligations than a mother in the control group whose youngest child is 6 years old. Finally, the last column presents the results from estimating equation 1 including all municipalities in the national sample in a staggered-roll out DD estimation. This estimation includes municipalities who introduced Estancias Infantiles after 2007 or not at all. In addition to increasing the variation in treatment intensity across municipalities, the inclusion of non-treated municipalities results adds a control group of municipalities that were untreated during the study period.

Overall, the results suggest that the estimated effects are robustly insignificant to alternative definitions of the control group. Table A10 indicates that the estimated associations between increased access to affordable childcare services and business performance remain statistically insignificant and for all the alternative specifications based on the national sample.

Appendix table A11 replicates the estimations in appendix table A10 using the urban sample. Again, the findings are similar across columns, suggesting that the results are not sensitive to the choice of control group. The estimated associations between Estancias Infacties and the number of hours worked and the likelihood of applying for a credit are no longer statistically significant when estimated using a sample of entrepreneurs whose youngest child is between 3-4 years old (column 4). This could potentially be explained by the small number of observations remaining in this sample.

On the other hand, the results suggest a negative and precisely estimated association between the program and the likelihood of operating the business from home for this sample. Overall, the results suggest that few exceptions of marginally significant coefficients, the estimated association between Estancias Infantiles and firm performance remains insignificant. Again, the overall results indicate that the estimated treatment effects are robust to different definitions of control groups.

The staggered roll-out DD estimates are not presented for the urban sample as all municipalities introduced Estancias Infantiles directly in 2007.

5.5.1 Individual panel

I complement the robustness analysis by estimating the association between Estancias Infantiles and female entrepreneurship using individual panel data. As mentioned earlier, the ENOE survey is a rotating panel which provides the opportunity to follow an individual for 5 quarters. I estimate the association between Estancias Infantiles and firm performance by following individuals during a period starting one quarter before the introduction of the program in 2007 and up to 1 year after the program was introduced. The estimation strategy builds on Equation 1, but adds individual fixed effects and quarter fixed effects.

The 2-period DD estimation strategy, table 2, produces unbiased estimates given the assumption that all unobservable cofounders are time-invariant during the study period. While the covariate-balance regressions, presented in table A4, provide convincing evidence that the composition of entrepreneurs is stable over time with regards to observable characteristics, variation in unobservable characteristics remains a potential source of bias. For example, underlying preferences and abilities among eligible entrepreneurs might change between the pre- and post-periods due to changes in the group of eligible entrepreneurs, causing different trends in firm performance between eligible and ineligible women that are unrelated to Estancias Infantiles. Individual fixed effects control for all such unobservable time-invariant characteristics. As a result, unobservable and time-varying confounders remain the only potential source of bias.

Table A12 presents the results from estimating equation (1), including individual fixed effects. The results suggest that Estancias Infantiles was not significantly associated with entrepreneurship. The results support earlier analysis and suggest that the DD analysis presented in tables 3 and 2 is robust to the inclusion of individual fixed effects. The results indicate that the DD estimates were not driven by time-invariant confounders.

6 Conclusion

Microentrepreneurship represent an important source of income for many women in developing countries (Razavi, 2012). Earlier research indicates that female entrepreneurs often underperform their male counterparts in relation to a number of indicators such as business profits and sales. This gender gap is substantial and has been documented in a large number of developing countries across the globe (Fiala, 2018; Hardy & Kagy, 2018).

In this paper, I have examined the importance of childcare obligations as a hindrance for female microentrepreneurship. Despite previous evidence that childcare obligations could represent a key constraint for female microentrepreneurship (Bruhn, 2009; Fitzpatrick & Delecourt, forthcoming.), I find little evidence that increased access to formal and subsidized childcare affects female entrepreneurship. Overall the results suggest that Estancias Infantiles had no effect on business performance among self-employed women in vulnerable households. The findings are consistent across both the DD and DDD estimation strategies. I do find tentative evidence that the childcare program was associated with an increase in hours worked among entrepreneurs among urban entrepreneurs, but this association is not consistently significant across specifications.

No previous study has estimated the effect of increased access to subsidized childcare on female-run businesses. However, in earlier work,Calderon (2014) found that increased availability of Estancias Infantiles was not significantly associated with hours worked among eligible women who were working in the previous period. Importantly, the study did not specifically focus on self-employed but rather on mothers of young children in general. Furthermore, evidence from Chile suggested that the expansion of a public daycare program was associated with a decrease in working hours among eligible women Medrano (2009). On the other hand, another study in Chile showed that the offer to participate in an after-school program was associated with an increase in hours worked among mothers, albeit insignificant (Martínez & Perticará, 2017). (Berlinski et al., 2011) found that women worked 7.8 more hours per week as a consequence of their youngest child attending pre-school in Argentina.

One explanation for the lack of results could be a relatively weak identification of treated households for the intention-to treat (ITT) estimates. Although the Estancias Infantiles program quickly became the most important provider of childcare services in Mexico for children 1 to 3 years old, the average intensity of the treatment reached only about 6.5% on average in a municipality. The ITT estimate measures the average treatment effect among all women who were eligible for Estancias Infantiles, implying that the treatment intensity might have been too weak. Although the DDD design adjust for the overall low treatment intensity in the DD estimations, the overall low treatment intensity in all municipalities might still be too weak to capture any potential treatment effects.

Another plausible explanation is that increased access to subsidized childcare services made it possible for women to accept wage jobs in the formal or informal sector that were less flexible and hard to combine with childcare obligations. A vast body of literature suggests that increased access to subsidized childcare has a positive effect on female labor market participation (Baker et al., 2008; Halim et al., 2019) and that women who entered the labor market were likely to obtain more jobs in the formal sector (Calderon, 2014). I provide evidence that my estimated treatment effects are not likely to have been driven by the entry and exit of eligible women to self-employment as a result of the implementation of Estancias Infantiles. I find that the introduction of the program was not associated with the average likelihood of being self-employed. Furthermore, I show that the program was not significantly associated with changes in the composition of self-employed individuals, suggesting that there are no confounding effects.

Informal self-employment has often been described as a strategy to combine childcare obligations and market employment. According to the analysis, entrepreneurs worked on average more than 30 hours per week before the introduction of Estancias Infantiles. This could suggest that self-employed mothers were bringing their children to work prior to the introduction of Estancias Infantiles. Another plausible explanation is that Estancias Infantiles substituted for already existing formal or informal childcare arrangements. This study shows that while many women decided to start their own businesses in order to enjoy time flexibility and be able to balance household and business obligations, the main reason was to supplement household income.

In the light of worldwide social norms that place a large share of the responsibility for childcare and housework on women, future policymakers should pay greater attention to the necessity for many women to balance household and market work. This is becoming more and more important as a result of shifts in women's labor market engagement, improvements in girls' education, growth in migration and urbanization, and changes to family structure, all of which have contributed to putting childcare on the policy agenda (Samman et al., 2016). In this context, self-employment could represent an important source of income for women in developing counties. Finding the key to improving opportunities for female entrepreneurs is an important goal in order to increase female income and empowerment.

References

Allen, W. D., & Curington, W. P. (2014). The self-employment of men and women: What are their motivations? *Journal of Labor Research*, *35*(2), 143–161.

Ángeles, G., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P., & Seira, E. (2011). *Evaluación de impacto del programa estancias infantiles para apoyar a madres trabajadoras* (Informe Final de Impacto). México: Instituto Nacional de Salud Pública.

Attanasio, O., & Vera-Hernández, M. (2004). *Medium-and long run effects of nutrition and child care: Evaluation of a community nursery programme in rural Colombia* (Working Paper EWP04/06). London: Institute for Fiscal Studies.

Baker, M., Gruber, J., & Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, *116*(4), 709–45.

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, *7*(1), 22–53.

Bastian, G., Iacopo, B., Goldstein, M., & Montalvao, J. (2018). *Short-term impacts of improved access to mobile savings, with and without business training: Experimental evidence from Tanzania* (Working Paper No. 478). Washington, DC, USA: Center for global development.

BenYishay, A., & Pearlman, S. (2014). Crime and microenterprise growth: Evidence from Mexico. *World Development*, *56*, 139–52.

Berlinski, S., & Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, *14*(3), 665–80.

Berlinski, S., Galiani, S., & Mc Ewan, P. J. (2011). Preschool and maternal labor market outcomes: evidence from a regression discontinuity design. *Economic Development and Cultural Change*, *59*(2), 313–44.

Boden, R. J. (1996). Gender and self-employment selection: An empirical assessment. *Journal of Socio-Economics*, *25*(6), 671–682.

Bruhn, M. (2009). *Female-owned firms in Latin America: Characteristics, performance, and obstacles to growth* (Policy Research Working Paper No. 5122). Washington, DC, USA: World Bank.

Bruhn, M., & Zia, B. (2013). Stimulating managerial capital in emerging markets: The impact of business training for young entrepreneurs. *Journal of Development Effectiveness*, *5*(2), 232–66.

Calderon, G. (2014). *The effects of child care provision in Mexico* (Working Paper No. 2014-07). México: Banco de México.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., & Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in West Africa. *Science*, *357*(6357), 1287–90.

Cassirer, N., & Addati, L. (2007). *Expanding women's employment opportunities: Informal economy workers and the need for childcare*. Geneva, Switzerland.

Crépon, B., Devoto, F., Duflo, E., & Parienté, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics*, *7*(1), 123–50.

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *Quarterly Journal of Economics*, *123*(4), 1329–72.

Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, *6*(2), 1–31.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, *50*(4), 1051–79.

Dupas, P., & Robinson, J. (2013). Savings constraints and microenterprise development: Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*, *5*(1), 163–92.

Fafchamps, M., McKenzie, D., Quinn, S. R., & Woodruff, C. (2011). *When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana* (Working Paper No. 17207). Cambridge, MA, USA: National Beureau of Economic Research.

Fiala, N. (2018). Returns to microcredit, cash grants and training for male and female microentrepreneurs in Uganda. *World Development*, *105*, 189–200.

Fitzpatrick, A., & Delecourt, S. (forthcoming.). *Childcare matters: Female business owners and the baby-profit gap.* (Management Science)

Fitzpatrick, M. D. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten. *Journal of Labor Economics*, *28*(1), 51–85.

Friedson-Ridenour, S., & Pierotti, R. S. (2019). Competing priorities: Womens microenterprises and household relationships. *World Development*, *121*, 53–62.

Halim, D., Johnson, H. C., & Perova, E. (2019). *Preschool availability and female labor force participation: Evidence from indonesia.* Washington, DC, USA: World Bank.

Hallman, K., Quisumbing, A. R., Ruel, M., & de la Briere, B. (2005). Mothers' work and child care: Findings from the urban slums of Guatemala City. *Economic Development and Cultural Change*, *53*(4), 855–85.

Hamilton, B. H. (2000). Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political Economy*, *108*(3), 604–31.

Hardy, M., & Kagy, G. (2018). Mind the (profit) gap: Why are female enterprise owners earning less than men? *AEA Papers and Proceedings*, *108*, 252–55.

Havnes, T., & Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, *3*(2), 97–129.

Hundley, G. (2000). Male/female earnings differences in self-employment: The effects of marriage, children, and the household division of labor. *Industrial & Labor Relations Review*, *54*(1), 95–114.

Hundley, G. (2001). Domestic division of labor and self/organizationally employed differences in job attitudes and earnings. *Journal of Family and Economic Issues*, *22*(2), 121–39.

INEGI. (2019). *Encuesta nacional de ocupación y empleo, Nota sobre los ingresos, sueldos y salarios de la población ocupada* (Tech. Rep.). Mexico: National Institute of Statistics and Geography (INEGI). Retrieved from `https://www.inegi.org.mx/contenidos/programas/enoe/15ymas/doc/enoe_nota_ingresos.pdf` (accessed 4 April 2021)

Jayachandran, S. (2020). *Microentrepreneurship in developing countries* (Working Paper No. 26661). Cambridge, MA, USA: National Beureau of Economic Research.

Karlan, D., & Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and statistics*, *93*(2), 510–27.

Karlan, D., & Zinman, J. (2011). Microcredit in theory and practice: Using randomized credit scoring for impact evaluation. *Science*, *332*(6035), 1278–84.

Klapper, L. F., & Parker, S. C. (2011). Gender and the business environment for new firm creation. *The World Bank Research Observer*, *26*(2), 237–257.

Lefebvre, P., & Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, *26*(3), 519–48.

Manley, J., & Vásquez Lavín, F. (2013). *Childcare availability and female labor force participation: An empirical examination of the Chile Crece Contigo program* (Working Papers No. 2013-03). Townson, MD, USA: Townson University, Department of Economics.

Marshall, M. I., & Flaig, A. (2014). Marriage, children, and self-employment earnings: An analysis of self-employed women in the US. *Journal of Family and Economic Issues*, *35*(3), 313–22.

Martínez, C., & Perticará, M. (2017). Childcare effects on maternal employment: Evidence from Chile. *Journal of Development Economics*, *126*, 127–37.

McKenzie, D., & Woodruff, C. (2013). What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, *29*(1), 48–82.

Medrano, P. (2009). *Public day care and female labor force participation: Evidence from Chile* (Working Paper No. 306). Santiago: University of Chile.

OECD. (2014). *Balancing paid work, unpaid work and leisure.* Retrieved from `https://www.oecd.org/gender/data/balancingpaidworkunpaidworkandleisure.htm` (accessed May 18 2020)

Potter, J., Marchese, M., Feldman, M., Kemeny, T., Lawton-Smith, H., & Pike, A. (2013). *The local dimension of SME and entrepreneurship issues and policies in Mexico* (OECD Local Economic and Employment Development (LEED) Papers No. 2013/14). Paris: OECD Publishing.

PROFECO, P. (2014). *Mamás multitarea* (Revista del Consumidor No. 447).

Razavi, S. (2012). World development report 2012: Gender equality and development. *Development and Change*, *43*(1), 423–437.

Rijkers, B., & Costa, R. (2012). *Gender and rural non-farm entrepreneurship* (Research Working Paper No. 6066). Washington, DC, USA: World Bank.

Samman, E., Presler-Marshall, E., Jones, N., Bhatkal, T., Melamed, C., Stavropoulou, M., & Wallace, J. (2016). *Women's work: Mothers, children and the global childcare crisis* (Report 2016). London: Overseas Development Insstitute. Retrieved from `https://bettercarenetwork.org/sites/default/files/Women%E2%80%99s%20work-%20Mothers%2C%20children%20and%20the%20global%20childcare%20crisis.pdf` (accessed May 2020)

Secretaría de Gobernación. (2007). *Reglas de operacion del programa de estancias infantiles para apoyar a madres trabajadoras, para el ejercicio fiscal 2008* (Diario Oficial). Mexico: Secretaria de Desarrollo Social.

Staab, S., & Gerhard, R. (2010). *Childcare service expansion in Chile and Mexico: For women or children or both?* (Gender and Development Programme Paper No. 10). Geneva Switzerland: United Nations Research Institute for Social Development.

Staab, S., & Gerhard, R. (2011). Putting two and two together? Early childhood education, mothers' employment and care service expansion in Chile and Mexico. *Development and Change*, *42*(4), 1079–1107.

Valdivia, M. (2015). Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Peru. *Journal of Development Economics*, *113*, 33–51.

World Bank. (2012). *World development report 2012, Gender equality and development* (Word Development Report). Washington, DC, USA: World Bank.

Appendix

Tables

Table A1: Summary statistics for the national and urban samples

	Panel A: National sample				Panel B: Urban sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable	N	Mean	Min.	Max.	N	Mean	Min.	Max.
Age	16 437	32.532	16	63	2 334	32.976	17	63
		(6.202)				(6.230)		
No education	16 437	0.021	0	1	2 334	0.007	0	1
		(0.146)				(0.087)		
Primary school	16 437	0.222	0	1	2 334	0.192	0	1
		(0.416)				(0.394)		
Secondary school	16 437	0.348	0	1	2 334	0.293	0	1
		(0.476)				(0.455)		
> secondary education	16 437	0.406	0	1	2 334	0.465	0	1
		(0.491)				(0.498)		
Married	16 437	0.904	0	1	2 334	0.851	0	1
		(0.293)				(0.355)		
Single	16 437	0.035	0	1	2 334	0.059	0	1
		(0.186)				(0.235)		
Widow	16 437	0.008	0	1	2 334	0.015	0	1
		(0.093)				(0.124)		
Divorced	16 437	0.050	0	1	2 334	0.073	0	1
		(0.219)				(0.261)		
Adults in HH	16 437	2.260	1	11	2 334	2.211	1	7
		(0.811)				(0.800)		
Tot. children	16 437	2.260	1	9	2 334	2.203	1	7
		(1.013)				(0.977)		
Income	16 437	3 118.677	100	129 000				
		(4 189.887)						
Rural	16 437	0.246	0	1				
		(0.430)						

Notes: Summary statistics for the national (columns 1–4) and urban (columns 5–8) samples. Standard deviations appear in parenthesis. For each sample, the table presents means (columns 2 and 6), and minimum and maximum values (columns 3–4 and 7–8). The sample includes female entrepreneurs whose youngest child is 1 to 6 years old. Because of the structure of the questions in the 2002 urban survey, it was not possible to calculate a monthly income for urban households.

Table A2: Pre-treatment summary statistics for the national sample (ENOE)

Variable	(1) All	(2) Treatment group (child 1–3)	(3) Control group (child 4–6)	(4) Difference
Age	32.592	31.399	34.218	2.819***
	(6.118)	(5.956)	(5.961)	
No education	0.029	0.025	0.035	0.009*
	(0.170)	(0.158)	(0.184)	
Primary school	0.264	0.245	0.291	0.046***
	(0.441)	(0.430)	(0.454)	
Secondary school	0.314	0.315	0.313	0.002
	(0.464)	(0.464)	(0.464)	
> secondary education	0.390	0.413	0.359	-0.053***
	(0.487)	(0.492)	(0.480)	
Married	0.905	0.919	0.887	-0.031***
	(0.292)	(0.272)	(0.315)	
Single	0.033	0.028	0.039	0.011*
	(0.178)	(0.165)	(0.195)	
Widow	0.010	0.010	0.010	0.000
	(0.103)	(0.103)	(0.104)	
Divorced	0.050	0.041	0.061	0.020***
	(0.218)	(0.200)	(0.240)	
Adults in HH	2.265	2.229	2.314	0.085***
	(0.841)	(0.103)	(0.871)	
Tot. children	2.326	2.370	2.265	-0.105***
	(1.060)	(1.095)	(1.007)	
Income	2 899.714	2 867.153	2 944.614	77.461
	(3 935.312)	(3 647.93)	(4 301.088)	
Rural	0.230	0.223	0.240	0.017
	(0.421)	(0.416)	(0.427)	
Observations	3 684	2 125	1 559	

Notes: Baseline summary statistics for the national sample in 2006, including the whole sample (column 1) as well as the treatment (column 2) and control (column 3) groups. The treatment group consists of female entrepreneurs whose youngest child is 1–3 years old and the control group consists of entrepreneurs whose youngest child is 4–6 years old. Column 4 presents differences in means between the treatment and control groups. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Pre-treatment summary statistics for the urban sample (ENAMIN)

Variable	(1) All	(2) Treatment group (child 1–3)	(3) Control group (child 4–6)	(4) Difference
Age	33.309	32.253	34.650	2.396**
	(5.968)	(5.692)	(6.049)	
No education	0.014	0.008	0.023	0.015
	(0.087)	(0.089)	(0.151)	
Primary school	0.258	0.205	0.325	0.119***
	(0.437)	(0.404)	(0.469)	
Secondary school	0.204	0.229	0.172	-0.057**
	(0.403)	(0.421)	(0.378)	
> secondary education	0.421	0.452	0.380	-0.072**
	(0.494)	(0.498)	(0.486)	
Married	0.838	0.851	0.822	-0.028
	(0.368)	(0.356)	(0.382)	
Single	0.038	0.045	0.038	-0.007
	(0.202)	(0.209)	(0.193)	
Widow	0.026	0.035	0.016	-0.019
	(0.161)	(0.184)	(0.128)	
Divorced	0.091	0.067	0.122	0.054***
	(0.288)	(0.251)	(0.327)	
Adults in HH	2.173	2.087	2.283	0.196***
	(0.753)	(0.669)	(0.836)	
Tot. children	2.359	2.485	2.200	-0.285***
	(1.040)	(1.082)	(0.963)	
Observations	817	457	360	

Notes: Baseline summary statistics for the urban sample (ENAMIN) in 2002, including the whole sample (column1) as well as the treatment (column 2) and control (column 3) groups the treatment group (column 2) and the control (column 3) group. The treatment group consists of female entrepreneurs whose youngest child is 1–3 years old and the control group consists of entrepreneurs whose youngest child is 4–6 years old. Column 4 presents differences in means between the treatment and control groups. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Correlation between municipality characteristics and the rollout of Estancias Infantiles

	(1)	(2)
Variables	EI treatment	Start date EI
Area	0.000**	-0.000
	(0.000)	(0.000)
Urban municipality	-0.022	13.089***
	(0.117)	(4.076)
Semiurban municipality	0.181***	-2.268
	(0.027)	(1.637)
Mixed municipality	0.188***	-6.566***
	(0.042)	(1.292)
Illiteracy (>15 yrs)	0.478***	10.837
	(0.153)	(13.970)
Avg. years schooling	0.093***	-1.150
	(0.022)	(1.022)
Occupation rate	0.000	-0.000***
	(0.000)	(0.000)
Primary sector	-0.191**	0.811
	(0.092)	(7.171)
Tertiary sector	0.290	-21.052**
	(0.240)	(8.217)
Share self-employed	-0.360*	19.768**
	(0.182)	(8.075)
Income below 1 minimum wage	0.145**	-1.373
	(0.058)	(11.028)
Observations	2,442	1,442
R-squared	0.472	0.235
State FE	Yes	Yes

Notes: Column 1 presents the estimated correlation between municipality characteristics and the likelihood that it introduced Estancias Infantiles (EI) based on all municipalities in Mexico. Column 2 presents the association between municipality characteristics and the timing of the introduction of EI, conditional on municipalities being treated. The timing of the introduction of the program is measured by an index that takes the value 1 in January of 2007, 2 in February 2007, and so on. In urban municipalities, more than 50% of the population lives in localities with ≥100,000 individuals; in semi-urban municipalities more than 50% of the population lives in localities with 15,000–100,000 inhabitants; and in rural municipalities more than 50% of the population lives in localities with less than 2,500 inhabitants. Mixed municipalities do not have a clear urban or rural profile. All estimations include state fixed effects, and standard errors are clustered at the state level. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Local effects of Estancias Infantiles on firm performance, by firm sector and education level, national sample (ENOE)

	(1)	(2)	(3)
Variables	Hours worked	Operated from home	Had paid worker
Panel A: Services			
Treatment effect (multiple period DD)	1.541	-0.003	0.006
	(1.732)	(0.050)	(0.032)
Observations	2,957	2,957	2,957
R-squared	0.039	0.050	0.139
Number of clusters	230	230	230
Panel B: Sales			
Treatment effect (multiple period DD)	-1.111	0.009	-0.003
	(1.340)	(0.019)	(0.011)
Observations	9,244	9,244	9,244
R-squared	0.034	0.012	0.061
Number of clusters	457	457	457
Panel C: Secondary education			
Treatment effect (multiple period DD)	2.610**	0.027	0.019
	(1.261)	(0.029)	(0.023)
Observations	6,430	6,430	6,430
R-squared	0.023	0.012	0.103
Number of clusters	321	321	321
Year FE	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: The results from estimating equation (1) using a number of subgroups of the national sample (2005–14). The results are estimated separately for entrepreneurs and businesses in the service sector (panel A), those in the sales sector (panel B), and entrepreneurs with at least a secondary education (panel C). All regressions control for municipality and year fixed effects. Individual and household controls include age, household income, household size, total number of children, and dummies for education level (primary school, secondary school, or post-secondary education) and marital status. Estimations are conditional on being self-employed. Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Local effects of Estancias Infantiles on firm performance, by firm sector and education level, urban sample (ENAMIN)

Variables	(1) Hours worked	(2) Operated from home	(3) Had paid worker	(4) Applied for credit	(5) ln(physical capital)
Panel A: services					
Treatment effect	6.276*	-0.011	-0.009	-0.099	0.548
(multiple period DD)	(3.226)	(0.089)	(0.058)	(0.066)	(0.702)
Observations	731	731	731	731	731
R-squared	0.046	0.064	0.053	0.062	0.243
Number of clusters	54	54	54	54	54
Panel B: sales					
Treatment effect	2.256	0.006	-0.001	-0.052	-0.768
(multiple period DD)	(3.930)	(0.071)	(0.025)	(0.061)	(0.617)
Observations	1,174	1,174	1,174	1,174	1,174
R-squared	0.078	0.021	0.011	0.033	0.197
Number of clusters	64	64	64	64	64
Panel C: Secondary education					
Treatment effect	4.634	-0.075	0.037	0.013	-0.075
(multiple period DD)	(3.178)	(0.086)	(0.048)	(0.065)	(0.608)
Observations	1,059	1,059	1,059	1,059	1,059
R-squared	0.051	0.021	0.010	0.044	0.108
Number of clusters	65	65	65	65	65
Year FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: The results from estimating equation (1) using a number of subgroups of the urban sample (2002, 2008, 2010, 2012). The results are estimated separately for entrepreneurs and businesses in the service sector (panel A), those in the sales sector (panel B), and entrepreneurs with at least a secondary education (panel C). All regressions control for municipality and year fixed effects. Individual and household controls include age, household income, household size, total number of children, and dummies for education level (primary school, secondary school, or post-secondary education) and marital status. Estimations are conditional on being self-employed. Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Robustness: baseline DD estimations using different control groups, national sample (ENOE)

Variables	(1) Baseline (mult.period DD)	(2) Control group children 5–6 yrs.	(3) Drop those with siblings 4–6 yrs.	(4) Youngest child 3–4 yrs.	(5) All municipalities
Working hours					
Treatment effect	-0.432	0.235	-0.793	-1.627	-0.747
	(0.894)	(1.118)	(0.923)	(1.270)	(0.874)
Observations	16,437	12,963	14,218	6,786	17,803
R-squared	0.024	0.024	0.025	0.026	724
Number of clusters	546	546	546	518	0.024
Operated from home					
Treatment effect	0.001	-0.020	0.019	0.040	0.009
	(0.016)	(0.019)	(0.017)	(0.026)	(0.016)
Observations	16,437	12,963	14,218	6,786	17,803
R-squared	0.022	0.023	0.024	0.024	0.022
Number of clusters	546	546	546	518	724
Had paid worker					
Treatment effect	0.005	-0.000	0.005	0.002	0.005
	(0.010)	(0.013)	(0.011)	(0.017)	(0.009)
Observations	16,437	12,963	14,218	6,786	17,803
R-squared	0.097	0.095	0.097	0.092	0.096
Number of clusters	546	546	546	518	724
Year FE	yes	yes	yes	yes	yes
Municipality FE	yes	yes	yes	yes	yes
Control	yes	yes	yes	yes	yes

Notes: The results from estimating equation (1) using alternative definitions of the control group in the national sample (2005–14). All regressions control for municipality and year fixed effects. Individual and household controls include age, household income, household size, total number of children, and dummies for education level (primary school, secondary school, or post-secondary education) and marital status. Column 1 shows the preferred estimation in table 2 (panel B), column 2 restricts the control group to include only entrepreneurs whose youngest child is 5–6 years old, and column 3 drops entrepreneurs in the treatment group who also has a child 4–6 years old. The results in column 4 are estimated using a restricted sample of entrepreneurs whose youngest child is 3–4 years old. Column 5 includes all municipalities in the sample, including those who never introduced Estancias Infantiles or introduced it later than 2007, resulting in DD estimation with a staggered roll-out. All estimations are conditional on self-employment. Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Robustness: baseline DD estimations using different control groups, urban sample (ENAMIN)

Variables	(1) Baseline (mult. period DD)	(2) Control group children 5–6 yrs.	(3) Drop those with siblings 4–6 yrs.	(4) Youngest child 3–4 yrs.
Working hours				
Treatment effect	4.244*	3.878	4.780**	2.464
	(2.197)	(3.189)	(2.356)	(3.072)
Observations Observations	2,334	1,829	2,061	995
R-squared	0.034	0.035	0.035	0.041
Number of clusters	84	84	84	80
Operated from home				
Treatment effect	-0.051	-0.052	-0.067	-0.198***
	(0.050)	(0.050)	(0.055)	(0.070)
Observations	2,334	1,829	2,061	995
R-squared	0.015	0.014	0.014	0.051
Number of clusters	84	84	84	80
Had paid worker				
Treatment effect	0.017	0.005	0.030	0.009
	(0.027)	(0.032)	(0.030)	(0.049)
Observations	2,334	1,829	2,061	995
R-squared	0.017	0.023	0.019	0.021
Number of clusters	84	84	84	80
Applied for credit				
Treatment effect	-0.057*	-0.078**	-0.063*	0.013
	(0.033)	(0.038)	(0.037)	(0.058)
Observations	2,334	1,829	2,061	995
R-squared	0.043	0.043	0.043	0.051
Number of clusters	84	84	84	80
ln(physical capital				
Treatment effect	0.023	0.036	0.036	-0.326
	(0.422)	(0.596)	(0.476)	(0.650)
Observations	2,334	1,829	2,061	995
R-squared	0.124	0.130	0.119	0.121
Number of clusters	84	84	84	80
Year FE	yes	yes	yes	yes
Municipality FE	yes	yes	yes	yes
Control	yes	yes	yes	yes

Notes: The results from estimating equation (1) using alternative definitions of the control group in the urban sample (2002, 2008, 2010, 2012). All regressions control for municipality and year fixed effects. Individual and household controls include age, household size, total number of children, and dummies for education level (primary school, secondary school, or post-secondary education) and marital status. Column 1 shows the preferred estimation presented in table 2 (panel B), column 2 restricts the control group to include only entrepreneurs whose youngest child is 5–6 years old, and column 3 drops entrepreneurs in the treatment group who also has a child between 4–6 years old. The results in column 4 are estimated using a restricted sample of entrepreneurs whose youngest child is 3–4 years old. All estimations are conditional on self-employment. Standard errors are clustered at the municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Difference-in-difference (DD) estimations with individual fixed effects, national sample (ENOE)

Variables	(1) Hours worked	(2) Operated from home	(3) Had paid worker
Treatment effect	-0.343	-0.028	-0.002
	(1.101)	(0.027)	(0.020)
Observations	8064	8064	8064
R-squared	0.007	0.002	0.002
Number of id	4608	4686	4686
Year FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: The results from estimating Equation (1) using individual panel data that includes 5 quarters between the 4th quarter 2006 and the 4th quarter 2007. All regressions control for municipality and year fixed effects. Individual and household controls include age, household income, household size, total number of children, and dummies for education level (primary school, secondary school, or post-secondary education) and marital status. Furthermore, the equation controls for individual fixed effects and quarterly fixed effects. Standard errors are clustered at the individual level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figures

Figure A1: Total numbers of daycare centers enrolled in Estancias Infantiles, 2007-–15

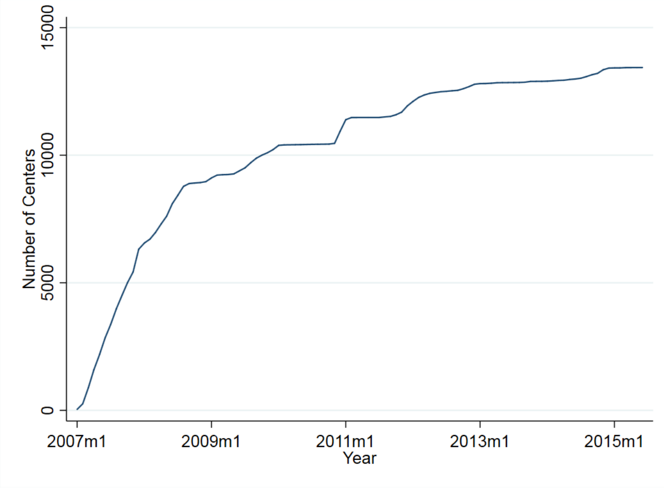
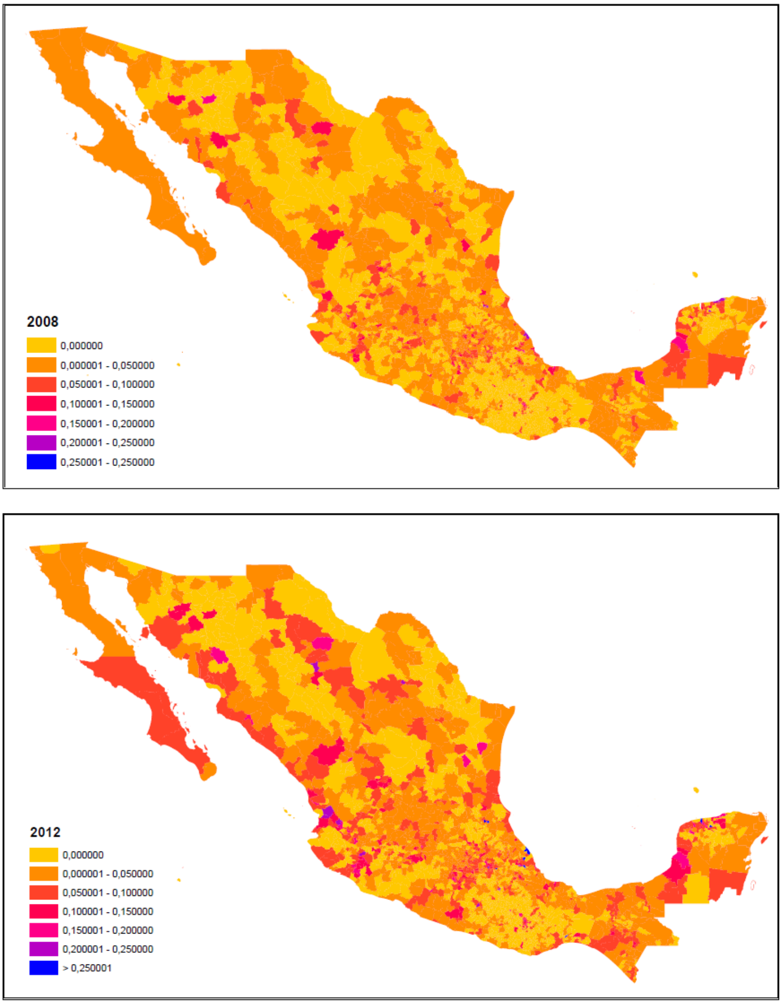


Figure A2: Geographic variation in availability of Estancias Infantiles in Mexico at two points in time, 2008 and 2012



Note: The program was first implemented in 2007.

Table A4: Correlations between Estancias Infantiles and entrepreneur characteristics, —DD framework

VARIABLES	(1) Age	(2) > secondary school	(3) Primary school	(4) Secondary school	(5) Married	(6) Widow	(7) Divorced	(8) Number adults	(9) Number children
Panel A: ENOE									
Child1–3*post	0.026	-0.028	0.016	0.005	-0.009	-0.006	0.004	-0.025	0.019
	(0.241)	(0.018)	(0.015)	(0.016)	(0.012)	(0.004)	(0.009)	(0.035)	(0.039)
Observations	16,437	16,437	16,437	16,437	16,437	16,437	16,437	16,437	16,437
R-squared	0.051	0.003	0.007	0.002	0.003	0.002	0.002	0.005	0.007
Number of clusters	546	546	546	546	546	546	546	546	546
Mean baseline	32.592	0.390	0.264	0.314	0.905	0.010	0.050	2.265	2.326
Panel B: ENAMIN									
Child1–3*post	-0.727	-0.019	0.094**	-0.074	-0.029	-0.022	0.047	0.132	-0.251**
	(0.640)	(0.057)	(0.038)	(0.046)	(0.044)	(0.015)	(0.036)	(0.083)	(0.099)
Observations	2,334	2,334	2,334	2,334	2,334	2,334	2,334	2,334	2,334
R-squared	0.051	0.007	0.017	0.016	0.003	0.006	0.006	0.010	0.025
Number of clusters	84	84	84	84	84	84	84	84	84
Mean baseline	33.309	0.421	0.258	0.204	0.838	0.026	0.091	2.173	2.359
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Panel A presents the estimated correlation between Estancias Infantiles and individual and household characteristics, comparing eligible and ineligible entrepreneurs before and after the implementation of the program in a DD estimation strategy. Panel B identifies the corresponding correlations by comparing eligible entrepreneurs with ineligible entrepreneurs, in municipalities with different program intensities, before and after program implementation, usin a DDD framework. All results are estimated using the national sample. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Summary statistics for microenterprises, national sample (ENOE)

Variable	(1) Mean	(2) Std. dev.	(3) Min	(4) Max
Working hours per week	30.514	21.738	1	112
Employees				
Workers	0.242	0.428	0	1
Paid workers	0.087	0.282	0	1
Location				
Specific premises	0.299	0.458	0	1
Work from home	0.313	0.464	0	1
Walking vendor	0.127	0.333	0	1
Improvised stand	0.053	0.224	0	1
Stand	0.038	0.192	0	1
Client's home	0.141	0.348	0	1
Vehicle	0.012	0.109	0	1
Sell directly to client	0.974	0.156	0	1
Type of business				
Service	0.196	0.397	0	1
Sales	0.573	0.397	0	1
Manufacturing	0.186	0.389	0	1
Observations	16 437			

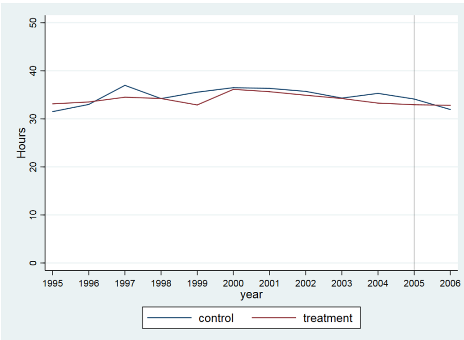
Notes: The results present sample means (column 1), standard deviation (column 2) and sample minimum (column 3) and maximum (column 4).

Table A6: Summary statistics microenterprises, urban sample

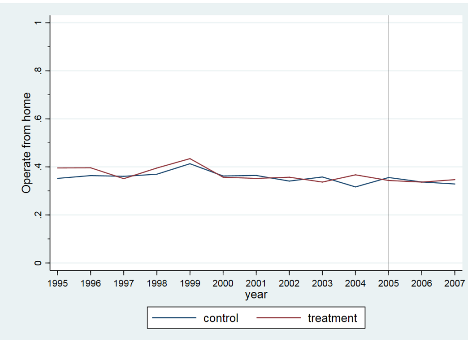
Variable	(1) Mean	(2) Std. Dev.	(3) Min	(4) Max
Working hours per week	31.516	20.644	1	112
Employees				
Workers	0.189	0.391	0	1
Paid workers	0.074	0.262	0	1
Location				
Specific premises	0.275	0.446	0	1
Work from home	0.323	0.467	0	1
Walking vendor	0.105	0.307	0	1
Client's home	0.191	0.393	0	1
Vehicle	0.010	0.101	0	1
Type of business				
Service	0.335	0.472		
Sales	0.514	0.499	0	1
Manufacturing	0.109	0.312	0	1
Individual business	0.910	0.286	0	1
Family business	0.047	0.212	0	1
Business started by owner	0.806	0.395	0	1
Reason for starting business				
Flexible working hours	0.060	0.239	0	1
Good opportunity	0.054	0.226	0	1
Supplement family income	0.519	0.499	0	1
Higher income	0.099	0.299	0	1
Credit				
Applied for credit	0.172	0.377	0	1
Amount of credit	2 120.12	6 510.60	0	100,000
Capital				
Start-up capital	0.862	0.344	0	1
Physical capital	9 684.42	1 9679.02	0	103 300
Observations	2 334			

Note: The results present sample means (column 1), standard deviation (column 2) and sample minimum (column 3) and maximum (column 4).

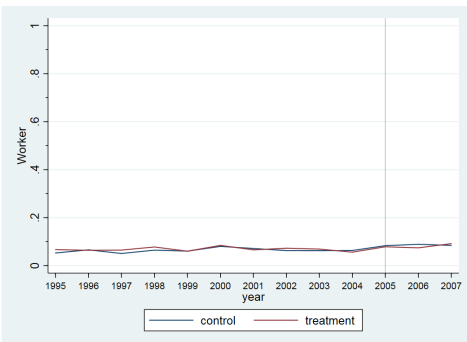
Figure A3: Pre-trends for outcome variables, national sample



Working hours



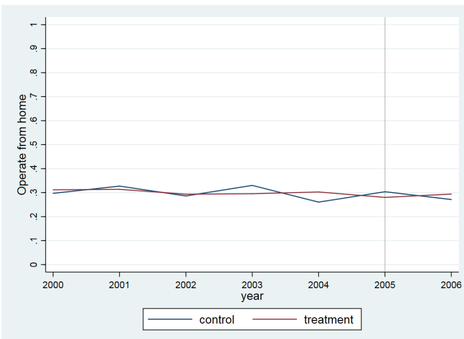
Working from home



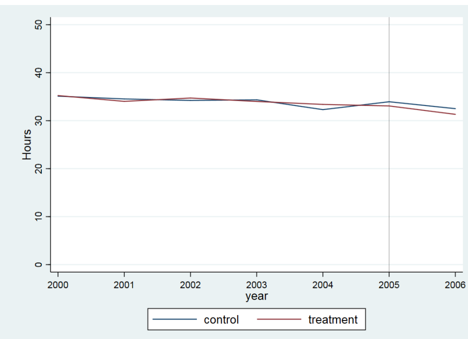
Worker

Notes: The graphs show parallel trends in outcome variables between 1995 and 2007 for the national sample. The treatment group refers to female entrepreneurs whose youngest child was 1–3 years old and the control group include female entrepreneurs whose youngest child was 4–6 years old. The vertical line indicates the restructure of the National Survey of Occupation and Employment (ENOE) in 2005.

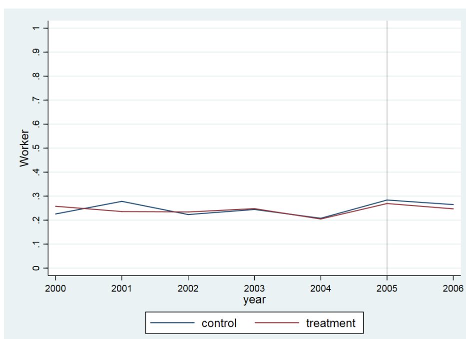
Figure A4: Pre-trends for outcome variables, urban sample



Working hours



Working from home



Worker

Notes: The graphs show parallel trends in outcome variables between 2000 and 2007 for the urban sample. The treatment group refers to female entrepreneurs whose youngest child was 1–3 years old, and the control group include female entrepreneurs whose youngest child was between 4–6 years old. The vertical line indicates the restructure of the National Urban Employment Survey (ENEU) in 2005.

III

Variation in the Quality of Primary Healthcare

Evidence from Rural and Urban Health Services in Rwanda

Carolin Sjöholm*

Abstract

Disparities in access to quality healthcare within countries represent a potential impediment to reaching the sustainable development goal of better health and well-being for all. In this paper, I first identify a disparity in the quality of primary healthcare between rural and urban primary health facilities in Rwanda. Then I study the importance of differences in structural inputs and contextual factors in explaining this outcome. To measure quality, I construct a quality score that summarizes both structure and process quality indicators. I use administrative data from the performance-based financing scheme to calculate the quality scores. These data were collected during unannounced evaluations of public health centers, performed by teams of professional hospital staff. The results confirm a small but significant quality gap between rural and urban health centers. Rural centers obtain a 1% lower quality score at the mean, or 0.3 standard deviations, compared with urban centers, potentially masking important differences in the delivery of health services for patients. I find that differences in structural and contextual inputs, such as access to drugs and clinic beds, wage expenditure, and distance to the nearest clinic and district hospital, explain only a small share of the difference in quality between rural and urban areas. The results indicate that investment in such factors might not represent an efficient policy tool to eliminate within-country inequalities in access to quality healthcare.

Keywords: healthcare quality, structural inputs, process quality, Rwanda
JEL classification: I11, I14, I18

*University of Gothenburg, carolin.sjoholm@economics.gu.se. I would like to thank my advisors Måns Söderbom and Annika Lindskog for their invaluable input. I also thank Andinet Woldemichael and Gustav Kjellsson for comments and support. I am also grateful to Pascal Kayobotsi, Andrew Muhire, and Emmanuel Ntawuyirusha at the Ministry of Health in Kigali for sharing their knowledge and expertise regarding the Rwandan health sector, as well as the HMIS and PBF data.

1 Introduction

United Nations Sustainable Development Goal 3, Good Health and Well-being, positions equity as a central issue of the global health agenda by aspiring to ensure healthy lives and promote well-being for all people at all ages. Disparities in access to quality healthcare within many low-income countries represent one of the major impediments to reaching this goal. One source of disparity that many studies have documented is that between rural and urban areas (Kruk et al., 2017; Gage et al., 2017; Das et al., 2012; Scheil-Adlung, 2015). Rural areas often face substantial shortages in health infrastructure compared with urban areas, such as a lack of equipment, health workforce, and physical infrastructure (Leslie, Sun, & Kruk, 2017; Scheil-Adlung, 2015). However, while it has been widely accepted that structural inputs represent a necessary condition to deliver quality care (World Health Organization, 2018), less is known about their importance in determining the quality of the actual health services that patients receive when they visit a health facility.[1]

This study investigates the gap in healthcare quality between rural and urban public health centers in Rwanda and how much of this gap can be attributed to differences in structural inputs. I use data from the performance-based financing (PBF) national monitoring system to assess the quality of healthcare. PBF is a health financing scheme that provides funding for health services based on their quality. The PBF quality data come from unannounced evaluations at public health centers across the country during 2013–18. District hospitals are responsible for the quality evaluations of all health centers in their district and the evaluations are carried out by a team of specialists from the hospital. The data were collected using direct observations during patient visits, chart examinations, and facility checklists (Ministry of Health, 2018b).

I create two measures of quality—a general quality score and a patient-focused quality score—based on 12 different PBF quality indicators. The general quality score measures the overall quality of the health clinic, whereas the patient-focused score includes only patient-focused activities. The quality scores summarize both structural and process measures of quality, describing the context in which care is delivered as well as all acts of healthcare delivery (Donabedian, 1988). The quality scores are matched

[1]Health facilities in many developing countries often face substantial gaps in their readiness to provide basic healthcare services (Leslie, Spiegelman, et al., 2017; K. L. Leonard & Masatu, 2007). The World Health Organization (WHO) estimates that approximately 40% of health facilities in low- and middle-income countries lack access to essential infrastructure such as improved water and nearly 20% lack sanitation (World Health Organization & UNICEF, 2015).

with data on structural inputs and contextual factors from the Rwandan Integrated Health Management Information System (HMIS). Structural inputs include measures of clinic size, access to drugs, and access to health workers, proxied by the clinic's total number of beds, total number of yearly outpatient visits, average drug expenditure per patient visit, and total wage costs. Additionally, I include a number of contextual factors that estimate the importance of external market factors such as demand and competition. These variables measure total population in clinic catchment area and the distances to the closest neighboring clinic and the nearest district hospital. To test how much of the inequalities in healthcare quality can be attributed to differences in structural and contextual factors, I measure whether the quality gap between rural and urban clinics decreases once I control for these factors in the estimations.

The results show that there is a small but significant difference between rural and urban clinics in the quality of care provided. Rural clinics received approximately 1%, or 0.3 standard deviations, lower quality scores than urban clinics. The results confirm previous empirical evidence from India (Das et al., 2008), Haiti (Gage et al., 2017), and Indonesia (Barber et al., 2007), indicating that the quality of care in rural areas is considerably lower than in urban areas. Furthermore, the results suggest that the structural inputs and contextual factors explain only a small share of the existing differences in quality. For example, differences in structural variables such as the total number of beds, staff, and medicines represented approximately 9% of the difference in quality scores. The results are in line with previous research that indicates that the structural inputs are only weakly associated with the process of care (Leslie, Sun, & Kruk, 2017).

This paper relates to the literature that studies variation in the quality of healthcare within low-income countries. The literature has documented significant disparities in the quality of healthcare by a number of stratifiers such as the type of healthcare providers (Barber et al., 2007; Das et al., 2012), household wealth level (Barber et al., 2007; Sharma et al., 2017), and rural and urban location (Barber et al., 2007; Das et al., 2012; Gage et al., 2017). Empirical evidence suggests that medical expertise is higher among urban than rural healthcare providers: in one study in India 52% of the healthcare providers in urban areas reported having a medical degree, compared with only 11% in rural areas (Das et al., 2012). Furthermore, other studies suggest that rural care providers are less likely than urban health workers to give a correct diagnosis and treatment (Das et al., 2012), and to adhere to clinical guidelines (Leslie, Spiegelman, et al., 2017).

I add to this literature by providing evidence of disparities in health quality between rural and urban areas in Rwanda. Given the country's remarkable improvements in public health outcomes in recent decades, Rwanda represents a compelling case study of the quality of care in a developing country context. Over two decades ago, genocide left almost 1 million dead and a legacy of poverty and human devastation. Rwanda's under-five mortality rate was the highest in the world, and life expectancy at birth was the lowest (Binagwaho et al., 2014). Since then, Rwanda has achieved impressive health gains, surpassing the performance of neighboring countries.[2] However, despite remarkable improvements in health outcomes, health inequalities within the country remain and continue to represent a major challenge (Liu et al., 2019; Pose & Samuels, 2011).

Moreover, this paper contributes to the literature that investigates which factors explain disparities in the quality of care within low-income countries. Previous research that has studied the association between health infrastructure and the quality of process care is inconclusive: while some studies find a positive correlation between medical infrastructure and clinical quality (Kruk et al., 2017), many others suggest that structural inputs do not predict the quality of care provided in consultations (Leslie, Sun, & Kruk, 2017; Das et al., 2008, 2012; Das & Hammer, 2014; K. L. Leonard & Masatu, 2007). Instead, these studies find that provider effort represents the primary determinant of the quality of care received by patients. Empirical evidence suggests that low practitioner effort is frequently found in health markets in many low-income countries. For example, a study in India showed that public care providers spent on average 2.4 minutes with the patient and completed 16% of checklist items (Das et al., 2016). Another study of 7 Sub-Saharan countries showed that healthcare providers on average performed 62% of the recommended antenatal care actions and approximately half of the suggested actions related to sick-child care (Kruk et al., 2017). In this context, low provider effort is likely to constrain potentially important impacts of structural inputs on the quality of care provided to patients.

[2] In Rwanda, between 1996 and 2018, life expectancy almost doubled, from 35 to 69 years, and the under-five mortality rate dropped from 196 to 35 per 1000 births, and between 2000 and 2018, the maternal mortality rate dropped from 1160 to 248 per 100,000 live births. In comparison, between 1996 and 2018, life expectancy increased from 49 to 65 years in Tanzania and from 45 to 61 years in Burundi, and under-five mortality dropped from 154 to 53 in Tanzania and from 175 to 59 in Burundi. Between 2000 and 2018, the maternal mortality rate dropped from 854 to 524 in Tanzania and from 1010 to 548 in Burundi (World Bank, 2020).

I contribute to previous research by providing evidence on the importance of structural inputs in explaining healthcare quality from a health sector with strong monitoring and accountability mechanisms. The PBF scheme represents a public mechanism for accountability by introducing a financial payment structure that conditions funding on health service quality and provider effort (Ministry of Health, 2018b). In addition to the PBF scheme, the central government monitors service delivery through a unique traditional system of performance contracts between local government agencies and the president of Rwanda, called Imihigo. The contracts include health-related performance targets such as reduction in morbidity and mortality, and access to care (Versailles, 2012).

The importance of provider effort as a determinant for the quality of care exceeds the scope of this paper. The quality measures used in this analysis were collected through direct observations. In the presence of an observing enumerator, caregivers are likely to alter their behavior in order to comply with clinical guidelines (K. Leonard & Masatu, 2006). Therefore, lack of provider effort is not likely to be fully captured by these measures. However, the Rwandan health sector provides a setting for estimating the importance of structural factors in a market where such effects are less likely to be constrained by low levels of provider effort.

The remainder of this paper proceeds as follows. Section 2 describes the study setting and the PBF scheme. Section 3 provides the data and defines the quality measures used in the analysis. Section 4 presents the empirical analysis, and Section 5 concludes.

2 The Rwandan Health Sector

Rwanda has a decentralized health system, in which the administrative responsibility for health service delivery, facility management, and infrastructure investment is centered in 30 district health departments. The districts are responsible for the health facilities and services provided therein, making them the organizational unit of primary health services provided at health centers and district hospitals (Versailles, 2012).

This study focuses on the quality of primary health services. Health centers are the gatekeepers of the health system and the focal point for primary care. Health centers provide basic primary care including promotional activities, and preventive and curative health services, such as normal deliveries, minor surgical interventions, management of noncommunicable and communicable diseases, and laboratory testing (Kalisa et al., 2015). For needs beyond these services, patients are referred to district hospitals, provincial hospitals, or referral facilities as needed (African Strategies for Health, 2015). In 2016 Rwanda counted with 499 health centers throughout the country, each serving a catchment area of several thousand people. Furthermore, the health system consisted of 8 national referral hospitals, 4 provincial hospitals and 36 district hospitals. Approximately 10% of all healthcare providers were private for-profit. Private healthcare providers are located mainly in Kigali, while the rest of the country is underserved by the private sector (African Strategies for Health, 2015).

One of the main objectives of the Rwandan health sector strategic plan is to ensure universal access to the highest attainable quality of health services at all levels. To expand the availability of healthcare to the Rwandan population, the government has implemented a number of health policies that aim to increase both the demand and the supply of healthcare. In 2006, the Rwandan government introduced a national community-based health insurance (CBHI) scheme that offered financial protection against healthcare expenditure for those who lacked access to health insurance. The insurance covered all health services provided at public and private nonprofit health facilities. The CBHI scheme contributed to increased geographic and financial accessibility, as well as usage of health services, among Rwandans (Lu et al., 2012). As a result, Rwanda has made impressive steps towards universal health coverage (World Health Organization, 2017).

The introduction of the CBHI scheme was accompanied by an increase in total healthcare expenditure in Rwanda, which grew from approximately 4% of GDP at the beginning of 2000 to 8.5% by 2012. On average, between 2000 and 2017, healthcare expenditure in Rwanda represented approximately 6.5% of GDP, higher than in several Sub-Saharan countries, such as Ghana (3.8%), and Ethiopia (4.4%), but less than in others, such as Uganda (8.8%). Healthcare expenditure represented approximately 5% of GDP during the period among all countries in Sub-Saharan Africa (World Bank, 2020).

In addition to the policy efforts that aimed to increase the demand for healthcare, the government has put an emphasis on increasing the supply of healthcare services across the country. To ensure service delivery, the central government introduced a number of public accountability systems (Pose & Samuels, 2011). The PBF scheme represents one of the key accountability systems in the health sector. The program was introduced as a national health financing policy at all health facilities in Rwanda in 2006 (Ministry of Health, 2018b). The performance-based contracts attach payments for specific health services, conditional on quality standards. The quality standards are assessed using a number of indicators that are defined according to qualitative and quantitative quality measures, outlined in the national PBF evaluation tool (Ministry of Health, 2018b; Basinga et al., 2011). For example, the evaluation of deliveries considers the correct use of medicines and medical supplies such as anesthesia, sterile gloves, and emergency kits; presence of qualified personnel; measures to prevent infection; and accurate decisions regarding referral of patients to referral hospitals. The evaluation of vaccines puts a large emphasis on storage and correct management of the vaccine stock. Additionally, the quality index includes the direct observation of four vaccinations of children, evaluating indicators such as the right dosage, appropriate injection sites, correct use of garbage cans, systematic reporting of side effects, and correct registration of vaccination. The registration of tuberculosis, use of the technical manual, and the stock of antituberculosis medication are other quality indicators included in the tuberculosis section of the evaluation. Importantly, the care of any cases of diagnosed tuberculosis are evaluated according to the guidelines of the national tuberculosis division (Ministry of Health, 2018b). For a detailed description of all quality indicators, please consult Table A1 in Appendix.

Health center performance is regularly measured against quality indicators during quarterly quality evaluations conducted by the district hospitals. At the end of each evaluation session, the points awarded for every quality indicator are summed up and used as an overall quality score for the health facility. Each facility's total score is used to determine the final PBF reimbursement it will receive. Every health facility prepares an annual action plan integrating all expected income sources, including PBF revenues and expenses. Each individual caregiver signs a contract with the health center that states the performance requirements and the monetary contribution paid for performance, measured by the caregiver's adherence to the checklist of quality indicators.

In conjunction with the introduction of the PBF, the Rwandan government introduced another type of performance contract called Imihigo with the aim of improving the efficiency of service delivery. Imihigo is an accountability system based on performance contracts, detailing specific performance targets that the local governments set. The contracts are signed between local and national governments. Local ministries, public agencies and districts are required to sign formal public service agreements to deliver key specific outputs each year. Activities included in the contracts are derived partly from the national economic development and poverty reduction strategy, and partly from demands and priorities in local communities.

The district governments develop five-year and one-year plans and targets based on these national and local demands and conversations. Each target is linked to a specific performance indicator. At the end of the process, the local governments sign the Imigiho contracts with the president in a public ceremony. The documents are available to the general public (Byamukama & Makonnen, 2012). The Imihigo work process is monitored closely and evaluated yearly by evaluation teams put together by the national government, to hold local governments accountable for meeting their targets. Annual evaluation is conducted by the national quality assurance team to determine the extent to which districts have achieved their objectives and contributed to improvements in the socioeconomic well-being of citizens. District mayors are held to account for their *Imihigo* performance twice a year in public sessions in Kigali, which are chaired by the president. As a result, Imihigo put pressure on public officials to fulfill policy goals and to provide quality services (Versailles, 2012).

3 Data

The quality measures in this analysis are based on data from the PBF national monitoring system. These data contains information on 12 indicators that measure health center adherence to quality checklists, developed by the Ministry of Health. The data were collected quarterly through unannounced clinic visits by professional staff. Each administrative district in Rwanda has a district hospital that is responsible for evaluating health centers in its catchment area. The quality indicators cover a wide range of administrative and clinical activities, including antenatal care, family planning, and deliveries; diagnosis and treatment of tuberculosis; providing vaccinations; financial, data and general management; management of laboratory and pharmacies; and hygiene. To ensure the accuracy of the quality evaluations, the Ministry of Health conducts biannual counter verification of health facility data. A district steering committee may sanction or even exclude a health center from the PBF strategy for reasons such as collusion with or corruption of the evaluators (Ministry of Health, 2018b).

I categorize rural health centers based on their geographic distance from one of 10 major cities in Rwanda: Kigali, Huye, Muhanga, Musanze, Rubavu, Byumba, Risuzu, Kibuye, Rwamagana, and Kibungo. I use the geographic coordinates of each health facility to measure the distance between the health center and the closest city center. The baseline definition of a rural health center uses a buffer zone of 15 kilometers (km) around each of the 10 cities. All health clinics that are situated outside the buffer zones are defined as rural. Kigali is the largest city in the country and accounts for more than 50% of the total urban population. As of 2012, Kigali had approximately 845 000 inhabitants. The second-largest city, Rubavu, had around 150 000 inhabitants, while 4 of the 10 largest cities had less than 60 000 inhabitants. Consequently, the urban system in the country is dominated by Kigali (Rajashekar et al., 2019). I study the sensitivity of the results to the choice of rural-urban cutoff by using alternate radii of the urban zones to 10 km, 20 km and 25 km. The results are presented in the next Section.

I use data from the HMIS on facility characteristics and structural inputs. The HMIS is a national data base containing monthly information for all health facilities in the country on a number of in- and outpatient statistics, ante- and postnatal care, and human resources, as well as financial information. The data are recorded electronically at each individual health facility and coordinated by the Ministry of Health. By matching data from the PBF national monitoring system and the HMIS, I was able to match the quality score for each health center with information on medical infrastructure and patient visits.

I construct a yearly quality score for each health center by taking the mean of all quarterly indicators during each year. The data contains information on 499 health centers, representing the complete population of clinics in 2018. I dropped 29 health centers that did not have any information on quality scores in the PBF data, as well as 74 clinics that lacked geographic coordinates, which prevented me from categorizing them as rural or urban clinics, and information on structural inputs in the HMIS data. The first year of data (2012) was of overall low quality, likely because electronic reporting in the HMIS was introduced during that year. Therefore, I dropped the data from 2012. The resulting sample consists of 396 health centers, covering the period 2013–18, with a total of 2113 observations.

Appendix table A2 provides a brief analysis of the missing values. Overall, the results suggest that rural clinics are approximately 7 percentage points more likely than urban to have missing values. Furthermore, the results suggest that the missing values are not random; compared with other clinics, clinics missing values are likely to have a larger number of hospital beds, higher drug expenditure per patient visit, and a different case mix. To evaluate whether the pattern of missing values is likely to affect the results, I further analyze the correlation between clinic characteristics and the likelihood of having a missing value separately for rural and urban health centers. The results show a positive correlation between missing values and the number of clinic beds among urban health centers, suggesting that larger urban clinics are more likely to be missing values in the sample. Furthermore, there is a significant positive correlation between the likelihood of missing values and total wage expenditure among urban clinics, whereas this correlation is significantly negative among rural clinics. Taken together, the results indicate that the missing values are likely to contribute to an underestimation of disparities in health infrastructure between rural and urban clinics. Under the assumption that increased access to health infrastructure is associated with higher health services quality, this would result in an underestimation of the quality gap between rural and urban health facilities.

3.1 Quality measures

Measures of quality of care have traditionally been divided into three domains: structure, process and outcome. Structural quality refers to the material and human resources, as well as the physical and organizational factors of the healthcare provider, and the process component describes the quality of health service delivery to patients such as adherence to clinical guidelines. Outcome reflects the impact of care on population health status (Donabedian, 1988). For this analysis, I create two measures of quality—a general quality score and a patient-focused score—based on 12 different PBF quality indicators (listed in table 1). The general quality score measures the overall quality of the health clinic, including activities such as data and financial management and hygiene and laboratory management, as well as the quality of the care provided, focusing on antenatal care, family planning, deliveries, tuberculosis, vaccination and supervision of community health workers. The patient-focused quality score is restricted to include only quality indicators related to patient-focused activities, such as antenatal care, family planing, deliveries, tuberculosis, and vaccinations. The aim of this measure is to provide a quality index that exclusively describes the direct care provided to patients. The quality scores summarize 140 structure and process measures of quality (see appendix table A1 for more detailed description of the quality indicators).

Structural measures evaluate the context in which the health services are provided, including facilities, personnel, and management related to the delivery of care, whereas the process component assesses the accuracy of the clinical encounter—that is, whether health services provided to patients are consistent with national guidelines (Donabedian, 1988). The process-related indicators measure adherence of health workers to the clinical guidelines for best practice of care defined by the Rwandan government in the clinical practice guidelines. The quality of health is measured by the PBF checklists, defined by the Rwandan government and applied during the quality evaluations at the health centers. Approximately 80% of the patient-related quality measure consists of process-related indicators, corresponding to about 60% of the general quality score.

Table 1 summarizes the 12 quality indicators used to construct the general and patient-focused quality scores, as well as the aggregated total quality scores. Column 1 describes the average values for all health centers, and columns 2 and 3 describe the values for rural and urban clinics. Columns 4 and 5 present sample maximum and minimum scores, and column 6 indicates the maximum quality score a health clinic could potentially receive according to the PBF financing rules (Ministry of Health, 2018b).

The results indicate that the average scores exceed 80% of the total score (column 6) for nearly all indicators. General organization, laboratory management, and the oversight of community health workers reached approximately 75% of the total quality score on average. The two rows under the individual quality indicators present the aggregated total quality scores. The general quality measure was 404.05 on average during the study period, and the patient-focused total score was 191.30. This corresponds to 88% on the general quality score and 80% on the patient-focused quality score, on average. The Ministry of Health categorizes health centers with a quality score above 80% as high-quality health facilities, whereas those that receive a quality score of less than 60% are described as low-quality providers. Health centers in this study received a general quality score ranging from 73% to 95% (333.75 to 437.5 score points) and a patient-focused score ranging from 57% to 85% (137 to 205 points). Less than 5% received a quality score lower than 5%, both general and patient focused, that is lower than 80%, either general of patient-focused.

Importantly, even small deviations from the maximum score could potentially mask quality deficits that could have large implications on the quality and safety of the care provided at the health clinic. For example, according to administrative documentation from the PBF, the gap between the sample average and maximum score in relation to the hygiene indicator could imply that a clinic lacks access to a toilet or latrine with soap and running water; that it has no contract with a cleaning service and no formal record of cleaning products such as soap, bleach, and chlorine; or that it lacks access to a water source such as running water, or a well or tank. Similarly, a difference between average and maximum quality scores in relation to deliveries could imply that a health center lacks routines for infection prevention, that it lacks local anesthesia and saline, or that deliveries were not done by qualified personnel.

The results in columns 2 and 3 of table 1 indicate that urban clinics on average had significantly higher quality scores for a number of quality indicators. Urban clinics received significantly higher quality scores on all indicators related to clinical care, with the exception of vaccinations. However, the differences are small and do not exceed one score point.

Table 1: Descriptive statistics: quality indicators and quality scores

	Indicators and scores	(1) All	(2) Rural	(3) Urban	(4) Min (sample)	(5) Max (sample)	(6) Max PBF score
General score	Data management	35.713	35.658	35.809	18.5	40	40
		(3.480)	(3.511)	(3.427)	(46%)	(100%)	
	Financial management	23.041	22.993	23.124	2	25	25
		(2.024)	(2.094)	(2.005)	(8%)	(100%)	
	General organization	28.367	28.209	28.640**	11.5	37	37
		(4.282)	(4.345)	(4.161)	(46%)	(100%)	
	Hygiene	25.092	24.973	25.298***	14.25	30	30
		(2.328)	(2.350)	(2.277)	(47%)	(100%)	
	Laboratory	17.933	17.866	18.050	11.5	25	25
		(4.106)	(4.079)	(4.151)	(46%)	(100%)	
	Pharmacy	30.739	30.623	30.938***	19.25	33	33
		(1.941)	(1.951)	(1.908)	(58%)	(100%)	
	Community Health Workers	23.058	22.893	23.344**	10.25	30	30
		(5.037)	(5.021)	(5.054)	(34%)	(100%)	
Patient score	Antenatal Care (ANC)	60.579	60.358	60.960***	42.75	63	63
		(2.437)	(2.561)	(2.155)	(68%)	(100%)	
	Deliveries	60.612	60.494	60.817*	38.75	65	65
		(4.005)	(3.939)	(4.110)	(60%)	(100%)	
	Family Planing	52.690	52.323	53.323***	6.25	57	57
		(5.103)	(5.034)	(5.163)	(10%)	(100%)	
	Tuberculosis	17.418	17.241	17.722***	6	20	20
		(2.095)	(2.282)	(1.681)	(30%)	(100%)	
	Vaccination	28.801	28.689	28.996	14.25	35	35
		(4.914)	(4.933)	(4.880)	(40%)	(100%)	
Total	General score	404.051	402.328	407.026	333.75	437.5	460
		(17.051)	(17.533)	(15.759)	(73%)	(95%)	
	Patient-focused score	191.300	190.418	192.823	137	205	240
		(8.869)	(9.017)	(8.398)	(57%)	(85%)	
Observations		2113	1338	775			

Notes: Summary statistics of two quality scores: the general score, which includes all quality indicators, and the patient-focused score, which is restricted to include only quality indicators related to antenatal care, deliveries, family planning, tuberculosis and vaccinations. Column 1 provides average values for all health centers in the sample, whereas columns 2 and 3 present values for rural and urban clinics separately. Columns 4 and 5 show sample maximum and minimum, and column 6 gives the maximum quality score according to Ministry of Health guidelines. *** p<0.01, ** p<0.05, * p<0.1.

3.2 Explanatory variables

Table 2 shows summary statistics for a number of structural inputs and contextual factors. The total number of beds measure health center bed capacity and is a proxy for clinic size, total number of outpatient visits in a year measures the capacity of the health clinics to produce output, demand for services, and also clinic size; and total wage expenditure is used as a measure of access to health staff. Wage levels for health workers in public health facilities in Rwanda are centralized and defined by the Ministry of Health (Ministry of Health, 2019), implying that variation in wage expenditure among health centers measure differences in number of health workers or variation in the composition of workers among centers rather than disparities in salary across the country. Drug expenditure per outpatient visit is a proxy for access to drugs and medical supplies or the structure of drug prescription at each clinic. The population of the catchment area represent an estimation of the potential demand for healthcare faced by each clinic.

Column 1 describes all health centers in the sample, whereas columns 2 and 3 give values for rural and urban clinics separately. On average, health centers had 22 beds, performed approximately 25 700 outpatient visits, spent on average RwF 554 (USD 0.6) per outpatient visit on drugs, and had a total wage expenditure of approximately RwF 3.94e+07 per year (USD 40,300). Moreover, the results suggest that there is a statistically significant difference between rural and urban health centers in relation to both health infrastructural and contextual variables. Columns 2 and 3 suggest that urban health centers had significantly higher wage expenditure and fewer beds. Furthermore, urban health centers have significantly more populated catchment areas, and the distances to the closes neighboring clinic and district hospital were significantly shorter in urban areas. There is no significant difference between the groups in total outpatient visits per year or the average drug expenditure per outpatient visit.

In addition to the structural inputs and contextual factors, table 2 includes health clinic case mix. The case mix refers to the composition of patient diagnoses related to a number of priority health problems that were determined at a clinic during one year and describes the demand for healthcare at each health center and the complexity of the service provided. Acute respiratory infection and pneumonia represented the most common group of diagnoses at health centers, accounting for approximately 21% of all outpatient visits. Maternal health services such as ante- and postnatal care represented 3.3% and 3.1% of all visits, respectively, and deliveries accounted for 2% of the visits during one year. The results suggest a relatively similar case mix between urban and rural regions. However, respiratory diseases, malaria, and diarrhea and parasites were significantly less common in urban areas, whereas deliveries were more common in rural areas. There are no significant differences between ante- and postnatal care.

Table 2: Descriptive statistics: urban and rural health centers

Variables	(1) All	(2) Rural	(3) Urban	(4) Min	(5) Max
Medical infrastructure					
Beds	21.989	22.457	21.180***	0.41	124
	(0.352)	(11.166)	(9.817)		
Total outpatient visits	25695.75	25369.72	26258.61	2236	102631
	(13723.39)	(14221.9)	(12806.63)		
Drugs per outpatient visit	553.673	558.033	546.146	44.49	2494.81
	(219.152)	(230.685)	(197.580)		
Total wage expenditure	3.94e+07	3.89e+07	4.03e+07**	5274461	9.90e+07
	(1.56e+07)	(1.52e+07)	(1.62e+07)		
External market factors					
Population catchment area	22 868.34	21865.3	24600.05***	3293	62847
	(10375.43)	(9370.286)	(11718.82)		
Nearest clinic (km)	4.692	5.093	4.000***	0.001	17.252
	(1.890)	(2.001)	(1.441)		
Nearest district hospital (km)	10.696	11.113	9.977***	0.001	30.421
	(5.799)	(0.159)	(0.204)		
Service-mix					
Deliveries (%)	0.018	0.019	0.017***	0.003	0.093
	(0.011)	(0.011)	(0.010)		
Antenatal care (%)	0.033	0.034	0.032	0.004	0.147
	(0.017)	(0.018)	(0.016)		
Prenatal care (%)	0.031	0.031	0.030	0.00	0.245
	(0.026)	.026(0)	(0.024)		
Malaria (%)	0.125	0.133	0.111***	0	0.617
	(0.119)	(0.121)	(0.113)		
Respiratory (%)	0.214	0.217	0.211	0.015	0.724
	(0.098)	(0.103)	(0.089)		
Diarrhea and worms (%)	0.084	0.086	0.081**	0.007	0.409
	(0.051)	(0.055)	(0.044)		
Oral/eye/ear (%)	0.044	0.045	0.043	0.000	0.335
	(0.039)	(0.042)	(0.034)		
Integrated management of	0.099	0.098	0.100	0.00	0.846
childhood illness (IMCI) (%)	(0.062)	(0.062)	(0.063)		
HIV (%)	0.001	0.001	0.001*	0.00	0.058
	(0.002)	(0.002)	(0.001)		
Observations	2 113	480	775	1338	

Notes: Summary statistics of the structural inputs and contextual factors included in the analysis. Column 1 provides average values for all health centers in the sample, whereas columns 2 and 3 present values for rural and urban clinics separately. Columns 4 and 5 show sample maximum and minimum values. *** p<0.01, ** p<0.05, * p<0.1.

4 Empirical Analysis

In this section, I study the differences in quality of care between rural and urban households. I use a linear regression to estimate disparities in quality. Furthermore, I investigate how much of the inequality can be attributed to differences in health infrastructure and external contextual variables. To do this, I decompose the quality gap by comparing estimations with and without controls for these factors. This strategy has previously been used by Fafchamps et al. (2009) to investigate the importance of job sorting in African labor markets.

I estimate the following regressions:

$$ln(Q)_{it} = \alpha_2 Rural_i + \delta_2 W_{it} + \zeta_t + \epsilon_{it} \tag{1}$$

$$ln(Q)_{it} = \alpha_3 Rural_i + \delta_3 W_{it} + \gamma_3 X_{it} + \zeta_t + \epsilon_{it} \tag{2}$$

$$ln(Q)_{it} = \alpha_4 Rural_i + \delta_4 W_{it} + \gamma_4 Z_{it} + \zeta_t + \epsilon_{it} \tag{3}$$

$$ln(Q)_{it} = \alpha_5 Rural_i + \delta_5 W_{it} + \beta_5 X_{it} + \gamma_5 Z_{it} + \zeta_t + \epsilon_{it} \tag{4}$$

where $(Q)_{it}$ denotes the quality score of health clinic i at time t. $Rural_i$ is a dummy that takes the value 1 if a health center is situated outside the 15 km buffer zone that surrounds one of the 10 major cities in Rwanda. W_{it} is a vector of covariates controlling for health center case mix, thus controlling for the disease burden of each clinic. This is important since patients with more complicated symptoms might choose particular healthcare providers, putting additional pressure on clinic resources and affecting quality, confounding the true difference in provider quality with differences in patient characteristics. As a result, I control for case mix in all specifications. X_{it} is a vector of structural inputs at time t, including total wage expenditure, number of beds, total number of outpatient visits during one year, and the average cost of drugs per visit.

When controlling for clinic size, the total number of outpatient visits can be viewed as a proxy for the efficiency of the clinic in producing outpatient visits, as well as the demand for healthcare services. As previously mentioned, the total wage expenditure is a proxy for number of health workers, and drug expenditure per outpatient visit measures the relationship between access and use of drugs and quality. Z_{it} is a vector of contextual factors, including the total population in health center catchment area, and the distances to the closest neighboring clinic and district hospital. The population size of the catchment area is used as a proxy for the demand of health services faced by the health centers. When controlling for demand, the distance to nearest clinic also aims to measure market competition. Additionally, this vector controls for access to provider knowledge proxied by the geographic distance to the closest district hospital. The majority of health centers do not have medical doctors (Collins et al., 2011). Medical doctors from the district hospitals periodically travel around their corresponding districts to support the health center health staff. I control for potential changes in the quality scores common to all health clinics by including year fixed effects, δ_t. All standard errors are clustered at the health facility level.

By comparing α across the different models it is possible to decompose the urban-rural quality gap into portions attributed to the different groups of factors. All estimations control for variation in disease burden across rural and urban regions. Comparing (α_1–α_2) allows us to evaluate how much of the gap is due to differences in the structural inputs between rural and urban areas, and comparing (α_1–α_3) estimates how much can be attributed to differences in external market factors. Comparing (α_1–α_4) provides an estimate of how much of the disparity in quality between rural and urban clinics is jointly explained by differences in both structural and external factors.

5 Results

Table 3 shows the estimated quality gap between rural and urban health centers, using the estimated models presented in equations (1–4). Column 1 controls exclusively for year fixed effects and disease burden. The results indicate a small but significant difference in the quality of health services between rural and urban health clinics. Rural health clinics have on average 4.2 points lower quality score than urban facilities, representing approximately 0.3 standard deviations or 1% at the mean. Although the data do not allow me to further investigate the implications of the difference in quality scores in terms of actual health services, appendix table A1 can provide some guidance on what this quality gap could mean. For example, the difference in quality scores between rural and urban clinics could potentially correspond to any of the following:(i) the absence of a water source; (ii) a lack of access to skilled health personnel during deliveries or the availability of delivery emergency kits, (iii) incorrect management of cases with complications during antenatal care visits, or (iv) incorrect storage of vaccines, or v) medications being out of stock, such as a lack of tracer drugs. Again, this description suggests that even a relatively limited difference in actual quality scores (4.2 out of the total score of 460) could potentially mask important differences in the quality of health services.

In columns 2–4, I decompose the quality gap into parts that can be attributed to differences in structural inputs and external market factors. Overall, the results indicate that medical infrastructure and external market factors do not explain much of the rural and urban quality gap. Differences in structural inputs explain just over 12% of the difference in quality between rural and urban health centers (column 2), external factors explain approximately 11% of the gap (column 3), and internal and external factors jointly explain about 8% of the quality gap (column 4). Comparing the estimates in columns 2 and 4 suggests that the structural inputs explain a smaller portion of the quality gap when the contextual factors are controlled for.

The results are in line with recent evidence of the potential of clinic readiness to predict the quality of clinical process. This research suggests that the capacity of a health clinic to produce quality says little about the quality of the actual care provided (Leslie, Sun, & Kruk, 2017; Das & Hammer, 2014).

Table 3: Total score: general quality score

	(1)	(2)	(3)	(4)
Rural	-4.610***	-4.051***	-4.111***	-4.230***
	(1.280)	(1.171)	(1.273)	(1.189)
Constant	418.3***	196.0***	378.6***	212.5***
	(3.256)	(26.95)	(16.27)	(27.34)
Observations	2,113	2,113	2,113	2,113
R-squared	0.088	0.167	0.114	0.176
Year FE	Yes	Yes	Yes	Yes
Case mix	Yes	Yes	Yes	Yes
Structural inputs	No	Yes	No	Yes
External context	No	No	Yes	Yes

Notes The results from estimating equations (1–4) for the general quality score. Column 1 controls for time fixed effects and case mix; column 2 adds controls for health infrastructure by including number of beds, population catchment area, total wage expenditure, and drugs per visit; and column 3 controls for external factors such as catchment area population and distance to closest clinic and district hospital. Column 4 includes a complete set of all covariates. Standard errors are clustered at household level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4 repeats the analysis in Table 3 using the patient-centered quality score, which almost exclusively contains quality indicators focusing on the clinical process. The results indicate that the patient-centered quality score is significantly lower among rural clinics than urban clinics. Rural health centers receive approximately 2.1 lower quality scores than those in urban areas, corresponding to 1% of the mean and approximately 0.25 standard deviations. As with the general quality score, columns 2–4 suggest that structural inputs and external market factors explain only a reduced share of the difference in quality scores between rural and urban health centers. Comparing the estimates in columns 1 and 4, suggests that case mix, health infrastructural factors, and external market factors jointly explain approximately 11% of the quality gap between rural and urban areas (α_2–α_5).

Table 4: Total score: patient-focused quality score

	(1)	(2)	(3)	(4)
Rural	-2.378***	-2.098***	-2.101***	-2.113***
	(0.637)	(0.603)	(0.653)	(0.627)
Constant	197.9***	110.2***	172.4***	111.6***
	(1.542)	(13.44)	(8.714)	(13.91)
Observations	2,113	2,113	2,113	2,113
R-squared	0.057	0.102	0.078	0.103
Year FE	Yes	Yes	Yes	Yes
Case mix	Yes	Yes	Yes	Yes
Structural inputs	No	Yes	No	Yes
External context	No	No	Yes	Yes

Notes: The results from estimating equations (1–4) for the patient-focused quality score. Column 1 controls for time fixed effects and case mix; column 2 adds controls for health infrastructure by including number of beds, population catchment area, total wage expenditure, and drugs per visit; and column 3 controls for external factors such as catchment area population and distance to closest clinic and district hospital. Column 4 includes a complete set of all covariates. Standard errors are clustered at household level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

As described earlier, a number of health centers were not included in the analysis due to missing data. Further analysis of the omitted health facilities indicates that these health centers differed significantly from those included in the analysis. They were on average smaller and more likely to be in rural areas than those included in the study sample. The nonrandom exclusion of health centers is likely to contribute to underestimating the quality gap between rural and urban areas.

5.1 Extensions and robustness

In this section I give additional estimations of the quality gap between rural and urban clinics, using alternative strategies. The analysis in this section provides a deeper understanding of the nature of the disparities in quality, as well as the association between distance to cities and the quality of care, and evaluates the robustness of the results.

Appendix table A3 reports results from estimating equation (4) with alternative definitions of rural households. Columns 1–3 show estimations using cutoff distances of 10, 20, and 25 km. The results suggest that the estimated variation of quality between rural and urban households changes as I vary the distance cutoffs but remains negative for all definitions of rural households. The estimated quality gap decreases as the distance of the cutoff from a city increases. The 10 largest urban cities and the different cutoff distances are presented in the map in appendix figure A1. A comparison of the estimated quality gaps in columns 1–3 of table A3 suggests that the quality of care is relatively higher in the very center of a city. Furthermore, there is no significant difference in quality between the semiurban clinics within the 10–25 km buffer zone and those situated in remote rural areas farther than 25 km from the city. The results indicate that the association between distance to urban centers and quality is not linear and that the gap in quality of care is driven by a significantly higher quality at urban health clinics.

In column 4, I further investigate the relation between distance from cities and the quality of care, assuming a nonlinear effect of the distance between the city centers and the health clinics. The results confirm that the quality of care is significantly higher at health centers situated within 10 km of the center of a city an at clinics situated within a zone 10–25 km from the city center. Column 5 measures the association between the quality of care and the geographic distance from a health center to Kigali. The city has a high density of healthcare providers and expertise. For example, half of all eight referral hospitals in the country are situated in Kigali City. The estimates in column 5 support previous results by showing a negative association between the distance to Kigali and health center quality score, that is, distance from urban areas is negatively associated with the quality score. A 1% increase in the distance between a health center and Kigali is associated with a 2-point decrease in the quality score. The estimates in column 6 show that the Kigali effect remains even when I control for rural versus urban location. The results suggest that the distance from Kigali is negatively associated with health quality among urban as well as rural clinics.

A large share of the health centers have received quality scores that exceed the 80% ceiling used by the Rwandan government to define high-quality care. Appendix figure A2 shows the sample distribution of all separate quality indicators that make up the total quality scores. The graphs reveal that the majority of the indicators have a right-censored distribution; that is, a number of health centers have received a full quality score, suggesting that the quality measures are censored at the maximum score. The censored quality indicators imply that the true quality of each health clinic is not fully observed, meaning that I do not observe the variation in quality among clinics beyond the maximum score of the PBF indicators. As a result, the variation in quality scores across health centers comes from the lower tail of the distribution. The distributions reveal that despite high average quality scores, several clinics have low scores for one or more individual indicators. Given the construction of the quality scores, a clinic with important deficiencies in relation to a few specific quality indicators could receive the same total score as a clinic that has a relatively high quality score across all indicators. Although these two clinics could vary significantly in the overall levels of quality and patient security, such difference would not be captured by the total quality score.

I adjust for the top coding of the quality indicators by constructing binary response variables that take the value 1 if a quality indicator is below 80% of the total score, and zero otherwise. Although the true quality score for each clinic is unobserved in the data, the binary variable indicates whether the quality was above or below the cutoff point of good quality care. I use a probit model to estimate the likelihood that a clinic had at least one quality indicator with a score less than 80%. In appendix table A3, column 8, I present the results from estimating equation 4 using a probit model. The model estimates the difference between rural and urban clinics in having at least one quality score that is less than 80%. The results suggest that rural clinics are significantly more likely than urban clinics to have a low score in at least one area. In column 9, I construct a count variable that indicates the number of total quality indicators below 80% that a health clinic has received. I use a Poisson regression model to estimate the difference in the number of low-quality indicators between rural and urban clinics. The results indicate that a rural clinic has 0.11 additional low-score indicators compared with urban clinics.

In column 7, I estimate the baseline equation 4 using data from 2012—18. The data from 2012 were excluded from the main estimations due to low quality during this first year of data reporting. The results suggest that the estimated quality gap is stable when the additional year is included.

Appendix table A4 repeats the estimations presented in table A3, using the patient-centered quality score. The results are in line with the estimations of the general quality score, with the exception that the estimated quality gap is smaller for the patient-centered score. Again, the quality gap is larger when I consider a more restrictive definition of rural clinics (10 km cutoff, column 1) compared with longer distance cutoffs (columns 2 and 3). The results in column 4, support a nonlinear effect of distance between the health center and the center of a larger city. Columns 5 and 6 suggest that there is a negative association between quality and the distance from Kigali. As with the general quality score, this negative association remains when I control for urban-rural location. Columns 8 and 9 indicate that rural clinics are more likely to have reported at least one quality indicator below 80%.

Finally, appendix table A5 estimates potential development in the disparities in quality over time. The results suggest that there has been no significant change in the quality gap over time.

5.2 Interpreting the urban-rural quality gap

In this section I analyze the quality gap between rural and urban areas from a financial perspective, using the PBF reimbursement formula. The reimbursement formula includes the quality score as a multiplayer that adjusts the funding according to the quality level of the health services. This allows me to use the estimated gap in the quality score between rural and urban clinics to estimate the financial implications of such quality gaps. The financial implications of differences in quality provide additional information on how to interpret the meaning of the variation in quality measures across urban and rural health clinics.

The PBF reimbursement formula defines the total financial funding for each healthcare provider by multiplier the number of performed healthcare services covered by the scheme with a service-specific unit cost defined by the Ministry of Health. This amount is then multiplied by the PBF quality score,[3] providing a financial measure of differences in quality. I use the PBD reimbursement formula to estimate the average funding related to six health services covered by the PBF scheme: the number of women who received four antenatal care visits, newborns who received 4 postnatal care visits within six weeks of their birth, deliveries at health centers, and preventive and curative consultations. Importantly, the PBF includes a large number of reimbursable indicators that

[3]The following equation describes the PBF payment formula:

$$Bonus_{it} = Quality_{it} * \sum_j P_{jt} N_{ijt}$$

where i indexes health center, j a reimbursable service, and t time periods. $Quality_{it}$ is the quality score resulting from the quarterly independent evaluations at each health center, P_{it} is the unique unit cost for each reimbursable health service in time t, and N_{ijt} is the number of validated and reimbursable health services delivered at a health center in time t (Ministry of Health, 2018b).

generate PBF payments, and my calculations include only a portion of all reimbursable treatments.

Table 5 presents the estimated PBF funding in relation to each separate health service. Columns 1–4 show the estimated financial funding for an average clinic with an average quality score that provides an average number of health services, and columns 5–6 describe the financial implications of the PBF funding. The total amount of financial compensation is estimated by multiplying the service-specific unit cost by the number of treatments executed. The amount is then conditioned on quality by multiplying that amount by 0.88, the average quality score in the sample.

The results show that the total PBF payment corresponding to the six health services sums to Rwf 1 955 101 (USD 2,000). The reimbursement represents approximately one yearly salary of a nurse with the highest educational level (A1), about 2.5 million RwF. Furthermore, according to the national health center cost model, the average total cost for preventive services at health centers was Rwf 41,106,580 (USD 42,000), and the average total cost of salaries for technical staff was Rwf 26,559,747 (USD 27,000). The average total cost of a health center was Rwf 120 million (Collins et al., 2011). The results indicate that the PBF funding represents a relatively small share of the overall health center budgets.

Columns 5–6 show the financial implications of the gap in quality between rural and urban areas. I base my calculations on the estimated quality gap estimated in table 3. Given the calculation formula of the performance-based financing, the estimated 1% gap in the quality score resulted in a difference in PBF funding of the same magnitude. In relation to the six covered health services, the PBF scheme results in a difference in funding of approximately Rwf 19,500 due to the quality gap. The monetary difference caused by the quality gap represents less than 1% of the yearly salary of an A1 nurse.

Table 5: Financial reimbursement within the PBF scheme

Health service	(1) Unit costs (Rwf)	(2) Avg. services	(3) Score (%)	(4) Total reimbursement	(5) Quality gap (%)	(6) Diff. reimbursement (Rwf)
4 antenatal visits	1139	233	88	233 541	1	2 335
4 postnatal visits within 6 weeks	1139	252	88	252 585	1	2 528
Deliveries HC	1773	390	88	608 493	1	6 084
Preventive and curative consultations	30	30 079	88	794 086	1	7 940
Metal health consultations	296	225	88	58 608	1	586
Screening NCD and cancer	59	150	88	7 788	1	78
Total				1 955 101		19 551

Notes: Column 1 shows the unit costs for each treatment, fee for services; column 2 lists the yearly average number of each service performed at a health center; column 3 displays the average quality score among health centers; and column 4 presents the estimated amount of financial compenastion from the PBF scheme according to the payment formula. The figures are from my own calculations based on HMIS and PBF technical information (Ministry of Health, 2018b).

6 Conclusions

During the last decade, access to healthcare has improved greatly for people in many low-income countries. However, there is a global recognition that potential health achievements from increased access to healthcare have not been realized because of low quality of health services. Moreover, numerous studies have documented large disparities in the quality of health services both between and within countries (Kruk et al., 2018). Such inequalities represent important barriers in the process of achieving equity in health across population groups, one of the overall development goals defined by the Sustainable Development Goals.

In this paper, I have examined the quality gap between rural and urban areas in Rwanda. This study finds that there is a statistically significant difference in quality between rural and urban health clinics and that rural clinics systematically underperform in comparison with urban clinics in all dimensions. Rural health centers have lower quality scores than urban clinics, and the gap represents approximately 1% of the mean score. The results suggest that the estimated quality gap could potentially mask important differences that could have considerable effects on the quality and security of the health services provided to patients. Based on PBF quality check lists, the average gap in the general quality score could, for example, mean that the rural clinic lacked access to a water source, essential vaccines, routines for infection prevention, or local anesthesia and saline, or that deliveries were not done by qualified personnel.

Although earlier empirical evidence on the quality gap between rural and urban areas is inconclusive, some studies have found important disparities in quality between rural and urban health facilities in the low-income setting (Kruk et al., 2018; Gage et al., 2017). While this paper supports these findings, it is important to consider the context of the study when interpreting the results. Importantly, the present analysis is based on quality measures designed for the Rwandan PBF scheme. The primary aim of these indicators was to provide a quality measure that could be used to define the level of reimbursement for health centers within this scheme. The Ministry of Health is likely to have designed quality measures that, in addition to evaluating the level of quality of health services, also considered the distribution of PBF funds across clinics. As a result, the quality indicators are likely to measure the minimum standards for receiving financial reimbursement rather than the actual quality of the care provided. The right-censored distribution of the quality indicators described in the previous section, suggests that this could be the case. This would result in an underestimation of the variation in service quality across regions, implying that the results presented in this analysis are likely to measure the gap in the supply of minimum-standard healthcare between rural and urban areas.

Following this line of reasoning, the PBF scheme would provide monetary incentives for health providers and facilities to reach an established minimum level of quality, but would not incentivize caregivers to improve quality of care beyond these minimum levels. Given that healthcare expertise and knowledge are often concentrated in urban areas, this reasoning could explain the small gap in quality scores between rural and urban health clinics. A restructure of the PBF performance indicators in order to increase the system's ability to identify and reward high-quality facilities, however, could lead to the reinforcement of initial quality differences and create barriers to improvement for health facilities with lower quality scores. As a result, the PBF system could contribute to increased inequities in health quality among geographic areas. Overall, the government's 80% ceiling for the quality scores implies that PBF contracts can contribute to improved healthcare services up to a certain point, but otherwise represent an inefficient policy tool to improve quality.

In addition to estimating differences in the quality of care between rural and urban clinics, I have examined how much of the quality gap can be attributed to differences in structural inputs and contextual factors. The results suggest that variation in inputs explain only a small share of the differences in quality scores. The results are in line with earlier work that has found no correlation between structural inputs and quality (Das et al., 2008; Das & Hammer, 2014; K. L. Leonard & Masatu, 2007). Instead, these studies suggest that provider effort represents the key determinant of care quality. Low effort implies that even when providers have the knowledge of how to correctly treat a patient, they often fail to do so. Low levels of provider effort are likely in the absence of accountability of care providers (Das & Hammer, 2014). This is important, since low levels of provider effort are likely to limit any potential effects of increased structural inputs beyond a certain point. In fact, studying the importance of structural inputs in markets with low provider effort potentially says little about the actual importance of structural inputs in explaining disparities in care quality.

This analysis adds important evidence to this discussion by studying the importance of structural and contextual factors in explaining differences in the quality of care in a health sector with a long experience of public accountability and monitoring of quality. The PBF scheme introduced accountability of quality of care into the health sector through both economic incentives and monitoring and evaluations. As a result, the economic cost of low effort is relatively high for each healthcare provider in the Rwandan healthcare sector compared with markets that lack formal accountability mechanisms. Furthermore, the Imihigo system provides additional monitoring by making local authorities accountable for service delivery.

While the results indicate that structural inputs account for a reduced share of the variation in quality between rural and urban areas, the results also suggest that the differences in structural inputs between rural and urban areas are small. This is an interesting finding since previous research has suggested that differences in inputs between rural and urban facilities are significant in many low-income countries (Leslie, Spiegelman, et al., 2017). One potential explanation for the relatively uniform distribution of structural inputs is the PBF system, which rewards investment in structural inputs both directly and indirectly through the quality indicators. However, an earlier evaluation of the PBF scheme in Rwanda found no evidence that the program had a significant effect on the increase of structural health infrastructure. Another potential explanation for the relatively uniform geographic distribution of health infrastructure across the country has been the inclusion of health related targets in the performance-based contracts signed between the president of Rwanda and local governments.

The results presented in this study suggest that statistically significant differences in quality between rural and urban areas remain after controlling for structural inputs. One potential explanation for this could be differences in knowledge between rural and urban healthcare providers. Unfortunately, lack of access to data on health worker knowledge prevents me from further investigating its importance in explaining the quality gap. However, if one is willing to consider the distance between Kigali and each health clinic as proxy for access to knowledge, specialized knowledge in particular, the estimations presented in column 5 in appendix table A3 provide supporting evidence for this idea. The results indicate that a decrease in access to knowledge is associated with lower quality of care. Although I cannot rule out that healthcare quality and provider knowledge may be explained by factors other than the distance between a health facility and Kigali, the results are in line with the expected relation between knowledge and quality.

Another plausible explanation would of course be variation in provider effort between rural and urban areas. Low provider effort results in a gap between provider knowledge and their performance during patient visits, referred to as the know-do gap (Das et al., 2008; Das & Hammer, 2007). In a market with public accountability, I expect the know-do gap to decrease. In fact, previous evidence has indicated that the "know-do gap" decreases as a result of the introduction of performance pay (Gertler & Vermeersch, 2012; Ngo et al., 2016), suggesting that the quality gap in Rwanda is less likely to be less attributed to low effort among providers than in markets without accountability. However, knowledge could remain a constraining factor for quality of care in healthcare sectors with accountability and uniform access to structural inputs.

Despite increased interest in the quality of healthcare in low-income countries during recent years, little is known about the exact factors that are correlated with disparities across a number of dimensions within these countries. Knowledge of what factors are associated with the quality of the care process is essential to design effective policy interventions that contribute to achieving the Sustainable Development Goals. Future investigations on how to increase quality and provide populations in low-income countries with equal access to high-quality care will continue to be a key issue on the global policy agenda.

References

African Strategies for Health. (2015). *Rwanda health private sector engagement assess-*

ment (Tech. Rep.). Kigali, Rwanda: U.S. Agency for International Development.

Barber, S. L., Gertler, P. J., & Harimurti, P. (2007). Differences in access to high-quality outpatient care in Indonesia: Lower quality in remote regions and among private nurses is a manifestation of the educational, policy, and regulatory frameworks upon which the indonesian health system is based. *Health Affairs*, *26*(Suppl2), w352–66.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J., & Vermeersch, C. M. (2011). Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: An impact evaluation. *The Lancet*, *377*(9775), 1421–28.

Binagwaho, A., Farmer, P. E., Nsanzimana, S., Karema, C., Gasana, M., de Dieu Ngirabega, J., ... Nyatanyi, T. (2014). Rwanda 20 years on: investing in life. *The Lancet*, *384*(9940), 371–75.

Byamukama, B., & Makonnen, N. (2012). *Performance contracts and social service delivery–lessons from Rwanda* (Rwanda Field Office Policy Brief). Kigali: African Development Bank.

Collins, D., Mukunzi, J., Jarrah, Z., Ndizaye, C., Kayobotsi, P., Mukantwali, C., ... Cros, M. (2011). *Rwanda health service costing: Health centre analysis* (Tech. Rep.). Kigali, Rwanda: USAID Integrated Health System Strengthening Project.

Das, J., & Hammer, J. (2007). Money for nothing: The dire straits of medical practice in Delhi, India. *Journal of Development Economics*, *83*(1), 1–36.

Das, J., & Hammer, J. (2014). Quality of primary care in low-income countries: Facts and economics. *Annual Review of Economics*, *6*(1), 525–53.

Das, J., Hammer, J., & Leonard, K. (2008). The quality of medical advice in low-income countries. *Journal of Economic Perspectives*, *22*(2), 93–114.

Das, J., Holla, A., Das, V., Mohanan, M., Tabak, D., & Chan, B. (2012). In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Affairs*, *31*(12), 2774–84.

Das, J., Holla, A., Mohpal, A., & Muralidharan, K. (2016). Quality and accountability in health care delivery: audit-study evidence from primary care in india. *American Economic Review*, *106*(12), 3765–99.

Donabedian, A. (1988). The quality of care: How can it be assessed? *Journal of the American Medical Association*, *260*(12), 1743–48.

Fafchamps, M., Söderbom, M., & Benhassine, N. (2009). Wage gaps and job sorting in African manufacturing. *Journal of African Economies*, *18*(5), 824–68.

Gage, A. D., Leslie, H. H., Bitton, A., Jerome, J. G., Thermidor, R., Joseph, J. P., & Kruk, M. E. (2017). Assessing the quality of primary care in Haiti. *Bulletin of the World Health Organization*, *95*(3), 182.

Gertler, P., & Vermeersch, C. (2012). *Using performance incentives to improve health outcomes.* Washington, DC, USA: World Bank.

Kalisa, I., Musange, S., Saya, U., & Kunda, T. (2015). *The development of community-based health insurance in Rwanda: Experiences and lessons* (Tech. Rep.). Kigali, Rwanda & Medford, MA, USA: University of Rwanda College of Medicine and Health Sciences, School of Public Health and Management Science for Health.

Kruk, M. E., Chukwuma, A., Mbaruku, G., & Leslie, H. H. (2017). Variation in quality of primary-care services in Kenya, Malawi, Namibia, Rwanda, Senegal, Uganda and the United Republic of Tanzania. *Bulletin of the World Health Organization*, *95*(6), 408.

Kruk, M. E., Gage, A. D., Arsenault, C., Jordan, K., Leslie, H. H., Roder-DeWan, S., … et al. (2018). High-quality health systems in the Sustainable Development Goals era: Time for a revolution. *The Lancet Global Health*, *6*(11), e1196–e1252.

Leonard, K., & Masatu, M. C. (2006). Outpatient process quality evaluation and the hawthorne effect. *Social science & medicine*, *63*(9), 2330–2340.

Leonard, K. L., & Masatu, M. C. (2007). Variations in the quality of care accessible to rural communities in Tanzania: Some quality disparities might be amenable to policies that do not necessarily relate to funding levels. *Health Affairs*, *26*(Suppl2), w380–92.

Leslie, H. H., Spiegelman, D., Zhou, X., & Kruk, M. E. (2017). Service readiness of health facilities in Bangladesh, Haiti, Kenya, Malawi, Namibia, Nepal, Rwanda, Senegal, Uganda and the United Republic of Tanzania. *Bulletin of the World Health Organization*, *95*(11), 738.

Leslie, H. H., Sun, Z., & Kruk, M. E. (2017). Association between infrastructure and observed quality of care in 4 healthcare services: A cross-sectional study of 4,300 facilities in 8 countries. *PLoS Medicine*, *14*(12), e1002464.

Liu, K., Subramanian, S., & Lu, C. (2019). Assessing national and subnational inequalities in medical care utilization and financial risk protection in Rwanda. *International Journal for Equity in Health*, *18*(1), 1–10.

Lu, C., Chin, B., Lewandowski, J. L., Basinga, P., Hirschhorn, L. R., Hill, K., ... Binagwaho, A. (2012). Towards universal health coverage: An evaluation of Rwanda Mutuelles in its first eight years. *PloS One*, *7*(6), e39282.

Ministry of Health. (2018a). *Grille d'evaluation de la quantite et de la qualite du centre de sante.* Kigali, Rwanda.

Ministry of Health. (2018b). *Performance based financing procedures manual for health facilities (hospitals and health centers)* (Tech. Rep.). Kigali, Rwanda: Ministry of Health.

Ministry of Health. (2019). *Health labor market analysis report* (Tech. Rep.). Kigali, Rwanda: Ministry of Health.

Ngo, D. K., Sherry, T. B., & Bauhoff, S. (2016). Health system changes under pay-for-performance: the effects of rwandas national programme on facility inputs. *Health policy and planning*, *32*(1), 11–20.

Pose, R. R., & Samuels, F. (2011). *Rwanda's progress in health: Leadership, performance and insurance* (Tech. Rep.). London: Overseas Development Institute.

Rajashekar, A., Richard, M., & Stoelinga, D. (2019). *Economic geography of rwanda* (Laterite study). International Growth Centre, Kigali. Retrieved from `https://www.theigc.org/wp-content/uploads/2019/08/Rajashekar-2019-Final-report.pdf` (accessed April 2021)

Scheil-Adlung, X. (2015). *Global evidence on inequities in rural health protection: New data on rural deficits in health coverage for 174 countries* (Tech. Rep.). Geneva, Switzerland: International Labour Organization.

Sharma, J., Leslie, H. H., Kundu, F., & Kruk, M. E. (2017). Poor quality for poor women? inequities in the quality of antenatal and delivery care in Kenya. *PloS One*, *12*(1). (e0171236)

Versailles, B. (2012). *Rwanda: Performance contracts (imihigo)* (Country Learning Notes). London: Overseas Development Institute.

World Bank. (2020). *World development indicators.* Retrieved from `https://data.worldbank.org/indicator` (accessed June 2020)

World Health Organization. (2017). *Primary healthcare systems (PRIMASYS): Case study from Rwanda, abridged version* (Tech. Rep.). Geneva, Switzerland: World Health Organization.

World Health Organization. (2018). *Delivering quality health services: A global imperative for universal health coverage*. Geneva, Switzerland: World Health Organization.

World Health Organization, & UNICEF. (2015). *Water, sanitation and hygiene in health care facilities: status in low and middle income countries and way forward.* Geneva, Switzerland. Retrieved from `https://apps.who.int/iris/bitstream/handle/10665/154588/9789241508476_eng.pdf` (accessed April 20 2021)

Appendix

Table A1: Descriptions of the quality indicators

Quality indicator	Description
Antenatal care	- Functional equipment and drugs available (including folic acid, iron, and Mebendazole) - Observation of 5 new patients on their first visit: i) Examination (medical history, HIV testing, cervical and breast cancer screening) and physical examination (height and weight, mid-upper arm circumference, edema assessment, breast examination) ii) Complementary examination (hemoglobin, syphilis, albumin, glycosuria, blood grouping) iii) Immunization according to schedule iv) Correct prescription of a) iron and folic acid, b) Mebendazole (from the second quarter), c) insecticide-treated mosquito net v) Management of cases with risk factors: a) risk factors identified, b) decision made correctly according to the consultation sheet (CPN), c) information communicated to the woman - Observation of 5 cases on he second and third visits: i) Obstetrical examination ii) Administration of the tetanus vaccin according to directives iii) Correct prescription of: a) iron for pregnant women, b) folic acid, c) Mebendazole (from the second trimester) iv) Management of cases with complications: a) complications identified b) decision taken correctly according to the flowchart, c) information communicated to the woman v) Existence of a delivery plan: a) detection of signs of hazards: (abnormal presentation, edema, hypertension, anemia), b) guidance for delivery

Family planning	- Contraceptive methods: a) Contraceptive availability with theoretical stock corresponding to the physical stock, b) quantified alert thresholds determined and respected - Existence of a system of feedback and search of cases that stopped using contraceptive Analysis of 10 fact sheets: - Reason for the methods chosen by the client, methods for which the client is eligible, and the method offered in relation to one indicated by the interview, medical history, physical examination. - Monitoring and follow-up: check in the register and the card if the part followed was correctly and completely filled.	
Vaccination	- Availability of vaccines and diluents (BCG, OPV / IPV, MR, PNEUMO, Pentavalent, ROTA TEQ, VAT and diluents): a) physical presence of unexpired antigens with label, b) nothing was of out of stock during the last 3 months - Cold chain: a) temperature of the fridge within the limits (between +2C–+8C) b) no break in the cold chain during the last 3 months Direct observation of 4 children in receiving vaccination - Systematic BCG scar search - Vaccine preparation: a) vaccine control pellet (VVM) in good condition, b) dilution technique respected, c) use of a self-locking syringe, d) appropriate dose - Injection and asepsis: a) cleaning the injection site with cotton soaked with water b) use of appropriate routes and injection sites, c) correct use of receptacles and garbage cans - Systematically recalled side effects - Correct and complete registration: a) vaccination record b) vaccination record c) scorecard	

Tuberculosis	- The stock of antituberculosis medication is managed correctly - The data collection tools used by the TB Division are available and in use - The updated TB Division data collection tools are correctly completed according to standard operating procedures for monitoring and evaluation: a) laboratory voucher, b) base register, c) treatment sheets - Glutathione peroxidase result availability time at the requesting site for all samples sent : results available at the requesting site within 4 days of date of dispatch - HIV test for suspected TB cases - Contact examination done at the beginning and at the end of TB treatment for theory of planned behavior Analysis of chosen cases on the sheets and register: - Correct management for any case diagnosed TB according to the national directives of the TB division (choose at random 2 cases): a) laboratory voucher (TPB +) or proof of diagnosis must be attached to the patient's file b) treatment in accordance with the categorization of patients c) control sputum if indicated done in accordance with the bacteriological monitoring algorithm d) HIV test carried out e) for TB/HIV co-infected cases, if ARV has been initiated according to the guidelines of the HIV and TB program - FOSA entered its full-time quarterly report to R-HIMS before the fifth day of the month following the quarter evaluated	

Hygiene	- Presence of latrines and showers that a) are usable, b) have door that closes, c) have available water and soap, d) have toilet paper and water, e) have covered pit - Clean rooms, courtyard and grounds - All beds with sheets, blanket, long lasting bed nets, beds with plastic mattresses and not torn - Absence of organic waste, syringes and dangerous products in the yard, in the rooms, or any other easily accessible place of the CS enclosure by the population - Availability of a water source (running water or well or pump or castle / water tank) - Cleanliness in the delivery room	
Laboratory	- Available and functional equipment and materials - No shortage of reagents and consumables - Separate waste management (sharps, non-infectious and infectious objects) with color identifying each type of bin - Presence of the wastewater evacuation system guaranteeing environmental protection - Existence of the quality control register of rapid malaria tests carried out by community health workers in the villages - Existence of a separate and well ventilated room for collecting, spreading and coloring sputum samples	
Financial management	- Tariffs for procedures, laboratory, medicines and consumables, ambulance: a) displayed, b) legible, c) at reception and at the cash register d) respected - Receipts completely filled in with proof of daily payment - Surprise control of the cash register - Daily revenue journal: a) available and up to date b) concordant with the receipts, c) writing is legible and has no correction fluid d) daily payments correspond with the daily revenue journal - Expenditure journal available an up to date - Bank cash book: a) available b) consistent with the supporting documents for expenditure and bank statements and the revenue journal c) up to date d) without too many entries or spaces - Availability of monthly and annual financial reports sent to the competent authorities	

Pharmacy	- Pharmacy premises conforming to standards - General or stock pharmacy including all drugs: essential drugs, ARV, IMCI drugs, malaria, FP, availability of at least one psychotropic drug per category (neuroleptic, antiepileptic, and antidepressant) - Pharmacy cleanliness: no dust on shelves and products, no cobwebs, no water, no expired products in the pharmacy, no other things (food, juice etc.) in the fridge - Storage in accordance with standards - Availability of drugs and tracer consumables (take a sample of 10 products and also observe on the shelves) - Compliance with the procedure for destroying expired products - Observation of the dispensing pharmacy: - Equipment and materials available and used: a) water filter, b) spatulas, c) spoons, d) clean cups, e) cutting object, f) packaging - Use of tools and up-to-date filling - Hygiene rules observed when handling medicines: a) use of spatulas and spoons, b) medicine packaging, c) disposable towel to clean spoons - Administration of drugs to the outpatient: a) give the first dose of drug, b) explain how to take the drugs correctly and systematically reminded at the time of distribution, c) pack and label the drugs for patients according to the prescriptions	
Community Health Workers	- Calendar of quarterly activities including supervision, monthly CHW meeting, training & retraining, quarterly evaluation of CHWs - Quarterly report on community health activities carried out, sent to the sector/district hospital - Presentation of supervisory feedback during the monthly CHW meeting (see meeting report) - Medication management and tools for community health activities: a) drug stock sheet completely and correctly completed, b) absence of a break in the tools of the program, c) no out-of-stock drugs or consumables, d) concordance between physical stock and theoretical stock - Minutes of the 3 monthly meetings of the quarter evaluated on the analysis of community data with available CHWs - Referral to the community: a) existence of a register of cases referred by the community, b) concordance between register of reference cases (choose 3 coupons at random), c) feedback from the health center to the CHW notified in the register	

General	- Minutes of the 3 monthly meetings of the last three months of the quality improvement team available. Each report must fulfill the criteria listed in element 1 of the checklist - Minutes of the 3 monthly meetings of the last three months of available staff. - A quarterly meeting report of costs with acknowledgment of receipt of the sector. - Evaluation of the implementation of the activities of the previous quarter of the quarterly plan of the health center - Existence of budget line on equipment maintenance in the annual action plan - Clean water stations with liquid soaps available in consultation rooms, hospital rooms and laboratories	
Deliveries	- Conditions of confidentiality in the waiting room, during delivery and postpartum - Equipment, and material available and functional - Prevention of infections - Drugs and consumables available - Analysis of 10 partographs chosen at random: a) partograph filled according to the norms, b) decision made in case of exceeding the alert line within one hour, c) delivery by qualified personnel	

Source: (Ministry of Health, 2018a)

Table A2: Analysis: missing values

Variables	(1) All	(2) Rural	(3) Urban
Rural (15km)	0.0720**		
	(0.0332)		
ln(Beds)	0.0891***	0.0551	0.157***
	(0.0311)	(0.0390)	(0.0551)
ln(Population catchment area)	0.0297	0.0406	-0.0121
	(0.0476)	(0.0629)	(0.0675)
ln(Total outpatient visits)	-0.0848*	-0.0737	-0.115
	(0.0480)	(0.0576)	(0.0765)
ln(Medicine costs /outpatient visit)	0.0179	0.0300	0.0139
	(0.0325)	(0.0380)	(0.0612)
ln(Wages)	-0.0586	-0.147**	0.151**
	(0.0441)	(0.0583)	(0.0595)
Deliveries (%)	-3.250	-2.745	-1.943
	(2.658)	(3.127)	(4.907)
Malaria (%)	-0.333**	-0.561***	0.275
	(0.161)	(0.197)	(0.249)
Respiratory (%)	-0.309**	-0.391**	-0.0368
	(0.143)	(0.182)	(0.215)
Diarrhea/parasites (%)	-0.254	-0.215	-0.363
	(0.286)	(0.335)	(0.493)
Integrated management of	0.450**	0.473*	0.170
childhood illness (%)	(0.207)	(0.283)	(0.271)
Oral/ear/eye/PCT (%)	0.418	0.137	1.105*
	(0.277)	(0.296)	(0.564)
HIV (%)	-7.254*	-8.134*	-6.702
	(4.224)	(4.766)	(6.739)
Prental care (%)	0.802	0.529	1.173
	(0.697)	(0.752)	(1.212)
Antenatal care(%)	-4.709***	-6.004***	-2.029
	(1.381)	(1.577)	(3.162)
Constant	1.609**	3.121***	-1.745*
	(0.687)	(0.939)	(0.966)
Observations	2,616	1,709	907
R-squared	0.083	0.113	0.106
Year FE	Yes	Yes	Yes

Notes: Standard errors are clustered at the health center level. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Robustness analysis:- general quality score

Variables	(1) LPM	(2) LPM	(3) LPM	(4) LPM	(5) LPM	(6) LPM	(7) LPM	(8) Probit	(9) Poisson
Rural 10km	-6.304***			-5.578***					
	(1.285)			(1.334)					
Rural 20km		-1.743							
		(1.254)							
Rural 25km			-3.850**	-2.656					
			(1.578)	(1.626)					
ln(Km Kigali)					-2.073**	-1.745*			
					(0.938)	(0.909)			
Rural 15km						-3.969***	-4.552***	0.243***	0.113***
						(1.183)	(1.162)	(0.0831)	(0.0429)
Constant	214.1***	208.0***	206.9***	212.0***	218.1***	219.8***	206.4***	8.869***	4.050***
	(27.28)	(27.57)	(27.43)	(27.17)	(28.17)	(27.93)	(25.30)	(2.053)	(1.000)
Observations	2,113	2,113	2,113	2,113	2,113	2,113	2,457	2,113	2,113
R-squared	0.182	0.165	0.172	0.186	0.169	0.180	0.195		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case mix	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Structural input	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
External factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The results from estimating equation 4 using different cutoff distances for the definition of rural health clinics: 10 km (column 1), 20 km (column 2), and 25 km (column 3). Column 4 estimates the quality gap between rural and urban health centers allowing for a nonlinear relation between quality and distance from urban areas. Columns 5 and 6 estimates the association between health facility quality score and distance from Kigali. Column 7 presents the baseline estimation (table 3, column 4) using data from 2012–18. Column 8 presents the results from estimating equation 4 using a probit model to estimate differences in the likelihood that a health center has at least one quality indicator below 80% between rural and urban clinics. In column 9, I use a Poisson estimation strategy to estimate the difference in the total number of quality indicators below 80% between rural and urban clinics. All standard errors are clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Robustness analysis: - patient–focused quality score

Variabes	(1) LPM	(2) LPM	(3) LPM	(4) LPM	(5) LPM	(6) LPM	(7) LPM	(8) Probit	(9) Poisson
Rural 10km	-3.284***			-3.078***					
	(0.642)			(0.656)					
Rural 20km		-0.805							
		(0.660)							
Rural 25km			-1.413*	-0.754					
			(0.822)	(0.838)					
ln(Km Kigali)					-2.090***	-1.939***			
					(0.433)	(0.425)			
Rural 15km						-1.823***	-2.228***	0.281***	0.292***
						(0.621)	(0.611)	(0.0843)	(0.0918)
Constant	112.5***	109.5***	109.1***	111.9***	119.0***	119.8***	108.4***	8.130***	8.087***
	(13.73)	(13.88)	(13.89)	(13.72)	(14.10)	(14.12)	(12.83)	(1.947)	(1.948)
Observations	2,113	2,113	2,113	2,113	2,113	2,113	2,457	2,113	2,113
R-squared	0.109	0.093	0.095	0.110	0.112	0.121	0.110		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case mix	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Structural input	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
External factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

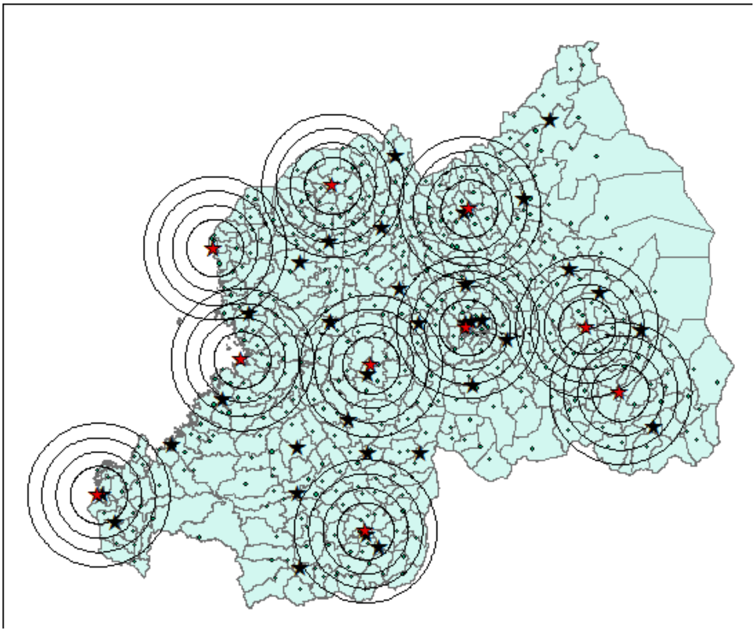
Notes: The results from estimating equation 4 using different cutoff distances for the definition of rural health clinics: 10 km (column 1), 20 km (column 2), and 25 km (column 3). Column 4 estimates the quality gap between rural and urban health centers allowing for a nonlinear relation between quality and distance from urban areas. Columns 5 and 6 estimate the association between health facility quality score and distance from Kigali. Column 7 presents the baseline estimation (table 4 column 4) using data from 2012–18. Column 8 presents the results from estimating equation 4 using a probit model to estimate differences in the likelihood that a health center has at least one quality indicator below 80% between rural and urban clinics. In column 9, I use a Poisson estimation strategy to estimate difference in the total number of quality indicators below 80% between rural and urban clinics. All standard errors are clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Changes over time in quality difference between rural and urban facilities

	(1)	(2)	(3)	(4)	(5)
Rural 15km	-5.445***	-4.831***	-5.127***	-3.623**	-4.552***
	(1.649)	(1.659)	(1.583)	(1.672)	(1.621)
2014 * Rural 15km	-0.875	-0.929	-0.363	-1.065	-0.553
	(1.388)	(1.457)	(1.508)	(1.503)	(1.545)
2015 * Rural 15km	1.976	1.229	2.376	0.870	1.913
	(1.895)	(1.990)	(2.013)	(2.070)	(2.118)
2016 * Rural 15km	1.102	0.829	2.124	0.287	1.258
	(1.951)	(2.046)	(2.052)	(2.068)	(2.083)
2017 * Rural 15km	2.580	2.074	3.250	1.128	2.229
	(2.171)	(2.305)	(2.317)	(2.304)	(2.355)
2018 * Rural 15km	-0.351	-0.460	0.321	-1.880	-0.827
	(1.979)	(2.073)	(2.051)	(2.072)	(2.076)
Constant	409.8***	417.9***	193.8***	368.0***	208.2***
	(1.285)	(3.799)	(40.48)	(22.69)	(41.72)
Observations	2,113	2,113	2,113	2,113	2,113
R-squared	0.034	0.104	0.190	0.137	0.206
Year FE	Yes	Yes	Yes	Yes	Yes
Case mix	No	Yes	Yes	Yes	Yes
Structural inputs	No	No	Yes	No	Yes
External context	No	No	No	Yes	Yes

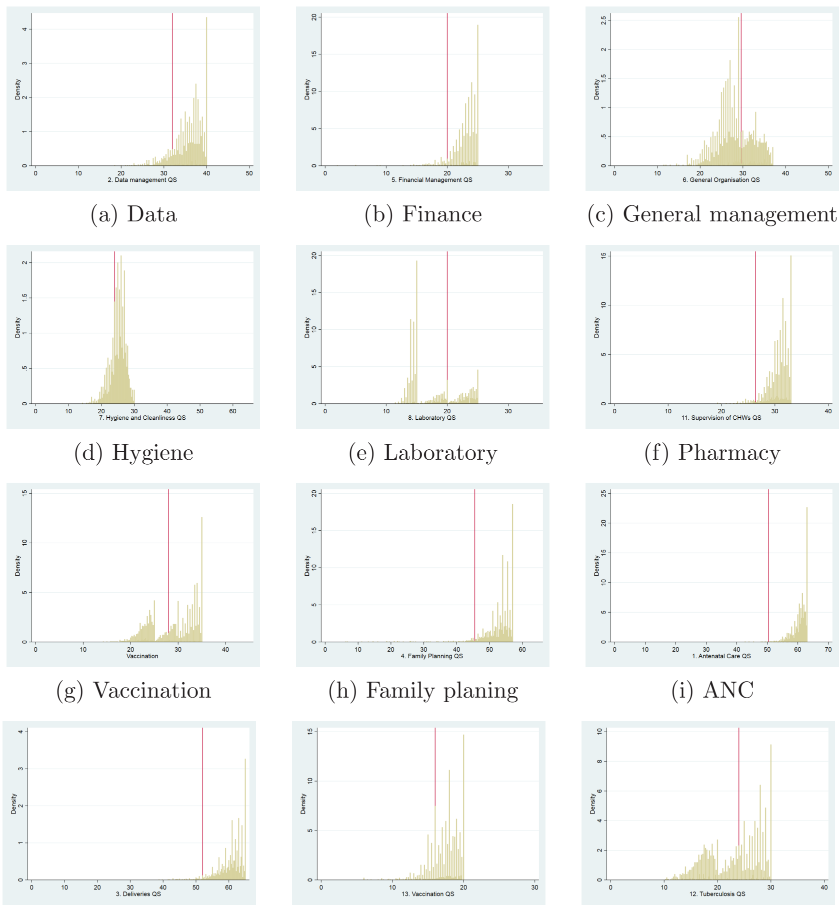
Notes: The results from estimating equation (1–4) are presented in columns 2–5 including rural-by-year interactions to account for differences in the quality gap between rural and urban clinics over time. Column 1 controls for time fixed effects, column 2 adds controls for case mix, and column 3 adds controls for health infrastructure by including number of beds, population catchment area, total wage expenditure, and drugs per visit. Column 4 controls for external factors, and casemix, and column 5 includes all previous covariates such as health infrastructure, external market factors, and casemix. All standard errors are clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1.

Figure A1: The 10 largest cities and buffer zones



Notes: The main specification uses a buffer zone with a radius of 15 km to classify rural and urban areas. In the robustness analysis, I include different buffer zones of 10, 20, and 25 km. The red stars indicate the location of the 10 major cities in Rwanda, while the black stars indicate the locations of district hospitals

Figure A2: Distribution of all quality indicators



(a) Data (b) Finance (c) General management

(d) Hygiene (e) Laboratory (f) Pharmacy

(g) Vaccination (h) Family planing (i) ANC

(j) Deliveries (k) Tuberculosis (l) CHW

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 arbetstidsförkortning (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
226. **Ruhinduka, Remidius Denis** (2015) Essays on Field Experiments and Impact Evaluation
227. **Persson, Emil** (2016) Essays on Behavioral and Experimental Economics, Cooperation, Emotions and Health.
228. **Martinangeli, Andrea** (2017) Bitter divisions: inequality, identity and cooperation
229. **Björk, Lisa** (2017) Essays on Behavioral Economics and Fisheries: Coordination and Cooperation
230. **Chegere, Martin Julius** (2017) Post-Harvest Losses, Intimate Partner Violence and Food Security in Tanzania
231. **Villalobos-Fiatt, Laura** (2017) Essays on forest conservation policies, weather and school attendance
232. **Yi, Yuanyuan** (2017) Incentives and Forest Reform: Evidence from China
233. **Kurz, Verena** (2017) Essays on behavioral economics: Nudges, food consumption and procedural fairness
234. **Yashodha, Yashodha** (2017) Contract Choice and Trust in Informal Groundwater Markets
235. **Meles, Tensay Hadus** (2017) Power Outages, Increasing Block Tariffs and Billing Knowledge
236. **Mühlrad, Hanna** (2018) The Impact of Reproductive and Birth Technologies on Health, Fertility and Labor Outcomes
237. **Gakii Gatua, Josephine** (2018) Primary Health Care Interventions and Social Ties in Kenya
238. **Felgendreher, Simon** (2018) Essays on behavioral economics and the effects of the colonial rule on Java
239. **Demeke, Eyoual** (2019) Essays on Environmental and Behavioral Economics
240. **Kiss, Tamas** (2019) Predictability in Equity Markets: Estimation and Inference
241. **Khomenko, Maksym** (2019) Essays on the Design of Public Policies and .

Regulations
242. **Munkanjari, Samson** (2019) Climate Policy and Financial Markets
243. **Schürz, Simon** (2020) Economic and Intergenerational Decision-Making in Families
244. **Lau, Debbie** (2020) Empirical Essays on Education and Health Policy Evaluation
245. **Ho, Hoang-Anh** (2020) Essays on Culture, Institutions, and Economic Development
246. **Rubio Ramos, Melissa** (2020) The Economics of Coercive Institutions, Conflict, and Development
247. **Lindahl, Anna** (2021) Empirical tests of exchange rate and stock return models
248. **Sjöholm, Carolin** (2021) Essays on public policy in the informal sector context