Universitat Autònoma de Barcelona

Essays in Applied-Microeconomics and Development

Alejandra Ramos

Universitat Autònoma de Barcelona

Doctoral Thesis:

Essays in Applied-Microeconomics and Development

Author: Alejandra Ramos

Advisors: Prof. Giacomo De Giorgi and Prof. Nezih Guner

A thesis submitted to the *Departament d'Economica i d'Història Econòmica* in fulfillment of the requirements for the degree of Doctor of Philosophy at the *International Doctorate in Economic Analysis* (IDEA).

April 21, 2017

A mi mamá, que se nos fue demasiado pronto...

Acknowledgments

I am deeply indebted to Giacomo De Giorgi and Nezih Guner for their invaluable supervision and superb guidance. I will be forever grateful for everything they taught me throughout these years; for their patience, their motivation, and encouragement. Words cannot express my gratitude for their continuous support of my Ph.D.

I would like to express my appreciation to my coauthor Samuel Berlinski. He not only provided me the amazing opportunity to join the Research Department of the Inter-American Development Bank, but also helped me to navigate through pain and grief. I have learned a lot from him, both as a researcher and as a human being.

All of these years would have been unberable without the laughter, talking, travels and support of my friends Ana Maria Manrique, Ana Maria Porras, Camila Fonseca, Camilo Pecha, Laura Ramos, Lavinia Piemontese, and Maria Paula Gandur.

To my colleagues and friends Alberto Ramirez, Alex Filatov, Alex Tagliabracci, Ana Maria Chirinos, Anna Houštecká, Annalisa Loviglio, Andrii Parkhomenko, Augustin Tapsoba, Codruta Boar, Davide Pietrobon, Edgardo Lara, Florence Nimoh, Francesco Cerigioni, Gabor Kocsis, Inês Black, Jose García-Louzao, and Yuliya Kulikova; thank you for our conversations, for the innumerable coffees, and for always providing me with comfort.

I would like to praise special thanks to IDEA. I could not possibly imagine being part of a better team. I am very thankful to Susanna Esteban and Luca Gambetti for their direction of the program; and to Mercè Vicente and Àngels López García for their administrative work. To Ada Ferrer-i-Carbonell, Albert Marcet, Alessandro Ruggieri, Antonio Miralles, Arnau Valladares-Esteban, Benjamín Tello, Carmen Beviá, Caterina Calsamiglia, Christopher Rauh, David Pérez-Castrillo, Dilan Okçuoğlu, Dmitry Matveev, Esther Hauk, Frances Obiols, Hannes Muller, Hugo Rodríguez Mendizábal, Inés Macho-Stadler, Isabel Melguizo López, Javier Fernández-Blanco, Joachim Jungherr, Joan Esteban, Joan Lull, Johannes Gierlinger, Jordi Caballé, Jordi Masso, Juan Carlos Conesa, Laura Mayoral, Lidia Farré, Luis Rojas, Matt Deventhal, Michael Creel, Miguel Angel Ballester, Pau Milán, Pau Olivella, Pau Pujolas, Pedro Rey Biel, Tomás Rodríguez Barraquer, Tongbin Zhang, Xavier Martínez-Giralt; thank you for always having the doors open. To Carmen, thank you for making my office the place to organize my thoughts.

Special thanks to Melissa Hidrobo for providing me the data of *Food, Cash, or Voucher*, and to the International Food Policy Research Institute (IFPRI) and the *World Food Programme* (WFP) Head Quarters and Ecuador's office. I also thank the Ministerio de Educación in Chile and its Centro de Estudios for sharing the data of *Asignación a la Excellencia Pedagógica* with me. I gratefully acknowledge financial support of Universitat Autònoma de Barcelona through the PIF Scholarship.

Nothing would have been possible without my family: they are the ultimate bone of what I do. I cannot stop thanking my father for his permanent interest in my research; my sisters for their loving presence regardless the oceans between us; my beloved friend Silvana Mantilla for being my refuge and bastion even in the darkest nights.

Contents

Introduction

In this doctoral thesis I use applied-microeconomics as a tool for policy re-design in developing countries. With it, I deep in the understanding of why policies work, and how we can use economic incentives to attain better outcomes. My goal is to propose strategies to leverage the resources devoted to social assistance programs, without dealing with the political economy of introducing new items in the development agenda. Specifically, I concentrate in the domains of gender and education. In the domain of gender, I address the issue of intimate partner violence and its interplay with transfers that governments give to women. In the domain of education, I focus on school teachers and a strategy to select and praise the best of them.

In Chapter 1, *Household Decision Making with Violence: Implications for Transfer Programs*, I study how intimate partner violence responds to transfers to women, and whether this response depends on the transfer being in-kind or in-cash. To this end, I develop and estimate a model of household decision making in which the husband and wife maximize a weighted sum of their utilities, and the weights of the agents are endogenously determined through violence. Only the husband can inflict violence to solve spousal disagreement and increase his relative weight. Yet violence comes at the cost of destroying female labor productivity. Under this framework, the utility gains the husband can appropriate through violence are lower when the transfers are in-kind than when the transfers are in-cash. As a result, depending on the level of disagreement, in-kind and cash transfers can have different effects.

I estimate this model using data from a randomized controlled trial of the *World Food Programme* in Ecuador that provides transfers to women of poor families, either in-kind or in-cash. The results indicate that, on average, violence destroys 4 percent of female labor productivity, with a market value of 10 dollars or one day of female labor a month. Violence also reduces female's relative weight in the household utility by 10 percentage points. Therefore, it is *as if* perpetrators were willing to sacrifice one day of female labor every month to reduce their victims' say in household decisions by 20 percent.

If the women beneficiaries of the program receive a cash transfer equivalent to 10 percent of the average household income, the prevalence of violence would reduce from 17 percent to 10 percent. If the same transfer were given in-kind, violence would decline by 3 additional percentage points, from 17 percent to 7 percent. These results also hold for all women receiving Ecuador's main social assistance program, *Bono de Desarrollo Humano*. Furthermore, the differential effect that in-kind transfers have on violence – relative to cash transfers– amplifies as the size of the transfer increases. From a policy perspective these results imply that, with the same resources, the goverment can reduce violence against the women by targeting transfers.

Chapter 2 and Chapter 3 of the thesis are joint work with Samuel Berlinski. In them, we analyze a program that rewards excellence in pedagogical practice in Chile. Similar to the United States, in Chile every year around 12 percent of the teachers transition out of the school system, and 9 percent move to a different school. To prevent good teachers from leaving the profession, in 2002 the Chilean government introduced *Asignación a la Excelencia Pedagógica* or AEP. This is a voluntary certification program designed to reward excellence in teaching practice, both economically and socially. Teachers who succeed on a set of assessments receive a 6 percent annual wage increase for up to 10 years; are invited to become mentors of other teachers in the *Red Maestro de Maestros*; and their names are announced in a ceremony with local authorities and media coverage.

In Chapter 2, *Does Rewarding Pedagogical Excellence Keep Teachers in the Classroom? Evidence from a Voluntary Award Program*, we evaluate the effects of the program on teacher retention and between school mobility. To identify the causal effect of receiving the award, we use a sharp regression discontinuity design. Our estimates indicate that, locally, the award does not alter transitions out of the school system. We observe, however, an increase in intra-system mobility among teachers that receive the award. Some of these mobility patterns are consistent with the award providing a signal of teacher quality.

To understand the nature of the findings we provide a simple model of teachers' quitting behavior. A teacher stays in the profession only if she is paid at least her reservation wage, and more productive teachers have higher reservation wages. To create incentives for more productive teachers to stay in the school system, the government pays a bonus to all those whom voluntarily take the test and score above a threshold. Teachers' decision is whether or not to take the test and, after observing the results, whether or not to quit. The retention effect of the program depends on the difficulty of the test. If the test is relatively easy, teachers around the threshold are paid above their reservation wage and will not quit. In contrast, when the test is difficult, teachers scoring just below the threshold will quit, but those scoring just above will not. Our results suggest that the difficulty of the test is rather low in the sense that teachers marginally failing to receive the award value their jobs more than their outside option. This suggests that, by increasing the difficulty of the exam, rents from low quality teachers can be captured and used to increase the size of the bonus.

In Chapter 3, *The Effects of Public Recognition of Teaching Excellence on Peers Voluntary Certification*, we analyze peers effects in the decision to apply for a certification. Specifically, we explore to what extent being publicly recognized as an excellence teacher affects peers' future application for the AEP certification. Similar to Chapter 2, we identify the effect of being AEP certified using a sharp regression discontinuity design. Our findings suggests that being certified as a excellence teacher doubles peers' application for the program without lowering the quality of the applicants, as measured by the certification rate. The results are particularly relevant for the Chilean context as, in spite of the attractive monetary incentives that AEP offers to awardees, during the first decade of the program less than 6 percent of the eligible teachers apply for it. Moreover, given that in 6 out of every 10 schools no one of the staff has ever applied for AEP, our results suggests that a strategy to increase the application rate is to advertise the program in these schools.

Chapter 1

Household Decision Making with Violence: Implications for Transfer Programs

1.1 Introduction

With the aim of fighting poverty and promoting gender empowerment, many governments around the world have programs that provide transfers to women. These policies implicitly assume that women with more control over economic resources have higher bargaining power, and can achieve better outcomes for themselves and for their children. The empirical evidence suggests that these programs have been successful at improving outcomes such as school attendance, health, and child nutrition.[1] Nevertheless, we know little about the transmission channels leading to these effects. In particular, it is not clear whether the improvement of a woman's utility outside the marriage unambiguously enhances her relative position in the household. On the one hand, women with more economic resources are less trapped in marriage. On the other hand, men can react to threats to their bargaining power by inflicting violence.

In this chapter I study how intimate partner violence responds to transfers to women, and whether this response depends on the transfer being in-kind or in-cash. To this end, I develop a model of household decision making in which the husband and the wife disagree on their preferences, and the husband can use violence to resolve the disagreement in his favor. Violence, however, is costly as it destroys female labor productivity. Under this setting, in-kind transfers with costly resale and cash transfers generate different effects.

I estimate the model structurally using data from a randomized intervention giving transfers to poor families in Ecuador, either in-kind or in-cash. The estimation provides a market value for the productivity cost of violence, quantifies how female's say in the household decisions responds to violence, and quantifies how different transfer regimes affect violence.

To the best of my knowledge, this is the first model of household decision making that acknowledges the fact that the marginal effect of transfers is agent specific, and that violence generates a productivity cost. The structural estimation of the model allows me to quantify men's willingness to pay to improve their relative position in the household, and how different transfers regimes affect violent behavior. This estimation technique

[1]See Duflo (2012) or Doepke and Tertilt (2011) for excellent reviews.

also complements the impact evaluation, and enables me to simulate a scale-up of the program at the national level, without requiring a nation wide randomization. None of these exercises could be done if we relied exclusively on a reduced-form estimation. Once randomized controlled trials are combined with other methods which impose more structure, they can help to better understand what works and why it works.[2]

In my model, a household is formed by a female and a male who derive utility from two public goods. One of the goods is a consumption good that can be bought in the market. The other is a home produced good whose production requires female labor and a market acquired input. We can think about children as the household public good, and food as the market input. Women care relatively more about home produced good. Households maximize a weighted sum of the utility of the female and the male. The couple disagrees on how to allocate income between the market good and the market inputs, and on how to allocate female labor between the market and home production. Only the male can use violence to resolve the disagreement in his favor and increase his relative weight, but this comes at the cost of reducing female labor productivity.

Households differ in the extent of the disagreement and in the female relative wages. The larger the disagreement, the larger the incentives for the male to use violence to align household allocations with his preferences. The effect of female relative wage is, however, ambiguous. On the one hand, when the female has a higher wage, violence sacrifices more female labor productivity. On the other hand, a female with a higher wage can make the household decisions more aligned with her preferences and violence becomes more likely.

This problem is similar to the set-up in Basu (2006), where the household also maximizes a weighted sum of the utilities, the weights are endogenous, and the outcomes are inefficient. In Basu (2006), however, even if current decisions affect the weights in the household utility function in the future, agents do not acknowledge the endogeneity when making decisions today. In my model, I focus on a static household problem where the agents fully internalize how their choices affect their relative weights. Furthermore, while in Basu (2006) the inefficiency comes from the over supply of goods, in my model the inefficiency comes from the destruction of resources.

Under this framework in-kind and cash transfers have different effects. To see why, consider a situation where the female receives a fixed value transfer, either in-kind or in-cash; and the in-kind transfer, in the form of the market input, is not marketable. Suppose that the transfer is infra-marginal for the female, but extra-marginal for the male. That is to say, the in-kind transfer is below what the female would like to spend on the market input if she received the transfer in-cash, but it exceeds what the male would like to spend on the good. These assumptions imply that even if the male uses violence to align the household allocation with his preferences, he can never appropriate the excess in-kind transfer. Therefore, the expected utility gains of violence are unambiguously lower when the transfer is in-kind. With less utilities to appropriate, in-kind transfers make violence less productive as an appropriation device.

Depending on the level of disagreement, any transfer is potentially extra-marginal. Given that the demand functions are defined at the individual level rather than at the household level, it is natural that the marginal effect of transfers (whether infra or extra-marginal) is also agent specific. As a result, as long as the transfer is extra-marginal for at least one of the members of the household, the utility gains that the male can appropriate through violence are lower when the transfer is in-kind than when the transfer is in-cash.

[2]Deaton and Cartwright (2016); Heckman (2010) and Deaton (2010) provide several arguments on how structural estimation techniques can complement randomized controlled trials.

As a result, in-kind and cash transfers can generate different effects.

An additional contribution of the chapter is that it provides a market value for the productivity cost of violence. Although I acknowledge that any form of violence is detrimental for women's well-being, I want to highlight that even abstracting from the obvious and immediate damage to victims, violence is costly for the society. The modeling assumption that enables me to do so is that, instead of violence generating a direct disutility to the female, violence destroys female labor productivity. This assumption is supported by empirical evidence associating intimate partner violence with long-term negative health effects. For instance, according to the WHO Department of Reproductive Health and Research, London School of Hygiene and Tropical Medicine, and South African Medical Research Council (2013), victims of intimate partner violence are at higher risk of future ill-health (including depression and HIV infection), are more likely to have more unwanted pregnancies, have more induced abortions, and are more likely to be long-term users of health services.

I estimate the model structurally using data from a randomized controlled trial of the *World Food Programme* in Ecuador: *Food, Cash, or Voucher.* The program was implemented in 2011 with the aim of reducing poverty and improving food security, and was targeted to women. Beneficiaries received a 40 dollars monthly transfer for 6 months. The transfer could be either in-kind (*food* or *voucher*) or in-cash. This data set is particularly well suited for my empirical exercise as it has a standardized measure of intimate partner violence collected from the same women at two different points in time. Besides, it has is an exogenous variation in the type of transfers.

The results indicate that, on average, violence destroys 4 percent of female productivity with a market value of 10 dollars a month. Violence also reduces female relative weight in the household decision making by 10 percent. Therefore, it is *as if* a perpetrator is willing to sacrifice 10 dollars to reduce his victim's say in the household decisions by 20 percent. Using the estimated distribution of disagreement in preferences and the empirical distribution of female relative wages, I also find that 17 percent of women are victims of intimate partner violence and a program such as *Food, Cash, or Voucher* reduces violence to 8 percent.

Hidrobo, Peterman, and Heise (2016) already performed an impact evaluation of the effect of *Food, Cash, or Voucher* on intimate partner violence. According to their findings, the program reduces physical or sexual violence from 16 percent to 10 percent, and the effects do not differ across treatment-arms. I complement the findings of the authors in three ways: (i) I model a mechanism leading to different effects of in-kind versus cash transfers that can explain heterogeneous effects, (ii) I quantify the cost of violence in productivity terms, and (iii) I quantify how female relative weights respond to violence.

As a policy counterfactual, I simulate how programs providing in-kind or cash transfers, exclusively, affect violence. While a cash transfer equivalent to 10 percent of the average household income would decrease violence by 7 percentage points, the equivalent in-kind transfer would decrease violence by 10 percentage points. To put it differently, if 17 out of every 100 women are victims of intimate partner violence, a cash transfer would guarantee that 7 of these 17 women would no longer be abused. If the exactly same transfers were given in-kind, at least 3 additional women would not be abused.

I take advantage of the model and the estimation results to simulate a scale-up of the program at the national level. This counterfactual analysis resembles Ecuador's anti-poverty program, *Bono de Desarrollo Humano*. Using the data from the *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres*, I simulate the effect of a program giving women of poor families a transfer equivalent to 10 percent of the average household income. The results indicate that a cash transfer program such as *Bono de Desarrollo Humano* reduces the prevalence of violence by 35 percent. If the same transfer were given the in-kind, violence would fall by almost 50 percent. This differential effect amplifies as the size of the transfer increases.

Related Literature

This research question is related to several lines of literature. First, it contributes to the policy debate on whether transfers should be made in-kind or in-cash. For instance, Cunha, De Giorgi, and Jayachandran (2015) explore the price effects of in-kind transfers in the context of Mexico's food assistance program, *Programa de Apoyo Alimentario*. In this very same context, Cunha (2014) tests paternalism and provide evidence on the distortion effect of extra-marginal transfers. Because extra-marginal transfers with costly resale constrain behavior, the paternalistic argument justifies the use of in-kind transfers to enforce outcomes that otherwise will not be attainable. Yet the relevant unit of analysis for the marginal effect of the transfer has been overlooked in the literature.[3] If, as shown by Attanasio and Lechene (2014), the demand functions are defined at the individual level, it is natural that the marginal effect of transfers is also agent-specific. My model of household decision making adds this additional dimension to the debate.

Second, it contributes to the empirical literature on the effect of Conditional Cash Transfers (CCT) on intimate partner violence. In the context of *Oportunidades*, Mexico's flagship poverty alleviation program, Angelucci (2008) finds that, while small transfers decreases violence, large transfers increase violence in households where the husband holds traditional views of gender roles. Also focusing on *Oportunidades*, Bobonis, González-Brenes, and Castro (2013) find that women beneficiaries of the progam are 5 to 7 percentage points less likely to be victims of physical abuse, but are more likely to be victims of emotional violence. In Ecuador, Hidrobo and Fernald (2013) find that the national CCT program, *Bono de Desarrollo Humano*, has no effect on physical violence. Moreover, when the woman is relatively more educated than her partner, the program can increase emotional violence by 9 percentage points. Also in Ecuador, Hidrobo, Peterman, and Heise (2016) evaluate *Food, Cash, or Voucher* – the same program analyzed in this chapter – and find that transfers reduce physical or sexual violence by 6 to 7 percentage points. My results show that *Food, Cash, or Voucher* reduces physical or sexual violence by 9 percentage points; an effect that is in line with the papers that find a reduction in violence due to CCTs.

Third, there is also a theoretical literature on household decision making with violence. In this literature violence is modeled as entering directly in the utility of the abuser, what is called *expressive violence*, or as a means to increase the bargaining power of the perpetrator,what is called *instrumental violence*. As for the victim, the literature models violence as generating direct dis-utility. In my model, violence is instrumental as it is a mechanism that males can use to improve their decision power within the household.[4] I model the cost of violence, however, as a productivity cost as opposed to a direct dis-

[3] see Currie and Gahvari (2008) for a review

[4] Farmer and Tiefenthaler (1997) and Card and Dahl (2011) are among the papers that model violence directly in the utility of the abuser. Anderberg and Rainer (2013), Bloch and Rao (2002) and Bobonis et al. (2013) build models with instrumental violence. When violence is purely expressive, Farmer and Tiefenthaler (1997) show that violence falls as female income increases.

utility. This adoption allows me to give a market value to the cost of violence that I could not attain otherwise.

My model is most related to Eswaran and Malhotra (2011) whom formulate a non-cooperative household model where violence is a means through which the husband enhances his bargaining power; not a psychopathology generating direct utility to the perpetrator.[5] Husband and wife decide on the consumption of public goods, "autonomy" (defined below), and violence. The husband decides how much violence to inflict and, after observing his decision, the wife chooses the bargaining power she would like to confer to herself, i.e. her autonomy. Even if autonomy is decreasing in violence, an increase in the reservation utility of the wife does not involve a monotonic decrease in violence. The reason is that the wife may rationally set a low autonomy to prevent violence, but can also accept more violence just to have more autonomy. In contrast to Eswaran and Malhotra (2011), my model departs from the non-cooperative approach and adopts a setting where, in the absence of violence, the outcomes would be efficient.

I also build upon the large literature on household decision making.[6] Within this literature my set-up is most related to Basu (2006) and Iyigun and Walsh (2007) where the household maximizes a weighted sum of the utilities, the weights are endogenous, and the outcomes can be inefficient. Basu (2006) studies how individual decisions about how much to work and earn today can affect the weights in household utility tomorrow. In Basu (2006), however, even if decisions today affect bargaining power tomorrow, the agents are not fully aware of this endogeneity when taking current decisions. In contrast, in my model the agents fully internalize how their choices affect the relative weights. In Iyigun and Walsh (2007) agents' income also improves their bargaining power, and they internalize this connection. Yet in Iyigun and Walsh (2007) the inefficiency takes the form of an oversupply of goods as agents work more to improve their bargaining power. In my model the inefficiency is captured by the destruction of resources.

Another key feature of the model is that the male faces a trade-off between his weight in household decisions and female labor productivity. On the one hand, since higher female relative wages make violence more costly, one can expect that reductions in female unemployment and in the gender wage gap reduce violence as in Anderberg, Rainer, Wadsworth, and Wilson (2015), Aizer (2010), and Chin (2012). On the other hand, if female labor market participation imposes a significant reduction in the male's relative weight, female employment does not necessarily reduce violence (e.g., Bowlus and Seitz (2006) and Alonso-Borrego and Carrasco (2016)).[7] All in all, the fact that males react to threats to their bargaining power and they do so by destroying female labor productivity is consistent with theories of economic sabotage such as Anderberg, Rainer, Wadsworth, and Wilson (2015). It is also in line with the empirical evidence presented by Bertrand, Kamenica, and Pan (2015), which suggests that husbands may have an aversion for their wives earning more than they do.

I also contribute to the gender empowerment literature. As in Anderson and Genicot

[5]As early as in the seventies, the research in sociology stop understanding domestic violence as rare phenomena confined to mentally disturbed people. From then on, family violence is approached as an extensive phenomenon which could not be explained solely as a consequence of psychological factors (Gelles, 1980).

[6]See Browning, Chiappori, and Weiss (2014) and Chiappori and Mazzocco (2016) for excellent reviews

[7]For Canada, Bowlus and Seitz (2006) suggest that female employment deters abuse only if it occurs before the onset of violence. In Spain, Alonso-Borrego and Carrasco (2016) find that male unemployment increases the risk of violence, while female employment reduces the likelihood of violence only if the male is also employed.

(2014), I provide a mechanism for why pro-woman redistribution policies can increase conflict within the family. Similar to Ashraf (2009), I show that the particular conditions for household decision making have effects on the household allocations. A key message that emerges from this chapter is that, since the type of transfers matters, not all forms of empowerment are equally relevant for all women. Finally, in my model males are willing to sacrifice resources to have more say in the household. This relates to the findings of Almas, Armand, Attanasio, and Carneiro (2015), whom study how a CCT in Macedonia influences women's willingness to pay to have more control of resources within the household.

The rest of the chapter is organized as follows. Section 1.2 presents the model. Section 1.3 describes the data used for the estimation. Section 1.4 presents the reduced-form evidence of the effect of the program on intimate partner violence. Section 1.5 states the assumptions under which the model is identified and how it is estimated. Section 1.6 presents the results. In Section 1.7 I perform some policy analysis. Section 1.8 present the simulations of a scale-up of the program at the national level, and Section 1.9 concludes.

1.2 A Model of Household Decision Making with Violence

In this chapter I develop a static model of household decision making with an explicit role for violence. This section presents a aversion of the model with public goods and no direct disutility of violence. In Appendix A.1 I relax both assumptions. Since they only reinforce the main conclusion, I use the two public goods version of the model to show how violence responds differently depending on the type of transfers. At the end of the section I describe the specific functional forms that will be used to take the model to the data.

General Framework

Consider a group formed by two agents, $s \in \{f, m\}$, who derive utility from a bundle of goods x. Call $u^s(x)$ the utility of agent s. Assume $u^s(x)$ is $\mathcal{C}^2$, strictly increasing, and strictly concave. I consider the interesting case where $u^f(x)$ is different from $u^m(x)$. The group maximizes a weighted sum of the utilities,

$$\mu(v,p)u^f(x) + (1-\mu(v,p))u^m(x),$$

where the weights depend on an endogenous variable v and an exogenous variable p. If I fix p, the weight of the agent f, $\mu(v)$, is $\mathcal{C}^2$, decreasing and strictly concave in v .

Only the agent m can use v, and he use it to shift the couple's decision towards his most preferred x disregarding the utility of agent f. Potentially he could choose a level of v such that $\mu(v) = 0$; however, using v destroys resources. For a given level of v, let $T(v)$ describe the set of feasible bundles. The higher the level of v, the higher the relative weight of m, but the smaller the set of feasible bundles. Therefore, $T : V \rightrightarrows X$ is a correspondence such that, for every $v, v' \in V$, if $v < v'$ then $T(v') \subset T(v)$. We can interpret $T(v)$ as a *resource destruction correspondence.*

The problem of the group is to

$$\max_{x \in T(v), v} \Omega(x,v) = \mu(v)u^f(x) + (1-\mu(v))u^m(x).$$

The strict concavity assumption guarantees that this problem has an interior solution. For a given $\tilde{v}$, the function $\Omega(x, \tilde{v})$ is also strictly concave. Therefore, with $T(\tilde{v})$ convex, there exists an interior maximum of $\Omega(\tilde{v}, x)$ called $x(\tilde{v})$.

$$x(\tilde{v}) = \arg \max_{x \in T(\tilde{v})} \Omega(x, \tilde{v})$$

Because $\Omega(v, x(v))$ is also strictly concave, agents can solve for

$$\max_{v} \Omega(x(v), v),$$

the solution of which is equivalent to:

$$\arg\max_{v} \Omega(x(v), v) \equiv \arg \max_{v \in T(v), x} \Omega(x, v)$$

It is *as if*, after the group achieves the Pareto-efficient allocation $x(v)$, agent m can use v to boost his weight. This outcome is no longer be Pareto-efficient as the agents could agree on same final weights without using v, and no resources would be destroyed.

Household's Problem

Consider now this framework applied to a household formed by a female and a male. Female f and male m derive utility from two public goods. One of the goods is a consumption good bought in the market (q), henceforth *market good.* The other good is a home produced good (Q), henceforth *home good.* The production of the home good requires a market input (d) and female labor ($1 - l_f$, where l_f denotes female labor supply).[8] The female splits her labor between the home production and the market, while the male devotes all his labor to the market. Female and male earn w_f and w_m per efficiency unit of labor supply. I assume that the market good and the market input have the same price, which I normalize to 1.

The government gives a transfer to the female, either in the form of the market input (t_k) or in-cash (t_c). Neither in-kind transfers or the home good are marketable, i.e. they cannot be sold and converted into cash.[9] As the female preserves the transfer even if she leaves the household, the transfer increases her reservation utility. Call $\widetilde{\omega}_f = \frac{w_f + t_k + t_c}{w_m}$ the potential income that the female could have outside the couple, relative to the income of the male. The relative weight of the female in the household decision making is a function of the violence her partner inflicts on her and of her potential income, i.e. $\mu\left(v, \widetilde{\omega}_f\right)$.

Only the male uses violence and he does so to increase his relative weight in the overall household utility. Violence, however, reduces female labor productivity by a fraction $e^{\gamma(v)}$, with $\frac{\partial e^{\gamma(v)}}{\partial v} < 0$ and $e^{\gamma(0)} = 1$. As a result, female labor income reduces to $e^{\gamma(v)} l_f w_f$. Likewise, violence affects the home good produced according to

$$Q = e^{\gamma(v)} F\left(d + t_k, (1 - l_f)\right).$$

I refer to the term $e^{\gamma(v)}$ as the productivity cost of violence.

[8] The market input (d) and the female labor used for home production ($(1 - l_f)$) enter utilities only through Q.

[9] The assumption that in-kind transfer are nor resalable is supported by the data used for the estimation of the model. Alternatively, one can relax the assumption and introduce a transaction cost. Yet this approach would require the estimation of an additional parameter.

The household decides how to allocate female labor between market and home production, how to spend the household income, and how much violence the male can inflict. This optimization problem can be written as

$$max_{q,d,l_f,v}\ \mu(v,\widetilde{\omega}_f)u^f(Q,q)+(1-\mu(v,\widetilde{\omega}_f))\,u^m\,(Q,q)\,, \tag{1.2.1}$$

subject to

$$q+d=e^{\gamma(v)}l_f w_f+w_m+t_c,$$

and

$$Q=e^{\gamma(v)}F\,(d+t_k,(1-l_f))\,.$$

Notice that, while cash transfers enter directly in to the household budget constraint, in-kind transfers enter directly in the home production.

This optimization problem is a short-cut for a marital bargaining where the bargaining power of the spouses depends on their reservation utility and violence. Clearly, not all weights are acceptable for the female. I concentrate in the set of weights that are incentive compatible and make the marriage sustainable. Although it is a strong assumption, it is consistent with the data used for the estimation. First, because in Ecuador 90 percent of the victims of violence are still in a intimate partner relation with the perpetrator (Instituto Nacional de Estadística y Censos, 2011). Second, because in the data only 4.5 percent of the couples dissolved, and among the couple that dissolved only 4.2 percent were violent households.

Female and male disagree on their preferences over the home good and the market good. In particular, the female cares relatively more about the home good. As a result, the male has incentives to use violence to shift the household allocation towards his preferred good, the market good. Violence, however, comes at a cost as it reduces female labor productivity.

To fix the ideas, think about a household deriving utility from healthy children (Q) and a TV (q). Although both husband and wife are enjoy having healthy children and the latest technology TV; the wife cares relatively more about the children. For the children to be healthy, they need food and their mother's attention. Therefore, when the resources are limited, the wife would rather have a simpler TV and spend more money in food. When facing such disagreement, the husband can use violence to reduce the wife's say and get a better technology TV. Violence, however, makes the wife less efficient in taking care of the children, and less productive at work. This in turn reduces the household's income and, overall, there are less resources available for the TV.

The problem of the household is to decide how much money to spend in food (d), how much time the wife spend at home ($1-l_f$), and the level of violence that the husband inflicts (v). These optimality conditions are given by

$$\frac{\partial Q}{\partial d}\left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial Q}+(1-\mu(v,\widetilde{\omega}_f))\,\frac{\partial u^m}{\partial Q}\right]=\left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial q}+(1-\mu(v,\widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q}\right]\frac{\partial q}{\partial d}, \tag{1.2.2}$$

$$\frac{\partial Q}{\partial 1-l_f}\left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial Q}+(1-\mu(v,\widetilde{\omega}_f))\,\frac{\partial u^m}{\partial Q}\right]=\left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial q}+(1-\mu(v,\widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q}\right]\frac{\partial q}{\partial(1-l_f)}, \tag{1.2.3}$$

and

$$\frac{\partial \mu(v,\widetilde{\omega}_f)}{\partial v}\Delta u_f^m = \left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial Q} + (1-\mu(v,\widetilde{\omega}_f))\frac{\partial u^m}{\partial Q}\right]\frac{\partial Q}{\partial v} + \left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial q} + (1-\mu(v,\widetilde{\omega}_f))\frac{\partial u^m}{\partial q}\right]\frac{\partial q}{\partial v}, \tag{1.2.4}$$

where Δu_f^m is the level difference between the utility of the male and the utility of the female, i.e. $\Delta u_f^m = u^m(Q,q) - u^f(Q,q)$.

Equations (1.2.2) and (1.2.3) require that the marginal benefit from devoting resources to the inputs of home production equals the opportunity cost of not spending the resource in the market good. This condition can also be interpreted in terms of the household's willingness to pay for the market good and the home good. Rearranging the terms of equations (1.2.2) and (1.2.3), we have that for any input of production $z \in \{d, 1-l_f\}$, at the optimum,

$$\frac{\frac{\partial Q}{\partial z}}{\frac{\partial q}{\partial z}} = \frac{\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial q} + (1-\mu(v,\widetilde{\omega}_f))\frac{\partial u^m}{\partial q}}{\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial Q} + (1-\mu(v,\widetilde{\omega}_f))\frac{\partial u^m}{\partial Q}}. \tag{1.2.5}$$

The LHS of equation (1.2.5) is the ratio between the marginal productivity of the input z and its marginal cost. The RHS of equation (1.2.5) is the ratio between the household's marginal willingness to pay for the market good and its marginal willingness to pay for home good. Because the RHS does not depend on z, the production of the home good is independent of preferences and weights,

$$\frac{\frac{\partial Q}{\partial d}}{\frac{\partial q}{\partial d}} = \frac{\frac{\partial Q}{\partial (1-l_f)}}{\frac{\partial q}{\partial (1-l_f)}}. \tag{1.2.6}$$

That is to say, the marginal rate of technical substitution between d and $(1-l_f)$ equals the relative marginal cost. This optimality condition holds as long as the inputs of home production do not generate direct utility to the agents.

Similarly, the optimality condition described in equation (1.2.4) requires that the marginal benefit of violence equals its marginal cost. The marginal cost of violence is defined in terms of the resources that are destroyed instead of being allocated to the consumption of Q and q. The first term of the RHS of equation (1.2.4), is the marginal cost of violence in terms of the forgone utility from the home good. The second term of the RHS of equation (1.2.4), is the marginal cost of violence in terms of the forgone utility from the market good. The marginal benefits of violence are the utility gains that the male extract from the wife by inflicting one additional unit of violence (LHS of equation (1.2.4)). This term depends on two elements: the additional weight the husband can confer himself by inflicting violence $\left(\frac{\partial \mu(v,\widetilde{\omega}_f)}{\partial v}\right)$, and the utilities at stake (Δu_f^m).

We can also interpret the optimality condition for violence in terms of the sensitivity of the productivity cost of violence and the female relative weight with respect to violence. Let $\varepsilon_v^Q = \frac{\partial Q}{\partial v}\frac{v}{Q}$ and $\varepsilon_v^q = \frac{\partial q}{\partial v}\frac{v}{q}$, be the elasticities of the home good and the market good with respect to violence. Using the home good technology production and the budget constraint it can be seen that the elasticity of Q and q with respect to v depend on the elasticity of the productivity cost of violence, i.e. $\varepsilon_v^Q = \gamma(v)\varepsilon_v^\gamma$ and $\varepsilon_v^q = \gamma(v)\frac{e^{\gamma(v)}l_f w_f}{q}\varepsilon_v^\gamma$, where $\varepsilon_v^\gamma = \frac{\partial \gamma(v)}{\partial v}\frac{v}{\gamma(v)}$. In a similar fashion, I express $\frac{\partial \mu(v,\widetilde{\omega}_f)}{\partial v}$ as $\frac{\mu(v,\widetilde{\omega}_f)}{v}\varepsilon_v^\mu$. Then, equation

(1.2.4) reads as

$$\mu(v,\widetilde{\omega}_f)\varepsilon_v^{\mu}\Delta u_f^m = \left\{ \left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial Q} + (1-\mu(v,\widetilde{\omega}_f))\,\frac{\partial u^m}{\partial Q}\right] Q \right. \tag{1.2.7}$$
$$\left. + \left[\mu(v,\widetilde{\omega}_f)\frac{\partial u^f}{\partial q} + (1-\mu(v,\widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q}\right] e^{\gamma(v)} l_f w_f \right\} \gamma(v)\varepsilon_v^{\gamma}.$$

The higher the elasticity of female relative weight, the higher the benefit of violence. Likewise, the higher the elasticity of the productivity cost of violence, the higher the cost of inflicting it.

Equation (1.2.7) also illustrates why different type of transfers can generate different effects on violence. Fix the marginal cost of violence (RHS of equation (1.2.7)). If the utility at stake (Δu_f^m) increases, for equality to hold it must be the case that $\frac{\partial \mu(v,\widetilde{\omega}_f)}{\partial v}$ decreases. Since $\mu(v,\widetilde{\omega}_f)$ is concave in v, then v must increase. In other words, with more utility gains to appropriate, the husband can allow himself to inflict more violence. As a result, if transfers generate different potential gains of violence, the level of violence will also differ.

1.2.1 In-kind versus Cash Transfers

Whenever the spouses disagree on their preferences and they bargain over how to allocate resources, the relevant demand systems should be defined at the individual rather than at the household level. A natural extension of this argument is that, whether a transfer is extra-marginal or infra-marginal should also be defined at the individual level.[10] Depending on the disagreement, any in-kind transfer has the potential of being extra-marginal for the male. If this is the case, he has an incentive to use violence to align household allocations with his preferences. As a result, in-kind and cash transfers will affect differently the level of utility gains that the male can extract from the female.

When the transfer is in-kind, even if the male would like to use violence to lower the demand of the market input, there is a minimal level of home good that must be produced. As a result, the utility gains that the male can extract through violence are less than what he could appropriate under a cash transfer. Figure 1.1 presents the argument graphically.

For a spouse $s \in \{f, m\}$, define the maximum utility that the he/she would attain if he/she were the sole decision maker of the household, i.e.

$$\overline{u}^s \equiv \max_{q,l_f,d} u^s(Q,q)$$

subject to

$$q + d = l_f w_f + w_m, \quad \text{and} \quad Q = F\left(d, 1-l_f\right).$$

The optimization problem described in equation (1.2.1) represents how the household bargain over these maximum utilities, taking into account that the husband can inflict violence and agents' reservation utilities. The black solid line depicts the utility possibility frontier of the household in the absence of violence and transfers.

Consider, now, a transfer to the female, either in-kind or in-cash. Fix the level of the transfer so that in-kind and cash have the same market value. Assume that the transfer

[10]A transfer is infra-marginal (extra-marginal) if it is below (above) what would be consumed of the same good if the transfer were given in cash.

is infra-marginal for the female and extra-marginal for the male. That is to say, if the female had the entire household income available and she received the transfer in cash, she would spend in d more than what the government transfers in-kind. In contrast, upon receiving the transfer in cash, the male would like to spend less than the equivalent in-kind transfer. The red dotted line and the blue dashed line in Figure 1.1 depict the household utility possibility frontier under an in-kind transfer and under a cash transfer. Regardless of the type of transfer, the household's utility possibility frontier expands; yet the expansion is not symmetric between the different transfer regimes.

Since the transfer is infra-marginal for the female, she is indifferent between receiving it in-kind or in-cash. This means that

$$\overline{u}_k^f \equiv \max_{q,l_f,d} u^f(Q,q),$$

subject to

$$q + d = l_f w_f + w_m =, \quad \text{and} \quad Q = F\left(d + t_k, 1 - l_f\right);$$

is equivalent to

$$\overline{u}_c^f \equiv \max_{q,l_f,d} u^f(Q,q),$$

subject to

$$q + d = l_f w_f + w_m + t_c, \quad \text{and} \quad Q = F\left(d, 1 - l_f\right).$$

I call this utility $\overline{u}_{k,c}^f$.[11]

As for the male, if he were the sole decision maker and if the female received the transfer in-cash, he would devote less resources to the market input d than what the government transfers in-kind. Call the associate utility levels

$$\overline{u}_k^m \equiv \max_{q,l_f,d} u^m(Q,q),$$

subject to

$$q + d = l_f w_f + w_m, \quad \text{and} \quad Q = F\left(d + t_k, 1 - l_f\right).$$

under an in-kind transfer, and

$$\overline{u}_c^m \equiv \max_{q,l_f,d} u^m(Q,q),$$

subject to

$$q + d = l_f w_f + w_m + t_c, \quad \text{and} \quad Q = F\left(d, 1 - l_f\right);$$

under cash. The above argument states that $\overline{u}_k^m < \overline{u}_c^m$.

The area between the dotted line and the solid line represents the potential utility gains that the male could appropriate through violence, under an in-kind transfer. The area between the dashed line and the solid line represents the analogous utility gains under a cash transfer. Clearly, when the transfer is in-kind there are less utilities that the male can extract. This in turn makes violence less productive as an appropriation device.

[11]Since we are in a two goods case, in-kind and cash transfers generate the same welfare for a female who is the sole household decision maker. With more goods, in-kind transfers are dominated (in welfare terms) by cash transfers. This would only reinforce the argument as the potential gains of violence under cash transfer will increase even more.

1.2.2 Taking the Model to the Data

To bring the model to the data, I assume that the home good is produced according to a Cobb-Douglas technology with parameter θ,

$$Q = e^{\gamma(v)} \left(d + t_k\right)^{\theta} \left(1 - l_f\right)^{1-\theta}.$$

Then, equation (1.2.6) reads as

$$\frac{1-\theta}{\theta}\frac{d+t_k}{(1-l_f)} \quad = \quad w_f e^{\gamma(v)}, \tag{1.2.8}$$

where $\frac{\partial q}{\partial d} = 1$ by normalization, and $\frac{\partial q}{\partial 1 - l_f} = w_f e^{\gamma(v)}$.

I assume that the utilities of both female and male are logarithmic, i.e.,

$$u^f(Q,q) = \alpha_i^f \log(Q) + \log(q) \quad \text{and} \quad u^m(Q,q) = \log(Q) + \log(q).$$

For females care relatively more about the home good, it must be that $\alpha_i^f > 1$ for any household i. These functional forms reduce all the extent of the disagreement between female and male to α_i^f.

I can now rewrite the optimality condition for violence. The difference between the utility of the male and the female that appears in the LHS of equation (1.2.7) becomes

$$\Delta u_f^m = \left(1 - \alpha_i^f\right) \log(Q).$$

As well, the first term inside the brackets on the RHS of equation (1.2.7), simplifies to

$$\left[\mu(v,\widetilde{\omega}_f)\frac{\alpha_i^f}{Q} + (1 - \mu(v,\widetilde{\omega}_f))\,\frac{1}{Q}\right] Q = \left[\mu(v,\widetilde{\omega}_f)\alpha_i^f + (1 - \mu(v,\widetilde{\omega}_f))\right].$$

Since both female and male have the same valuation for the market good, the second term inside the brackets on the RHS of equation (1.2.7) reduces to

$$\frac{\partial u^{f,m}}{\partial q} e^{\gamma(v)} l_f w_f = \frac{e^{\gamma(v)} l_f w_f}{q} = \rho,$$

where ρ is the ratio between female's labor income and the value of the market good. Putting the pieces together, equation (1.2.7) reads as

$$\mu(v,\widetilde{\omega}_f)\varepsilon_v^{\mu} \left(\alpha_i^f - 1\right)(-\log(Q)) \quad = \quad \left[(1+\rho) + \mu(v,\widetilde{\omega}_f)\left(\alpha_i^f - 1\right)\right]\gamma(v)\varepsilon_v^{\gamma}. \tag{1.2.9}$$

As both $\mu(v,\widetilde{\omega}_f)$ and $\gamma(v)$ are decreasing in v, and $\alpha_i^f > 1$, equation (1.2.9) only holds for $Q < 1$. This means that violent households are characterized by a relatively low production of household public good.

1.3 Data

Food, Cash, or Voucher Program

I estimate the model using data from a randomized intervention of the *World Food Programme* in Ecuador named *Food, Cash, or Voucher*. This is a randomized controlled trial implemented during 2011 with the aim of improving food consumption, increasing the role of women in food consumption decisions, and reducing tensions between Colombian refugees and Ecuadorians. Beneficiaries of the program received a monthly transfer of 40 dollars during 6 months. This transfer was equivalent to 10 percent of the average household monthly income. To receive the transfer, participants were required to attend a nutritional training program.

As the name of the program suggests, the transfer could take one of three forms: cash, food, or voucher. Cash transfers were distributed through ATMs. Food transfers consisted of a food basket containing rice (24 kg), lentils (8kg), vegetable oil (4 l) and can-sardines (8 cans).[12] The voucher transfer was a voucher on the name of the female, redeemable at local supermarkets from a list of pre-approved goods.[13] As the list of pre-approved goods maps to the food basket, both *voucher* and *food* constitute an in-kind transfer.

Food, Cash, or Voucher was implemented in seven urban centers of the provinces of *Carchi* and *Sucumbíos*. These urban centers had more than 10 percent of Colombian refugees, more than 50 percent of people living in poverty, a local provider to implement food distribution, and financial institutions to distribute cash via ATMs. The urban centers were divided into 84 neighbors, which were randomly assigned to treatment (61) and control (19) groups. Neighbors were divided into geographical units named clusters. Within the treated arm, 110 clusters were randomly assigned to food, cash, or voucher. The sample consists of 2357 households, 652 of which were assigned to the control group. From the 1,705 treated households, one third received the food transfers, one third the cash transfer, and one third received the voucher. All the households with at least one Colombian member, and all households considered as poor were enrolled in the program. Households receiving *Bono de Desarrollo Humano*, Ecuador's main poverty alleviation program, were excluded.

In March-April 2011, the 2357 households were interviewed. Out of them, 2122 were re-surveyed in October-November 2011. The final sample used for the reduced form estimation consist of 1,230 households where the female respondent of the intimate partner violence questionnaire was the head of household or the spouse of the head, she was between 15 and 70-years-old, married or at union at the beginning of the intervention, and alone at the time of both interviews (see Table A.1). I also limit the sample to the households where the domestic violence questionnaire was answered by the same woman at baseline and the follow-up. For the structural estimation I further restrict the sample to the set of households with complete information on violence, labor income, food expenditure, and time allocation, both at baseline and follow-up.

Definitions of Violence

For the reduced form estimation, I define physical violence as a dummy taking the value of 1 if the female reported being pushed, slapped, punched, kicked, strangled, or threatened or attacked with a weapon by her partner in the last 6 months.[14] Emotional violence is

[12] Hidrobo, Hoddinott, Peterman, Margolies, and Moreira (2014) report that, at baseline, the median household consumed 20 kgs of cereals, 0.35kg of fish and seafood, and 2 kg of pulses and legumes. Nevertheless, 40 dollars is less than the average household food consumption at baseline.

[13] Vouchers were serialized and printed centrally, and were non-transferable.

[14] In Ellsberg and Heise (2005), the World Health Organization define intimate partner violence as any act or omission by a current or former intimate partner which negatively affects the well-being, physical or psychological integrity, freedom, or right to full development of a woman. Physical violence includes

coded as 1 if the female reported being threatened with abandonment, threatened with being taken away from her children, threatened with being hurt, humiliated, or ignored by her partner in the last 6 months. Sexual violence is coded as 1 if during the last 6 months the intimate partner forced the female to have sex or to commit sexual acts she did not approve. A woman suffering from physical, emotional or sexual violence with her partner as the perpetrator was considered as a victim of intimate partner violence (*any violence*). To ensure comparability with most of the surveys on *Violence Against the Women*, I focus on physical or sexual violence.

For the structural estimation, I built an index of physical or sexual violence ranging between 0 and 1. This index captures the different forms of violence that the female experienced in the last 6 months by hands of her partner: pushed, slapped, punched, kicked, strangled, threatened with a weapon, attacked with a weapon, forced to perform sexual acts that she did not approved, and forced to have sex. For instance, a female who reported being punched and threatened with a weapon, but who did not suffer any of the other assaults listed, has an index of $\frac{2}{9} = 0.225$.

Descriptive Statistics

Table 1.1 presents some descriptive statistics of the sample. I begin by presenting basic household demographics (Panel A, Table 1.1). The average household is a 5 members family, with one child under 5 years and another under 14. The head of the household is a 39-years-old working man and his spouse is a 35-years-old woman. Around 40 percent of these couples are married. The remaining 60 percent live in cohabitation. Both female and male have on average 8 years of education. Yet 18 percent the women are more educated than their partners. Columns (5) and (6) of Table 1.1 present the p-value of a battery of randomization tests. Except the number of household's members and the number of children aged 6 to 15, none of the variables is statistically different between the treatment and the control group. Comparing in-kind and cash, apart from the percentage of households that have a male as the head of household, none of the variables is statistically different across treatment-arms at the 5 percent significance level.[15]

Panel B of Table 1.1 presents the prevalence rates of intimate partner violence. At baseline, 1 out of every 3 women experienced violence in the last 6 months. During the same period, 15 percent of the females have been physically abused by her partners. This means that in almost 60 percent of the cases of intimate partner violence, the abuse involved the intentional use of physical force with the potential of causing death, injury, or harm. Reports of sexual violence are scarce.

Panel C of Table 1.1 presents the main variables used for the estimation. Among the women who where physically or sexually abused by an intimate partner, the average index of violence is 0.25. That is to say, the average abused woman experienced about two out of the nine possible forms of physical or sexual violence surveyed. The average household

scratching, pushing, shoving, throwing, grabbing, biting, choking, shaking, poking, hair pulling, slapping, punching, hitting, burning, the use of restraints or one's body size or strength against another person, and the use, or threat to use, a weapon (gun, knife, or object). Emotional violence is any act or omission that damages the self-esteem, identity, or development of the individual; including humiliation, threatened loss of custody of children, forced isolation from family or friends, threatened to harm the individual or someone they care about, repeated yelling or degradation, inducing fear through intimidating words or gestures, controlling behavior, and the destruction of possessions.

[15]Refer to Hidrobo et al. (2014) and Hidrobo et al. (2016) for an in-dept validation of the randomization and the attrition analysis.

daily income amounts to 14 dollars. From those 14 dollars, 4 dollars are devoted to daily food expenses. For these families, a 40 dollars transfer is equivalent to 10 percent of the average household monthly income and is infra-marginal for the average household. Around 30 percent of women work in the labor market for an average of 5 hours a day. In addition to these 5 hours, the average woman devotes 7 hours to household chores. Almost all of the males in the sample work in the market, for an average of 7 hours a day. While the average male earns 12.40 dollars a day, the average female earns 6.55 dollars a day.

Female Wages

The data on female labor income captures $e^{\gamma(v)} l_f w_f$. For the estimation of the model, however, I need to disentangle w_f, $e^{\gamma(v)}$, and l_f. Since the model does not include leisure, for every household, I express the female time allocation relative to the total time she devotes to household work and to the labor market. This means that, for the estimation, a female that works 8 hours a day in the market and 4 hours at home is equivalent to a female that works 4 hours a day, devotes 2 hours to household chores, and use the rest of her time for leisure. Without this normalization, it would not be possible to distinguish between the forgone female labor income associated to violence and the forgone female labor income associated to leisure.

I interpret the female per hour wage rate from the data as $e^{\gamma(v)} w_f$. In the absence of violence, w_f is directly observed. I proceed as if the wages of working women who are abused were non observable, and use female per hour wage rate in non-violent households to predict their wages. I implement this strategy using a Heckman Two-Step procedure among the female-working households.[16] As exclusion restrictions, I use the number of children and the cohabitation status of the couple. Both in the literature and in the data, these two variables are identified as risk factors for violence.[17] As explanatory variables of the non-violent component of wages, I use demographics and the numbers of hours worked. Table A.2 presents the results of this approach in the original sample.

I use the Heckman procedure just described to predict female relative wages, $\widehat{w}_f$. The female wage variable used for the estimation is

$$w_f = \begin{cases} w_f & \text{if} \quad v = 0 \quad \text{and} \quad l_f > 0 \\ \widehat{w}_f & \text{if} \quad v = 1 \quad \text{and} \quad l_f > 0 \end{cases} \tag{1.3.1}$$

Figure 1.2 presents the predicted distribution of female relative per hour wage rate, in logarithms.

1.4 Reduced-Form Evidence

Hidrobo, Peterman, and Heise (2016) evaluate the effect of *Food, Cash or Voucher* on intimate partner violence. Their reduced-form estimates suggest that the overall treatment reduces violence by 6 to 7 percentage points. Nevertheless, the impacts do not vary across

[16] This is not a classical employment equation. Here the selection occurs on violence. I refrain from doing a selection labor market participation as the optimality conditions derived only hold for interior solutions.

[17] In the data, 20 percent of the women with children and 18 percent of the women living in cohabitation are abused. These rate decreases to 13 percent for women without children, and to 14 percent for married women.

the three modalities. This section presents the reduced-form evidence of the effect of the program combining the *voucher* and *food* treatment-arms in a unique in-kind treatment. Similar to Hidrobo et al. (2016), I estimate the following linear probability model:

$$v_{ij1} \;=\; c + \beta_{\text{cash}} T_i^{\text{cash}} + \beta_{\text{in-kind}} T_i^{\text{in-kind}} + v_{ij0} + \phi_j + e_{ij} \tag{1.4.1}$$

The violence v that the woman of household i located in province j faces at the end of the intervention is a function of the type of transfer she receives, the violence she was facing at baseline, the province where she lives, and an error term.[18] The coefficients of interest are β_{cash} and $\beta_{\text{in-kind}}$.[19]

Table 1.2 presents the estimates of equation (1.4.1). Columns (1) to (4) present the estimates for the entire sample. Columns (1) and (2) present the effect of the pooled treatment without and with controls. Columns (3) and (4) compare the in-kind and cash treatment arms without and with controls. At baseline, 16 percent of the women have been victims of physical or sexual violence in the last 6 months. Similar to Hidrobo et al. (2016), I find that *Food, Cash or Voucher* reduces physical or sexual violence by 6 percentage points and the effects is no different across treatment-arms.

In columns (5) to (8) of Table 1.2, I repeat the exercise for the sub-sample of households where the female was working at baseline. At the beginning of the intervention, 17 percent of these women were victims violence. Among them, *Food, Cash or Voucher* reduces violence by around 10 percentage points. The magnitude of the estimates of column (7) and (8) suggests the effect is concentrated among the set of households that received the in-kind transfer; yet the difference is only statistically significant at the 24 percent level and 30 percent level.

This evidence suggests that there is a substantial amount of heterogeneity in the effects of in-kind and cash transfers. Since the reduced- form evidence provides limited insides about the source of heterogeneity, I rely on the model presented in Section 1.2 to understand what are the underlying mechanism leading to the different outcomes of the transfers and which are the households that would benefit more from receiving the transfer in-kind rather than in-cash.

1.5 Identification and Estimation

In the data from *Food, Cash, or Voucher* I observe violence (v), the household food expenditure (d), female time allocation to home production and to the labor market (l_f and $1 - l_f$), male labor income (w_m), female labor income ($e^{\gamma(v)} l_f w_f$), the ratio between female labor income and the value of the market acquired good (ρ), the cash transfer (t_c), and the in-kind transfer (t_k). The remaining remaining parameters to be identified are θ, $\gamma(v)$,α_i^f and $\mu(v, \widetilde{\omega}_f)$. This section, first, I impose additional restrictions that will allow me to identify these parameters. Then I explain how I estimate the model.

[18] A dummy for province is included in the regression as it was used for the stratification of the sample.

[19] Instead of doing a conventional difference-in-difference approach where the left-hand side variable would have also been time indexed, here I avoid serial correlation by controlling the violence at baseline. This approach is preferred over the conventional difference-in-difference approach when the correlation of outcomes is low (McKenzie, 2012).

Identification

At the optimal level of home production, the relative marginal productivity of the female time and the market input equals its relative price (see equation (1.2.6)). Using the Cobb-Douglas functional form and taking logarithms, a rearranged version of equation (1.2.8) reads as

$$\log\left(\frac{d+t_k}{(1-l_f)\,w_f}\right)=\log\left(\frac{\theta}{1-\theta}\right)+\gamma(v).$$

If there is a random error in the measurement of the inputs of household production, I can estimate

$$\log\left(\frac{d_{i\tau}+t_{k,\tau}}{(1-l_{f,i\tau})\,w_{f,i\tau}}\right)=\log\left(\frac{\theta}{1-\theta}\right)+\gamma(v_{i\tau})+\epsilon_{i\tau}, \tag{1.5.1}$$

to recover θ and $\gamma(v)$. The sub-index i denotes a household, and the sub-index τ denotes if the observation corresponds to baseline ($\tau=0$) or to follow-up ($\tau=1$). Since the transfers were only active at follow-up, $t_{c,0}=t_{k,0}=0$.

The parameter θ is identified as a transformation of the constant term of equation (1.5.1). The functional form $\gamma(v)$ can be directly identified from the relation of the relative demand for the inputs of production and violence. The underlying assumption for identification is that the measurement error is uncorrelated with violence. If this is the case, θ and $\gamma(v)$ enable me to recover the home good (Q) using observations of the household food expenditure (d) and female labor devoted to home production ($1-l_f$).

To identify the effect of violence on the female weights, $\mu(v,\widetilde{\omega}_f)$, I begin by applying a logarithmic transformation to equation (1.2.9)

$$\log\left(1+\rho\right)=\log\left(\alpha_i^f-1\right)+\log\left(\mu(v,\widetilde{\omega}_f)\right)+\log\left(-\log(Q)\frac{\varepsilon_v^{\mu}}{\varepsilon_v^{\gamma}\gamma(v)}-1\right).$$

Even if classical measurement error was assumed, the estimation of

$$\log\left(1+\rho_{i\tau}\right)=\log\left(\alpha_i^f-1\right)+\log\left(\mu(v_{i\tau},\widetilde{\omega}_{f,i\tau})\right)+\log\left(-\log(Q_{i\tau})\frac{\varepsilon_{v_{i\tau}}^{\mu}}{\varepsilon_{v_{i\tau}}^{\gamma}\gamma(v_{i\tau})}-1\right)+\epsilon_{i\tau},$$

would not allow me to identify $\mu(v,\widetilde{\omega}_f)$ separately from ε_v^{μ}. To circumvent this issue, I assume that $\frac{\varepsilon_v^{\mu}}{\varepsilon_v^{\gamma}\gamma(v)}=\delta$ (with $\delta>2$) and do a linear approximation of $\log(-\log(Q)\delta-1)$ around δ, to arrive at

$$\log\left(1+\rho\right)+\log(Q)+1\simeq\log\left(\alpha_i^f-1\right)+\log\left(\mu(v,\widetilde{\omega}_f)\right)+\log(\delta-1).$$ [20]

The only functional form of $\mu(v,\widetilde{\omega}_f)$ that satisfies this assumption is

$$\mu(v,\widetilde{\omega}_f)=\exp(\delta\gamma(v)+\kappa(\widetilde{\omega}_f))=\left[\exp(\gamma(v))\right]^{\delta}\exp(\kappa(\widetilde{\omega}_f)).$$

Therefore, the response of the female relative weight to violence is a power function of the productivity cost of violence. Moreover, when $v=0$, $\mu(0)=\exp(\kappa(\widetilde{\omega}_f))$, so that $\exp(\kappa(\widetilde{\omega}_f))$ can be interpreted as the female weight in the absence of violence.[21] With

[20] The linearization around δ is as if $Q\simeq\exp\left(-\frac{\delta+1}{\delta}\right)$. Since $Q\in\left(e^{-\frac{3}{2}},e^{-1}\right)$, it is as if I approximated around the level of home production of a highly violent household.

[21] The concavity of $\mu(v,\widetilde{\omega}_f)$ is guaranteed as long as the productivity cost of violence is also concave, i.e. $\frac{\partial^2\gamma(v)}{\partial^2 v}<0$.

this additional assumption, the optimality condition of violence described in equation (1.2.4) reduces to

$$\log(1+\rho) + \log(Q) + 1 \simeq \log\left(\alpha_i^f - 1\right) + \delta\gamma(v) + \kappa(\widetilde{\omega}_f) + \log(\delta - 1),$$

where δ is a new parameter to be identified.

In the presence of classical measurement error, δ and $\mu(v, \widetilde{\omega}_f)$ can be identified using the following equation,

$$\log(1+\rho_{i\tau}) + (\log(Q_{i\tau}) + 1) \simeq \log\left(\alpha_i^f - 1\right) + \delta\gamma(v_{i\tau}) + \kappa(\widetilde{\omega}_f) + \log(\delta - 1) + \epsilon_{i\tau}. \quad (1.5.2)$$

Given that there are two observations per household, the disagreement in preference parameter, α_i^f, is identified through a household fixed effect. Once α_i^f and $\gamma(v)$ are identified, δ is non-linearly identified. I recover $\mu(v, \widetilde{\omega}_f) = \exp(\delta\gamma(v) + \kappa(\widetilde{\omega}_f))$ using δ, $\gamma(v)$ and $\kappa(\widetilde{\omega}_f)$. Although the transfer affects both v and $\widetilde{\omega}_f$ which in turn enter in $\mu(v, \widetilde{\omega}_f)$, *Food, Cash, and Voucher* gives me an exogenous variation in the gains of violence that allows me to separate the effect of violence on weights from the effect of higher relative income.

Estimation

I implement the above identification strategy in three blocks. The first block uses the data on household food expenditure ($d_{i\tau}$), household labor supply ($l_{f,i\tau}$), and female relative wages ($\frac{w_{f,i\tau}}{w_{m,i\tau}}$) to estimate the technology of the home production ($\widehat{\theta}$) and the productivity cost of violence ($\widehat{\gamma}(v_{i\tau})$) through an OLS regression. The second block uses data of $\rho_{i\tau}$ (the ratio between the female labor income and the market acquired good) and the first block estimates to construct the LHS variable of equation (1.5.2), and estimate a household fixed effect regression to recover $\widehat{\alpha}_i^f$. The third block uses the residuals of the second block to estimate a non-linear least squares regression on violence ($v_{i\tau}$) to recover $\widehat{\delta}$ and compute $\widehat{\mu}(v_{i\tau}, \widetilde{\omega}_{f,i\tau})$.

I begin by estimating equation (1.5.1) through an OLS regression of the relative demand for the inputs of home production, on a constant and a series of polynomials of violence.

$$\log\left(\frac{d_{i\tau} + t_{k,\tau}}{(1 - l_{f,i\tau})\,\widetilde{\omega}_{f,i\tau}}\right) = \beta_0 + \beta_1 v_{i\tau}^1 + \beta_2 v_{i\tau}^2 + \beta_3 v_{i\tau}^3 + \ldots + \epsilon_{i\tau}. \quad (1.5.3)$$

As this equation comes from the interior solutions described by equations (1.2.2) and (1.2.3), I only estimate it in the set of households where the female works in the labor market. The technology parameter θ is identified from the following transformation of the constant β_0 of equation (1.5.3),

$$\widehat{\theta} = \frac{\exp(\widehat{\beta}_0)}{1 + \exp(\widehat{\beta}_0)}.$$

Once $\widehat{\theta}$ is recovered, I can back-up $\widehat{\gamma}(v_{i\tau})$ by subtracting the constant $\widehat{\beta}_0$ from the linear prediction of equation (1.5.3). This difference is $\widehat{\gamma}(v_{i\tau})$. By construction, when there is no violence, $\widehat{\gamma}(v_{i\tau}) = 0$, so that $e^{\widehat{\gamma}(v_{i\tau})} = 1$.

Based on the estimated values of $\widehat{\theta}$ and $\widehat{\gamma}(v_{i\tau})$, and the data on $d_{i\tau}$ and $1 - l_{f,i\tau}$, I compute the predicted home good, $\widehat{Q}_{i\tau}$. Now, I proceed to estimate the following regression based on equation (1.5.2):

$$\log(1+\rho_{i\tau}) + \left(\log(\widehat{Q}_{i\tau}) + 1\right) \simeq \log\left(\alpha_i^f - 1\right) + \delta\widehat{\gamma}(v_{i\tau}) + \kappa(\widetilde{\omega}_f) + \log(\delta - 1) + \epsilon_{i\tau}. \quad (1.5.4)$$

As this equation comes from the interior solution described by equation (1.2.4), I only estimate it in the set of households where there is violence.

Given the non-linearity of δ, I estimate equation (1.5.4) in two-stages. In the first-stage, I use the panel structure of the data and run an OLS regression of the outcome variable against a household fixed effect,

$$\log(1+\rho_{i\tau}) + \left(\log(\widehat{Q}_{i\tau}) + 1\right) = a_i + \eta_{i\tau}. \quad (1.5.5)$$

The term $\eta_{i\tau}$ contains all of the arguments in the RHS the regression 1.5.4, except $\log(\alpha_i^f - 1)$. Under the classical measurement error assumption, $\eta_{i\tau}$ is uncorrelated with α_i^f, and this first-stage estimation is unbiased. The parameter α_i^f is recovered from

$$\alpha_i^f = \exp(a_i) + 1.$$

Next, I define $\kappa(\widetilde{\omega}_f)$ as the maximum weight a female with relative potential earnings $\widetilde{\omega}_f$ can attain under no violence, $\kappa(\widetilde{\omega}_f) = \exp(\mu_{max}(w_f))$. I set $\mu_{\max} = \exp(\kappa) = 0.5$ for any $\widetilde{\omega}_f$, and use the difference between the residuals $\widehat{\eta}_{i\tau}$ and κ as the outcome variable of the following regression on violence:

$$\hat{\eta}_{i\tau} - \kappa = \delta\widehat{\gamma}(v_{i\tau}) + \log(\delta - 1) + \epsilon_{i\tau}. \quad (1.5.6)$$

Equation (1.5.6) is the second-step of equation (1.5.4), and is estimated through non-linear least squares. With it, I recover $\widehat{\delta}$. The weight function $\mu(v, \widetilde{\omega}_f)$ is recovered from a transformation of the constant κ, the estimated parameter $\widehat{\delta}$, and the relation $\widehat{\gamma}(v_{i\tau})$:

$$\widehat{\mu}(v_{i\tau}, \widetilde{\omega}_{f,i\tau}) = \exp\left(\widehat{\delta}\,\widehat{\gamma}(v_{i\tau}) + \kappa\right).$$

1.6 Results

This section presents the results of the empirical strategy just described. First, I present the parameter estimates for θ, $\gamma(v)$, δ and $\mu(v, \widetilde{\omega}_f)$, as well as the distribution of α_i^f. Then, I present the prevalence of violence implied by the model.

Parameter Estimates

The predicted functional form of $\gamma(v)$ is

$$\widehat{\gamma}(v_{i\tau}) = \widehat{\beta} v_{i\tau}^2,$$

where β represents the parameter β_2 in equation (1.5.3).[22] From the constant $\widehat{\beta}_0$ of equation (1.5.3), I recover the associated technology parameter $\widehat{\theta}$, which has mean 0.86 and standard error 0.02. The predicted technology of the home production is of the from

$$\widehat{Q}_{i\tau} = e^{-0.85 v_{i\tau}^2} (d_{i\tau} + t_{k,i\tau})^{0.86} (1 - l_{f,i\tau})^{0.14}.$$

[22] Several other degrees of the polynomial were considered. The final specification is the one with the highest adjusted R^2 preserving the concavity of $\gamma(v)$. See Tables A.3 and A.4, and Figure A.1.

Figure 1.3 shows that at low levels of violence there is limited resource destruction. On average, violence destroys 4 percent of female productivity, but as it increases, it can destroy up to 38 percent of the female labor productivity.

One can use the productivity cost of violence to generate a market value of preventing one woman from being abused. Consider the average beneficiary of *Food, Cash, or Voucher*. This female earns around 2 dollars for each hour of efficient labor. Suppose that her intimate partner inflicts violence to the average index level. This means that instead of earning 2 dollars an hour, she would earn 1.9 dollars. If she works 5 hours a day and the household earns 400 dollars a month, violence would cost 10 dollars per month. For a husband inflicting the maximum level of violence, violence will account for the destruction of one fourth of the household monthly income or 95 dollars.

Figure 1.4 plots the estimated distribution of the disagreement in preference parameter $\widehat{\alpha}_i^f$.[23] As described in Section 1.5, I use the household fixed effect of equation (1.5.5) to recover the $\widehat{\alpha}_i^f$ for every household. The distribution of $\widehat{\alpha}_i^f$ comes from a bootstrap in which, for each sample draw we: (i) estimate the household fixed effect and recover $\widehat{\alpha}_i^f$, (ii) generate bins of equal size of $\widehat{\alpha}_i^f$, and (iii) for each bin, compute the median and the frequency. In the average household, the female prefers the home good over the market good 10 times more than the male. Nevertheless, there are households with low disagreement where the female cares about the home good only twice as much as the male. In other households, female valuation for the home good can be as high as 20 times the valuation of the male. These relative preferences for the home good are, on average, uncorrelated with female relative wages.[24]

Figure 1.5 presents the predicted relationship between violence and female weight in the overall household utility. Given the identification assumptions described in Section 1.5, the response of female relative weight to violence has the functional form of: $\mu(v, \widetilde{\omega}_f) = \exp(\delta\gamma(v) + \kappa)$. The estimation result of equation (1.5.4) suggests the parameter $\widehat{\delta}$ has mean 3.05 and standard error of 0.32 (see Table 1.3). By normalization, in the absence of violence, the female has the same say as the male, $\mu_{\max} = \exp(\kappa(\widetilde{\omega}_f)) = 0.5$ for every $\widetilde{\omega}_f$. If the husband inflicts violence at the average index, he reduces her say by 10 percentage points. As violence increases, it can reduce her relative weight in the household decision making by half.

Simulations

I use the above estimated parameters (see Table 1.3 and Figure 1.4) to simulate the model described in Section 1.2. As a comparative statics exercise, figures 1.6 to 1.9 present the graphical representation of how the choice variables (violence v, female labor supply l_f, and demand for market input d), the consumption of public goods (home good Q, and market good q), and utilities (u^f and u^m) vary with the female relative preferences over Q, fixing the female wages at $w_f = 0.5$. Later on, when computing the prevalence of violence and performing counterfactual analysis, I will use the entire distribution of female relative wages depicted in Figure 1.2.

Figures 1.6 to 1.9 plot how the choice variables (violence v, female labor supply l_f, and demand for the market input d) change with the extent of disagreement (α_i^f). In each figure, the black solid line depicts a situation under no transfers, the blue dashed line

[23]See Figure A.2 for the distribution of α_i^f in the original sample.

[24]The average correlation between the disagreement in preferences and the female relative wages is -0.0001654.

represents the outcome under a cash transfer equivalent to 10 percent of the male income, and the red dotted line represents the equivalent in-kind transfer. In the x-axis I plot the disagreement in preference parameter, α_i^f, ranging from 2 to 50, as in the estimated distribution of α_i^fs (see Figure 1.4).

In Figure 1.6, I observe how, when female's preference for the home good (Q) is relatively similar to that of the male, violence is not profitable because there are limited utility gains that he can appropriate. As her relative preference for the home good increases, there is more disagreement. From the male's perspective, without violence the household would devote too many resources to the home good production. Thus, he inflicts violence to align the household resource allocation with his preferences.

Transfers alleviate part of the tension in the family by increasing the resources available. With less disagreement, there is a larger range of α_i^fs for which there is no violence. There is also a range of α_i^fs for which in-kind transfers decreases violence while cash transfers do not. The reason is that, although there is disagreement, by making the transfer in-kind rather than in-cash, the government resolved part of the conflict.

Figure 1.7 plots how female labor supply changes with female relative preferences for the home good. As female's preference for Q increases, her opportunity cost of devoting her time to the labor market rather than to the home good production increases. Consequently, she is less willing to allocate her time to the labor market and female labor supply falls. The male does not like this shift as the reduction of female labor supply reduces her labor income, which the household could use to buy his preferred good. Although violence dampens female productivity, it allows the male to enforce a larger allocation of resources towards the market good. This effect compensates the productivity loss. Besides, as female labor supply decreases with α_i^f, the resource destruction effect of violence concentrates on the good the male cares relatively less (Q), so that from his perspective, violence is less costly. The demand for market input d depicted in Figure 1.7 reflects such non-monotonicity. When disagreement is low, higher preferences for the home good lead to higher demand for the market input. Once there is violence, the resource destruction lowers the marginal productivity of the market input, which in turn lessens its demand.

Figure 1.8 plots how these effects translate into the home good production and the demand for the market good. Up to the threshold where violence starts being profitable, the home good increases with female's preference for it. Once there is violence, the home good production falls. Relative to households with less disagreement, the resources liberated from the home production are devoted to the market good, which is relatively more preferred by the male. This transfer of utilities from the female to the male occurs as violence dampens female weight in the overall household utility.

Finally, Figure 1.9 plots how these changes in disagreement in preferences map to utilities. At low levels of disagreement, the fact that the female cares more about the home good improves everyone's well-being. Once violence is used, as disagreement increases, the utility of both the female and the male falls because more resources are destroyed.

Prevalence of Violence

I use the estimated distribution of α_i^fs depicted in Figure 1.4 and the empirical distribution of female relative wage presented in Figure 1.2 to compute the prevalence of violence. In the absence of transfers, 17.6 percent of the households in this economy would exhibit some violence. A program such as *Food, Cash or Voucher*, which randomly allocates two thirds of the transfers in-kind (*Food* and *Voucher* treatment-arms combined) and one third in-cash, would reduce violence by 9 percentage points.

1.7 Policy Experiments: In-Kind versus Cash

This section presents a policy experiment trying to quantify the differential effect that in-kind and cash transfers have on violence. Similar to Section 1.6, I use the estimated distribution of α_i^fs depicted in Figure 1.4 and the empirical distribution of female relative wage presented in Figure 1.2 to compute the rate of violence under different transfer regimes.

In the absence of transfers, 17.6 percent of the households experience intimate partner violence. If the government implements a program of transfers to the women that provides a cash transfer equivalent to 10 percent of the average household income, the rate of violence would decrease to 9.8 percent. The 7 percentage point reduction violence is consistent with the point estimates of other impact evaluations addressing the effect of CCTs on violence. For instance, in the context of Mexico *Oportunidades*, Bobonis et al. (2013) find that the program reduces physical abuse by 5 to 7 percentage points.

I also use this framework to analyze what would be the reduction in violence if the same transfers were given in-kind rather than in-cash. According to the simulations, if the transfers were given in-cash violence would fall by 44 percent. If the transfer were exclusively in-kind, violence would fall by 57 percent. To put it differently, 17 out of every 100 women beneficiary of *Food, Cash, or Voucher* are victims of intimate partner violence. With a cash transfer program, the program would guarantee that 7 of these 17 women would live in non-violent households. Yet if the same transfers were given in-kind, additional 3 women would no longer be abused.

A policy relevant question is to how to valuate the fact in-kind transfers would allow for 3 additional women to live in non-violent households. In addition to the transfer itself, to provide an in-kind transfer the government must incur in the additional cost of looking for local providers and paying the transportation cost of bringing the transfer to the beneficiaries. Therefore, we should compare the gains from in-kind transfers to the additional implementation cost. For *Food, Cash, and Voucher*, Hidrobo et al. (2014) suggest that the cost of providing a food transfer, a cash transfer, and a voucher transfer are 11.46 dollars, 2.99 dollars, and 3.27 dollars per household, respectively. Therefore, it seems to be the case that the 8.47 dollars difference in the implementation are offset by the 10 dollars monthly reduction of income associated to violence, as described in Section 1.6.

1.8 Scaling-up the Program

This section combines the results of *Food, Cash, and Voucher* with national level data to make out-of-the-sample predictions relevant for Ecuador's anti-poverty program, *Bono de Desarollo Humano*. *Bono de Desarollo Humano* was introduced in 2003 to replace Ecuador's main assistance program *Bono Solidario*. Under *Bono de Desarollo Humano*, around 40 percent of the families in Ecuador receive a transfer to the woman. In 2011, the program gave 35 dollars a month to every beneficiary family. Given the similarities with *Food, Cash, or Voucher*, I use the results of *Food, Cash, or Voucher* to simulate the effect that *Bono de Desarollo Humano* has on violence and the potential effects of transforming the program into an in-kind transfer one.

Predicting Disagreement in Preferences at the National Level

To replicate the exercise described in Section 1.5 at the national level, I would require a data set with repeated observations per household, containing information on violence, time allocation, wages, and food consumption. This type of data, however, is scarce because most of the surveys on *Violence Against the Women* abstain from making multiple visits to the same household to preserve the integrity of the respondents. To circumvent this issue I use the estimates presented in Section 1.6 to draw conclusions about a *Bono de Desarrollo Humano*, using a cross-sectional national representative data base.

I use data from the *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres*. This is a nationally survey representative of all women aged 15 or more, living in urban and rural areas of Ecuador. The survey was implemented in 2011 by the Instituto Nacional de Estadística y Censos (INEC), and collects information on family history, marital status, violence, time allocation, and labor market participation. To ensure comparability with the families beneficiaries of *Food, Cash or Voucher*, I restrict to the set of households where the respondent was a female head of household or the spouse of the head, she is aged between 15 and 70-years-old, and she was married or living in cohabitation. Because I am interested in drawing results relevant for *Bono de Desarollo Humano*, I concentrate in the households that report being beneficiaries of the program.

Table 1.7 presents some descriptive statistics of the sample. Similarly to the beneficiaries of *Food, Cash, or Voucher*, the average household is a family of 5 members. The head of the household is a 45-years-old working man. His spouse is a 41-years-old woman, and they have one child aged under 5 and another child under 14. Unlike the beneficiaries of *Food, Cash, or Voucher*, 65 percent of these families are married. The remaining 35 percent live in cohabitation. Both female and male have, on average, 4 years of education. Yet 22 percent of the women are more educated than their partners. Around 32 percent of women work in the labor market for an average of 8.5 hours a day. More than 90 percent of the partners of these women are employed, and they also work for 8.5 hours a day. Figure 1.10 plots the empirical distribution of female relative per hour wage rate in logarithms.

Regarding violence, almost 4 out of every 10 women (37%) have been victims of a form of physical violence by an intimate partner in the last 12 months. During this time period, none of the women reported an episode of sexual intimate partner violence. Even if this is a sub-sample of the population in Ecuador and in the descriptive statistics I do not use the expansion factors of the survey, the rate of violence in the sample is close to the 35 percent national prevalence of physical abuse (Instituto Nacional de Estadística y Censos, 2011).

As inputs for simulating the model for an economy with the characteristics described in Table 1.7, I use the estimated technology of home production ($\widehat{\theta}$), the productivity cost of violence ($\widehat{\gamma}(v)$), and the response of female relative weights to violence ($\widehat{\mu}(v, \widetilde{\omega}_f)$) presented in Table 1.3 and in Figures 1.4 and 1.5. A key element for the simulation is, however, missing: the disagreement in preferences parameter (α_i^f). The disagreement between female and male is observable for the household, but not for the econometrician, nor for the government. To overcome this issue, first, I run a regression of the unobserved heterogeneity term, $\widehat{\alpha}_i^f$, on household observable characteristics among the beneficiaries of *Food, Cash, or Voucher*. Then, I use these coefficients to produce the predicted disagreement parameters in the national data set.

Table 1.8 presents the results of the first step.[25] Each point estimate comes from a bootstrap in which, for every sample draw of the beneficiaries of *Food, Cash, or Voucher*, I recover the α_i^fs for every household (as explained in Section 1.5), and run a regression of α_i^fs on some of the household observable characteristics.Households with more members and a male as head of the household tend to have more disagreement. Females with higher of income have higher conflict at home. Nevertheless, as the male income increases, the conflict reduces.

I proceed by using these coefficients to predict the disagreement in preferences that would exhibit the beneficiaries of *Bono de Desarrollo Humano*. For each of the households in the sub-sample of the described in Table 1.7, I predict a parameter α_i^f using the household characteristics and the coefficients in Table 1.8. Figure 1.11 plots the predicted distribution of α_i^f.

In-Kind versus Cash at the National Level

I use the predicted distribution of α_i^fs depicted in Figure 1.11, and the empirical distribution of female relative wage depicted in Figure 1.10 to compute the prevalence of violence without transfers, and with a cash transfer corresponding to 10 percent of the average household income. This is equivalent to simulate the effect of *Bono de Desarrollo Humano* on intimate partner violence.

According to the simulation, 36 percent of the households in the economy experience a form of physical violence. A cash transfer regime transforms 13 of these 36 households into non-violent households. Therefore, it is as if *Bono de Desarrollo Humano* reduces physical intimate partner violence by 35 percent. This result differs from the impact evaluation of the program done by Hidrobo and Fernald (2013) according to which *Bono de Desarrollo Humano* has no effect on physical violence. Yet the impact evaluation was done using the data of 2003, when the size of the transfer was 15 dollars.

I also use the model to consider the additional reduction in violence that a program such as *Bono de Desarrollo Humano* could attain if the transfers were given in-kind. As described in Table 1.6, if the transfer of *Bono de Desarollo Humano* were in-kind, the prevalence of violence would decline from 36.10 percent to 19.40 percent. This means that the differential effect on violence of in-kind versus cash amounts for 4 additional women living in non-violent households.

Finally, I also use the model to simulate the prevalence of violence for different size of transfers (as share of the average household income). Figure 1.12 presents the results. The red dotted line represents an in-kind transfer, and the blue dashed line represents a cash transfer. As the size of the transfer increases, the differential effect of in-kind versus cash amplifies. This result derives directly from the argument exposed through Figure 1.1 as with larger transfers, the gap between the male's maximum utility under in-kind transfer ($\overline{u}_k^m$) and under a cash transfers ($\overline{u}_c^m$) increases. As a result, the additional gains of violence under cash also increase.

[25] Table A.6 presents the regression of the disagreement in preference parameter on household observable characteristics, in the original sample.

1.9 Conclusions

In this chapter I address how intimate partner violence responds to transfers whenever the male can use violence to boost his bargaining power. With this research question in mind, I construct and estimate a model of household with an explicit role of violence. I understand the household as a space of cooperation where there is also room for disagreement and conflict. If there were no violence, the household would attain efficient outcomes. Yet, in the presence of disagreement, the male has incentives to inflict violence to enhance his relative position in the household decision making; but at the expense of making his partner less productive.

Under this setting the type of transfer matters. The reason is that different type of transfers generate different gains of violence. Although both in-kind and cash transfers to the woman improve her outside option, in-kind transfers have an additional margin in the reduction of violence as they mitigate part of the spousal disagreement and make violence less profitable as an appropriation device.

I estimate the model using the data from *Food, Cash or Voucher*. With it, I am able to provide a market value for the cost of violence, and quantify the effect of violence on the female relative weight. According to the estimations, violence can destroy up to 38 percent of female productivity, and reduce her relative weight in the household decision making by almost half. I also find that 17.6 percent of the women are victims of violence and a program such as *Food, Cash or Voucher* reduces violence by 10 percentage points. Using the data from the *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres* in Ecuador, I extent the model to a much larger program such as *Bono de Desarrollo Humano*. The simulations suggests that the program reduces physical intimate partner violence by 35 percent, and additional improvements could be achieved if in-kind transfers were considered for the intervention.

Chapter 2

Does Rewarding Pedagogical Excellence Keep Teachers in the Classroom? Evidence from a Voluntary Award Program

*written jointly with Samuel Berlinski**

2.1 Introduction

Successful public schools systems retain the best teachers in their classrooms. Yet compensation policies in many countries do not provide much help in achieving this goal. In the US, for example, teachers earn 67 percent of what they could have earned in other career paths (OECD, 2013). Not surprisingly, seven percent of US teachers leave the profession every year, presumably with relatively higher separation rates among those with better outside options.[1] Because wage schedules are traditionally based on factors that are not necessarily related to classroom performance (e.g., education, experience, hours worked), they are not flexible enough to reward the best teachers.

A large literature in personnel economics focuses on the role that wages play in motivating, retaining and recruiting workers.[2] Tying wages to a performance measure may both motivate workers and allow firms to retain its most productive workers. These personnel policies crucially depend on workers knowing their ability/type and on higher ability workers being able to sort themselves at a relatively lower cost than lower ability ones. If ability can be revealed through observation and testing at a reasonable cost for schools and teachers, it may provide a suitable tool to keep the best teachers in the classroom.

In this chapter, we analyze the effects on teachers' retention and between school mobility of a program that rewards excellence in pedagogical practice in Chilean primary

*The views expressed herein are those of the authors and should not be attributed to the Inter-American Development Bank, its Executive Directors, or the governments they represent.

[1] Statistics from the U.S. Department of Education. http://nces.ed.gov/pubs2014/2014077.pdf

[2] See Prendergast (1999) for a review.

and secondary schools (a country with comparable teacher turnover rate to the US). The Pedagogical Excellence Award initiative aims to identify good teachers, prevent them from leaving the public school system, and allocate them where they are needed the most (Araya-Ramirez, Taut, Santelices-Etchegaray, Manzi-Astudillo, and Mino-Flores, 2012; Rodriguez, Manzi, Peirano, Gonzalez, and Bravo, 2015). Teachers apply voluntarily for the award which is allocated on the basis of teachers' knowledge of their field and their pedagogical skills. In order to receive the award, teachers must prepare a teaching portfolio and take a knowledge test. The results of both assessments are combined in a final score and only those scoring above a certain cut-off receive the award. The teachers that succeed on the assessments are awarded the equivalent of a six percent yearly wage increase for up to ten years, after which they need to re-apply for the program.

We formalize our identification conditions using a simple model of quit behavior (Guasch and Weiss, 1980, 1981) and teacher testing (Angrist and Guryan, 2004). Teachers decide whether or not to take the test and, after observing the results, they decide whether or not to quit. The retention effect depends on the test difficulty. The easier is the test, the more likely it is that teachers scoring around the threshold are being paid above their reservation wage and that they will be willing to stay in the school system irrespectively of receiving the award. In contrast, when the test is rather difficult the award can alter teachers' decisions to quit at all margins.

Using administrative data over eight cohorts of applicants, our estimates indicate that locally the award does not alter transitions out of the school system. This suggests that teachers marginally failing to receive the award value their jobs more than their outside option. We observe, however, an increase in mobility within the school system among teachers that receive the award. Some of these mobility patterns are consistent with the award providing a signal of teacher ability. Contrary to the spirit of the policy, awardees in schools with relatively low performing students and working conditions tend to move after receiving the award.

To our knowledge we are the first to provide estimates of the impact on teacher mobility of a voluntary award system that ties wages to an input measure of classroom performance. The education literature suggests that teacher separation responds to changes in basic compensation (Dolton and Van der Klaauw, 1995, 1999; Clotfelter, Glennie, Ladd, and Vigdor, 2008; Falch, 2011) and that teachers' effort increases when compensation is tied to student performance (Lavy, 2002, 2009; Muralidharan and Sundararaman, 2011). Our research is also related to the literature on occupational licensing. Compulsory licensing imposes a barrier to entry, which reduces the supply of labor and increases labor costs.[3]

The rest of the chapter is organized as follows. In Section 2.2 we provide some background on the Chilean education system and the design of the program. Section 2.3 describes the data used. In Section 2.4 we model the decision to quit teaching and study the margins at which a program with the basic features of the Chilean Pedagogical Excellence Award program can affect behavior. In section Section 2.5 we present our identification strategy and relate it to the model presented in Section 2.4. In Section 2.6 we present our results. Section 2.7 concludes.

[3]See Kleiner (2000) for a thorough discussion on occupational licensing.

2.2 Background

Primary and secondary education in Chile is provided by three type of institutions: municipal or public schools, private-subsidized schools, and private schools. Municipal schools are non-profit institutions that offer instruction to students for free. They receive a per-student subsidy from the Ministry of Education and are administered by municipalities. Private schools are for profit institutions that charge tuition to students. They receive no subsidies from the government and are administered as private corporations. Private-subsidized schools are run like private schools, they receive the same per-student subsidy than municipal schools and can also charge a tuition (Mizala and Schneider, 2014; Hsieh and Urquiola, 2006).[4] We refer to municipal and private-subsidized schools as the Voucher School System.

The contractual arrangements for teachers are different in the three type of providers.[5] In Figure 2.1, we use data on wages, age category and type of school to construct wage-age profiles for teachers. The data comes from *Encuesta Longitudinal Docente 2005: Análisis y Principales Resultados*, a national representative survey of that collects information on socio-demographic characteristics and employment history of 6,000 Chilean teachers. As one will expect, wages increase with age. Wages in the private sector are uniformly higher. For younger teachers, wages in private-subsidized schools are higher than in the municipal sector but wages increase faster in the municipal sector. In fact, the level of wages is practically equal for the 41-50 age group. After this age, municipal school teachers are paid a higher per hour wage rate than private-subsidized schools.

In Figure 2.2 we present the share of students enrolled and teachers employed in primary and secondary schools during 2004-2013. Enrollment in the Voucher System over this period is pretty stable at around 93 percent. However, there have been large compositional changes between municipal and private-subsidized schools. In 2004, 50.4 percent of the students were enrolled in municipal schools; while in 2013, it was only 39 percent. This has, of course, caused a commensurate shift in the share of teachers employed in municipal and private-subsidized schools.

In Chile, like in the US (Hanushek, Kain, and Rivkin, 2004), there is considerable teacher turnover, particularly among the least experienced teachers. As we show in Table 2.1, from all teachers employed in 2003, two years later 12 percent were no longer teaching and 9 percent have changed schools. These figures are even larger for those with less than 11 years of experience (18 and 15 percent, respectively).

The Chilean government perceived that many good teachers were leaving the profession and introduced a voluntary award program designed to reward, both economically and socially, excellence in teaching practice: The Pedagogical Excellence Award (*Asignación a la Excelencia Pedagógica*) or AEP (following its Spanish acronym).[6] Starting in 2002, teachers employed in municipal schools and private-subsidized schools could apply

[4]The fees that these schools can charge to students are regulated.

[5]The employment of teachers in municipal schools follows a union negotiated teacher statute. In the private sector, employment follows the standards established by common labor law. Employment of teachers in private-subsidized schools retain some aspects of the municipal school system Mizala and Romaguera (2005); Santiago, Benavides, Danielson, Goe, and Nusche (2013). For example, minimum wages, bonuses, and maximum working hours are determined by the Teachers Statute. Yet, after reaching the retirement age (60 years for women and 65 for men) teachers are no longer allowed to teach in municipal schools, but they can still teach in private-subsidized institutions.

[6]AEP was established by law in 2001 (Law 19715). Modifications to the law were introduced in 2006 (Law 20158) and 2011 (Law 20501).

for this award. Eligible candidates must teach at least 20 hours a week during the academic year in Voucher System schools. The award entitles beneficiaries with a 6 percent annual salary increase for up to ten years.[7][8] The magnitude of the bonus varies at four levels of experience: 0-11 years, 12-21 years, 22-30 years, and 31 plus years. In addition, those awarded the AEP are invited to become mentors of other teachers in the Network of Teachers of Teachers (*Red Maestro de Maestros*).[9] The awards are presented in a ceremony with local authorities and media coverage. Teachers can apply for an award only twice within each level of experience.

To receive the AEP award, teachers must prepare a teaching portfolio and take a written test in their main area of expertise. In the portfolio, teachers must demonstrate their teaching practices. This assessment requires a learning plan for the students, an evaluation strategy, a pedagogical reflection and a recording of a class. In the written test, teachers are evaluated on grounds of their knowledge. The results of these two assessments are combined in a final score ranging from 100 to 400. For the AEP rounds taking place between 2002 and 2011, the final score was a weighted average with 70 percent of the weight given to the portfolio and 30 percent to the written test. Only teachers with a final score of at least 275 receive the award.[10]

The application process for the AEP begins in April. The portfolio is prepared from July to October, and the written examination takes place in November. The school year starts in March and teachers learn about their score in April. For those who are successful, payments are done twice a year with the first installment in July. We present this time line in Figure 2.3.

There are other incentive mechanisms built into the Chilean education system. In 1996, the National System for Performance Evaluation (*Sistema Nacional de Evaluación del Desempeño*) or SNED (following its Spanish acronym) introduced collective performance incentives in the Voucher School System. Every two years, SNED gathers information about schools' performance at a standardized national examination, repetition and dropout rates, educational activities provided, parental participation in school activities, and overall working conditions. After grouping schools in sets with similar students' socioeconomic characteristics, SNED ranks schools using an aggregate index that combines the factors described. The best schools in each group (accounting for up to 35 percent of the enrollment in the set) receive a monetary transfer for two years. That transfer is distributed among teachers and accounts for a 50 percent to 70 percent of a monthly salary (Mizala and Urquiola, 2013).

In 2004 the Ministry of Education implemented a compulsory examination for municipal school teachers. Every 4 years, teachers of municipal schools are assessed through a written examination (*Evaluación Docente* (EV)).[11] Municipal school teachers with an

[7] AEP bonus is equivalent to 70 percent of a monthly salary.

[8] After 2011, the AEP award period was reduced to four years.

[9] AEP awardees willing to become members of the Network of Teachers of Teachers are required to present another portfolio. If they score above a certain threshold, they become permanent members of the Network. Teachers who are not selected can reapply every three years for as many times as desired. Members of the Network receive an additional monetary incentive tied to the hours worked. 40 percent of the AEP awardees become members of the Network at some point after receiving the award.

[10] This cut-off point was identified by inspecting the data and was confirmed by the *Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógica* (CPEIP) in internal correspondence. To our knowledge, there is no official document where the threshold is stated.

[11] Teachers failing the test can retake it up to three times. After a third failure, teachers are fired from the Voucher System. From 2004 to 2013, around 1.5 percent of the teachers that took the test fail it at the first attempt but less than 0.1 percent took the examination more than twice.

outstanding evaluation can apply to a performance award: *Asignación Variable al Desempeño Individual* or AVDI (following its Spanish acronym). For this purpose, teachers must take the same knowledge test than for AEP. Teachers can receive both the AEP and the AVDI award, and can apply to them simultaneously (although not many do). We focus the main body of the chapter on the AEP as the data suggests that the conditions for identification using a regression discontinuity design are not fulfilled for AVDI. For completeness, we provide the analysis of AVDI along the lines of our work for AEP in an online appendix.

2.3 Data

We use administrative data from all teachers in the school system published yearly by the Ministry of Education. The data set starts in 2003 and contains information on basic demographics, educational qualifications, experience, place and hours of work. We match it with the scores and award status of individual applicants to AEP and AVDI and with school level data from SNED. Figure 2.4 presents a sample flowchart. We start with the 13,098 teachers that applied for the first time for an AEP award between 2003 and 2011.[12] Further, we restrict to individuals who applied for an award as primary or secondary school teachers.[13] We match this data with administrative records and restrict our analysis to individuals that at the time of application were at least four years away from the retirement age (i.e., 56 for females and 61 for males). For brevity, our main results focus on the sample of 9,311 teachers that are not concurrently applying to AVDI.

We start by showing that the assignment rule was strictly enforced. In Figure 2.5, we plot the mean of a variable that takes the value of 1 if an individual has an AEP award and 0 otherwise for each possible score cell (circles). There is clearly a sharp discontinuity. Those who obtained the award have an aggregate score of 275 or more. In Table 2.2, we present the awardee rate and final scores by year. We divide the data in two samples, Panel A has the 9,311 teachers from our benchmark sample and Panel B has the 13,098 first time applicants. The table confirms the information on the graph: compliance with the allocation rule is above 99 percent regardless of the application wave or sample. Focusing on Panel A, 28 percent of the teachers that apply for an AEP obtained it. There are significant differences, however, in the passing rates over time; while 44 percent of the 2003 applicants received the award, less than 22 percent did so after 2007. Finally, the awardee rates in Panel B are slightly smaller than in Panel A, but we cannot detect systematic differences in the final score.

In the first column of Table 2.4, we present average information for all employed teachers in the Voucher School System during the 2003-2014 period. In the second column, we present the same information but only for those who have applied to AEP during the 2003-2011 window. Beginning with basic demographic and qualification variables, we observe that over the 2003-2014 period, the average Chilean teacher is a 44 years old woman with a degree in education and 17 years of teaching experience. Teachers work on average 35 hours a week, around 10 percent work in more than one school, 75 percent work as primary school teachers, 10 percent hold a managerial position, and 40 percent work at private-subsidized schools. Every year, 12 percent of the teachers

[12]We eliminate 2002 AEP applicants because of lack of administrative data.

[13]We eliminate those applying for the award in pre-primary education, adult education and special education as they face radically different inside and outside options that teachers in primary and secondary schools.

change schools and 7 percent move to a different municipality. Around 42 percent of the teachers work in municipalities considered as isolated and are monetarily compensated with an allowance.[14] We average schools' working conditions and students' performance using information from SNED between 2003 and 2014, and rank the schools based on these variables. Around 40 percent of the teachers work in schools ranked in the top 50 percentile in terms of working conditions and 63 percent in schools ranked in the top 50 percentile in terms of student performance.

Of the employed teachers, 6.5 percent applied to AEP at some point between 2003 and 2011. From these applicants, 42 percent also applied to AVDI. Only 1.2 percent of the teachers are recipients of the AEP award and 2.6 percent are AVDI recipients. Relative to the average Voucher System teachers, AEP applicants are slightly younger, more likely to have a degree in education, and more likely to work in a school with top-performing students. In the third and fourth column of Table 2.4, we describe the sample at the first time of application to AEP and two-years after. Two-years after applying to AEP, 4 percent of the teachers are not employed in the school system[15], 1 percent work in a private school, 10 percent change from municipality, and 16 percent moved to a different school from the one they were at when applying.[16] The number of contract hours, the rurality of the school, the percentage working in subsidized-private schools, the percentage working in school with better working conditions and student performance all fall relative to the baseline measure.

2.4 A Model of Teachers' Quit Behavior

One of the goals of AEP is to prevent good teachers from leaving the profession (Araya-Ramirez et al., 2012). With this aim, the program entitles whomever pass the assessment with a fix monetary award and a token of social recognition. These incentives increase the marginal benefit of being in the profession and raise the opportunity cost of quitting. In this section, we provide a simple model that captures these features and speaks to the margin of behavior that the identification strategy described in Section 2.5 is able to capture.

Similar to Guasch and Weiss (1980, 1981) and Angrist and Guryan (2004), we consider a continuum of teachers i, characterized by their productivity as teachers or ability, ω_i, where $\omega_i \in [0, 1]$. Teachers are risk neutral, and their productivity is distributed following $\omega_i \sim f(\omega)$. Teachers observe their productivity upon entering the profession, but the school system only observes the overall distribution.

Assume that ω_i also captures teachers' reservation wage. Teaching pays a fixed wage w. Without lost of generality assume that $E(\omega_i) \leq w$. A teacher i with reservation wage ω_i will stay in the profession if and only if

$$w \geq \omega_i.$$

The government wants to retain all teachers of at least productivity $\widetilde{\omega} > E(\omega_i)$,

[14]In Chile, the *D.L. 249* of the Fiscal Sector establishes a percentage increase in the RMBN for civil servants working in zones considered as isolated or with high cost of living. We extract allowance information at the municipality level from the *Ley 19.354* of 1994 and use those percentages along the 2003-2014 period.

[15]This variable takes a value of zero for those teachers who have dropped from the sample.

[16]For example, someone without a contract one year after applying but with a contract two years after applying is classified not at work in the first year and at work in the second year.

without increasing the base salary. Even if productivity cannot be directly observed, the government can design a test where the score, s_i, is an increasing function of teachers' productivity, ω_i, and some measurement error, ν_i. We assume that the measurement error follows a distribution $\nu_i \sim g(\nu)$, is symmetrically distributed around a mean of zero and orthogonal to ability. Furthermore,

$$s_i = \omega_i + \nu_i. \tag{2.4.1}$$

Taking the test is costly for teachers and we define this cost, in monetary terms, as c. To create incentives for more productive teachers to stay in the school system, the government pays a bonus b to all teachers that voluntarily take the test (i.e., pay the cost c) and score above the cut-off $\widetilde{\omega}$.

A teacher whose reservation wage is at most the fix wage, $\omega_i \leq w$, will sit the exam if the expected pay-off of taking the test is at least the fixed wage. In other terms,

$$\begin{aligned} p(\omega_i)\,(w+b-c) + (1-p(\omega_i))\,(w-c) &\geq w, \\ p(\omega_i) &\geq \frac{c}{b}, \end{aligned} \tag{2.4.2}$$

where $p(\omega_i)$ is the probability of passing the exam for a teacher of productivity ω_i. Likewise, a teacher with reservation wage above the fix wage, $\omega_i > w$, will sit the exam if the expected pay-off is at least as high as her reservation wage

$$\begin{aligned} p(\omega_i)\,(w+b-c) + (1-p(\omega_i))\,(\omega_i-c) &\geq \omega_i, \\ p(\omega_i) &\geq \frac{c}{w+b-\omega_i}. \end{aligned} \tag{2.4.3}$$

We can now characterize the teachers' decision to quit around the cut-off $\widetilde{\omega}$. Using equations (2.4.2) and (2.4.3), we define the probability of receiving the award for the lowest, $\underline{\omega}$, and highest, $\overline{\omega}$, productivity teachers that take the exam as:

$$\begin{aligned} p(\underline{\omega}) &= \frac{c}{b} \quad \text{and} \\ p(\overline{\omega}) &= \frac{c}{w+b-\overline{\omega}}. \end{aligned}$$

For a teacher of productivity ω_i, the probability of receiving the award is:

$$\begin{aligned} p(\omega_i) &= \begin{cases} 1 & \text{if } \omega_i + \nu_i \geq \widetilde{\omega} \\ 0 & \text{if } \omega_i + \nu_i < \widetilde{\omega} \end{cases} \\ &= \Pr(\omega_i + \nu_i \geq \widetilde{\omega}) = 1 - \Pr(\widetilde{\omega} - \omega_i \geq \nu_i) \\ &= 1 - G(\widetilde{\omega} - \omega_i) = G(\omega_i - \widetilde{\omega}), \end{aligned}$$

where $G(.)$ is the CDF of the measurement error term. Therefore, $\underline{\omega}$ and $\overline{\omega}$ can be expressed as,

$$G(\underline{\omega} - \widetilde{\omega}) = \frac{c}{b} \quad \text{and} \tag{2.4.4}$$

$$G(\overline{\omega} - \widetilde{\omega}) = \frac{c}{w+b-\overline{\omega}}. \tag{2.4.5}$$

Let $q(\omega_i)$ be the probability that a teacher of productivity ω_i stays in the profession. Notice that $q(\omega_i)$ depends both on the probability of taking the exam and the probability of passing it. In particular,

$$q(\omega_i) \;=\; \begin{cases} 1 & \text{if } \omega_i \leq w \\ p(\omega_i) & \text{if } w < \omega_i \leq \overline{\omega} \\ 0 & \text{if } \omega_i > \overline{\omega}. \end{cases} \tag{2.4.6}$$

Consider the case where b, c and $\widetilde{\omega}$ are such that equations (2.4.2) and (2.4.3) hold for a positive mass of teachers. In other words, some teachers whose outside option is below the current teacher wage apply for the award and some teachers who in the absence of the award will quit sit for the exam as well (i.e., $\underline{\omega} \leq w < \overline{\omega}$). It is worth pointing out that these set of conditions do not pose any restrictions on the relation between the difficulty of the test and the fixed teachers wage (this can be seen by manipulating equation (2.4.4)).[17]

Define an interval ϵ around the cut-off $\widetilde{\omega}$ and call $\Delta_\epsilon^{\widetilde{\omega}}$ the difference between the mass of non quitters above and below the cut-off.

$$\Delta_\epsilon^{\widetilde{\omega}} \;=\; \int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} q(\omega_i)f(\omega_i)d\omega_i - \int_{\widetilde{\omega}-\epsilon}^{\widetilde{\omega}} q(\omega_i)f(\omega_i)d\omega_i$$

Let $\widetilde{\omega}_L$ denote a test such that $\widetilde{\omega}_L \leq w$ and $\widetilde{\omega}_H$ denote a test such that $w < \widetilde{\omega}_H$. We now prove that the capacity of the award to affect teachers' quit behavior around the threshold depends on the difficulty of the test.

Proposition 1. For any $0 < \epsilon < \min\{w - \widetilde{\omega}_L, \widetilde{\omega}_H - w\}$ such that $\int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} f(\omega_i)d\omega_i = \int_{\widetilde{\omega}-\epsilon}^{\widetilde{\omega}} f(\omega_i)d\omega_i$, $\Delta_\epsilon^L = 0$ and $\Delta_\epsilon^H > 0$.

Proof. For relatively easy test, $q(\omega_i) = 1$ for all $\omega_i \in (\widetilde{\omega} - \epsilon, \widetilde{\omega} + \epsilon)$, so that

$$\Delta_\epsilon^L \;=\; \int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} f(\omega_i)d\omega_i - \int_{\widetilde{\omega}-\epsilon}^{\widetilde{\omega}} f(\omega_i)d\omega_i.$$

It is straight forward that $\Delta_\epsilon^L = 0$ whenever the distribution productivity is smooth around the cut-off. For a relatively difficult test, $q(\omega_i) = p(\omega_i)$ for all $\omega_i \in (\widetilde{\omega} - \epsilon, \widetilde{\omega} + \epsilon)$, so that

$$\Delta_\epsilon^H \;=\; \int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} p(\omega_i)f(\omega_i)d\omega_i - \int_{\widetilde{\omega}-\epsilon}^{\widetilde{\omega}} p(\omega_i)f(\omega_i)d\omega_i.$$

As the measurement error is independent of productivity,

$$\Delta_\epsilon^H \;=\; \left[\int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} p(\omega_i)d\omega_i - \int_{\widetilde{\omega}-\epsilon}^{\widetilde{\omega}} p(\omega_i)d\omega_i\right] \int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} f(\omega_i)d\omega_i > 0$$

[17] Equation (2.4.4) and $\underline{\omega} \leq w$ imply that

$$\begin{aligned} \underline{\omega} = \widetilde{\omega} + G^{-1}\left(\frac{c}{b}\right) &\leq w, \\ G^{-1}\left(\frac{c}{b}\right) &\leq w - \widetilde{\omega}. \end{aligned}$$

As $g(\nu)$ is symmetric around 0, $G^{-1}\left(\frac{c}{b}\right) \leq 0 \iff \frac{c}{b} < G(0) = \frac{1}{2}$. Therefore, if $2c \leq b$, $w \gtreqless \widetilde{\omega}$.

The first term on the RHS is positive as $p(\omega_i)$ is increasing in ω_i. Formally, let $h : [0, \epsilon] \to \mathbb{R}$ such that $h(x) = p(\widetilde{\omega} + x)$, and $k : [0, \epsilon] \to \mathbb{R}$ such that $k(x) = p(\widetilde{\omega} - x)$. Since $p(\omega_i)$ is strictly increasing in ω_i, $h > k$. Then, by monotonicity,

$$\int_{[0,\epsilon]} h(x)dx > \int_{[0,\epsilon]} k(x)dx \iff \int_{\widetilde{\omega}}^{\widetilde{\omega}+\epsilon} p(\omega_i)d\omega_i > \int_{\widetilde{\omega}-\epsilon}^{\widetilde{\omega}} p(\omega_i)d\omega_i$$

□

Proposition 1 is intuitive: if the difficulty of the test is rather low ($\widetilde{\omega}_L \leq w$), teachers around the threshold are currently being paid above their reservation wage and they will stay in the school system irrespectively of receiving the award. The decision of teachers that are not infra-marginal for a low difficulty test can be affected as long as the bonus b (relative to the cost c) is large enough (see equation (2.4.3)). In contrast, when the test is rather difficult ($w < \widetilde{\omega}_H$) the award can alter teachers' decisions to quit at all margins.

When the test difficulty is low, there is potential room to capture some of the rents of teachers with relatively low reservation wages by increasing the difficulty of the assessment. To see it clearly, we characterize the lowest and the highest productivity teachers taking the exam in terms of its difficulty. Differentiating equations (2.4.4) and (2.4.5) with respect to $\widetilde{\omega}$, we obtain

$$\begin{aligned} \frac{\partial \underline{\omega}}{\partial \widetilde{\omega}} &= 1 \quad \text{and} \\ \frac{\partial \overline{\omega}}{\partial \widetilde{\omega}} &= \frac{cg(\overline{\omega} - \widetilde{\omega})}{cg(\overline{\omega} - \widetilde{\omega}) - G^2(\overline{\omega} - \widetilde{\omega})}. \end{aligned}$$

Therefore, if the cost is sufficiently low, increasing the difficulty of the test *deters* low productivity teachers from applying more than what it *deters* higher productivity ones $\left(\text{i.e. } \left|\frac{\partial \overline{\omega}}{\partial \widetilde{\omega}}\right| < \left|\frac{\partial \underline{\omega}}{\partial \widetilde{\omega}}\right|\right)$.[18] The remaining funds can be used to increase the bonus which, *ceteris paribus*, may prevent teachers with the highest outside options from leaving the educational system. Differentiating equations (2.4.4) and (2.4.5) with respect to b, we obtain

$$\begin{aligned} \frac{\partial \underline{\omega}}{\partial b} &= -\frac{1}{b}\frac{G(\underline{\omega} - \widetilde{\omega})}{g(\underline{\omega} - \widetilde{\omega})} \quad \text{and} \\ \frac{\partial \overline{\omega}}{\partial b} &= \frac{G^2(\overline{\omega} - \widetilde{\omega})}{G^2(\overline{\omega} - \widetilde{\omega}) - cg(\overline{\omega} - \widetilde{\omega})}. \end{aligned}$$

The larger the bonus, the higher the incentives for teachers to take the exam.[19] Yet, if the bonus is sufficiently high, the *entry* effect dominates for high productivity teachers.[20]

Finally, the model speaks only to the decision to leave the school system. If passing the assessment for the award provides an otherwise unobservable signal of ability, the program may also boost mobility within the school system. To the extent that competition in wages among schools is coerced by wage rules, schools may still be available to compete

[18] The referred condition is $2c < \frac{G^2(\overline{\omega}-\widetilde{\omega})}{g(\overline{\omega}-\widetilde{\omega})}$.

[19] Notice $\frac{\partial \underline{\omega}}{\partial b} < 0$ for any set of parameters b,c, $\widetilde{\omega}$. For $\frac{\partial \overline{\omega}}{\partial b} > 0$, we require $c < \frac{G^2(\overline{\omega}-\widetilde{\omega})}{g(\overline{\omega}-\widetilde{\omega})}$.

[20] The referred condition is $\frac{1}{b} < \frac{g(\underline{\omega}-\widetilde{\omega})}{G(\underline{\omega}-\widetilde{\omega})} + \frac{g(\overline{\omega}-\widetilde{\omega})}{G(\overline{\omega}-\widetilde{\omega})}\frac{G(\underline{\omega}-\widetilde{\omega})}{G(\overline{\omega}-\widetilde{\omega})}$.

for teachers in amenities (e.g., working conditions, students ability, etc). Therefore, we will expect that independently of the quit decision the program may increase mobility between schools.

2.5 Identification Strategy

Our goal is to measure the causal effect of obtaining an award on teachers' retention and between school mobility. AEP is assigned using a performance measure which is likely to be associated with other determinants of teacher behavior. Therefore, a naive comparison of the outcomes of awardees versus non-awardees will provide biased and inconsistent estimates of the causal effect of the program. We tackle this issue by using a regression-discontinuity approach. We exploit the sharp discontinuity in the allocation of the award for teachers with 275 points or more in the aggregate evaluation score. In the absence of manipulation around the cut-off, teachers that obtained 275 should be similar to those that obtained 274.[21] As a result, any systematic differences in behavior after the award is granted could be attributed to the program.

We implement the regression discontinuity design using the following estimating equation for a teacher i who applied to AEP in wave τ:

$$Y_{i\tau}^{t} = \alpha + \beta D_{i\tau} + \gamma_{\tau} f(s_{i\tau}) + \delta_{\tau} D_{i\tau} \times f(s_{i\tau}) + \lambda_{\tau} + \varepsilon_{i\tau}^{t}. \tag{2.5.1}$$

$Y_{i\tau}^{t}$ is the outcome variable of interest t years after the candidate applied for the award (e.g., weekly hours of work in the voucher school system), $D_{i\tau}$ is a variable equal to 1 if the teacher composite score at the exam was at least 275 and 0 otherwise, $s_{i\tau}$ is the teacher's score centered around the 275 cut-off, $f(s_{i\tau})$ is a suitable polynomial function of the composite score and λ_{τ} is a set of wave fixed effects. We allow the effect of the running variable to differ across waves as well as at both sides of the cut-off.[22]

We are interested in the parameter β. Under suitable assumptions, β provides a local measure of the causal impact of obtaining the AEP award. The basic identifying assumption is that there is no systematic manipulation of the running variable around the cut-off. There are at least two strategies to test the plausibility of this assumption (Bloom, 2012; Hahn et al., 2001; Imbens and Lemieux, 2008; Lee and Card, 2008; Lee and Lemieux, 2010). First, there should be no kinks in the density of the score around the discontinuity. Second, predetermined factors ought to vary smoothly around the 275 cut-off.

In Figure 2.6, we plot the histogram of the final score for the pooled sample of applicants. In column one of Table 2.3, we present the results of testing for a discontinuity for the pooled sampled using the McCrary (2008) test and Frandsen (2014)'s approach for variables with discrete support. In the remaining columns we presents the McCrary (2008) and Frandsen (2014)'s p-values for each AEP wave. These tests do not reject the null hypothesis either yearly or pooling all the years together. As the estimated den-

[21] See Hahn, Todd, and Klaauw (2001); Lee (2008) for an interpretation of the regression discontinuity approach as a local randomization.

[22] The estimated regression functions do not fully saturate the model. Lee and Card (2008) show that one can interpret the deviation between the true conditional expectation function and the estimated regression function as random specification error that introduces a group structure into the standard errors for the estimated treatment effect. Thus, we always report standard errors clustered by test score integer bins.

sities to the left and to the right of the discontinuity overlap, we cannot reject the no discontinuity hypothesis.

In Table 2.5 we provide evidence on the continuity of baseline characteristics around the threshold. We estimate equation (2.5.1) using as outcome variables the characteristics of the teachers and their schools, at the time of application to AEP. The number of the column in this table indicates the order of the piece-wise polynomial of the score used in each specification. In general, there are few statistically significant differences and these differences are small in magnitude with respect to the mean of the variables involved. The second degree order polynomial seems to do best at eliminating individual differences in baseline characteristic between AEP awardees and non-awardees. We cannot reject the null hypothesis of continuity for the 15 variables presented, either using individual tests or a joint (Wald) test. Therefore, we adopt a polynomial of degree two as our benchmark specification. Additionally, we present specifications that control for baseline variables interacted with wave fixed effects.

2.6 Results

2.6.1 Teacher Retention and Labor Supply

We look now at the effect of receiving an AEP award on teacher retention.[23] As we showed in Section 2.4, the presence or absence of a local effect of an AEP award on teacher quitting behavior reflects the difficulty of the test relative to the fix wage. If we observe that the AEP award has no effect on quitting at the threshold, this would suggests that teachers marginally failing to receive the award value their jobs more than their outside option. In contrast, if we do observe an effect in teacher turnover, it is the case that teachers at the threshold were about to quit in the absence of the program.

In Figure 2.7, we summarize the relationship between the AEP aggregate score and teacher turnover, two years after applying for an AEP award. The circles represent the un-adjusted mean of this variable within bins of the score. The superimposed lines are fitted values from a piece-wise linear specification on the score. There is no visual evidence of breaks around the cut-off. In the light of the model presented in Section 2.4, this implies that the marginal teacher who obtains the award is receiving a rent as she was not at risk of quitting, even in the absence of the award.

The program may have affected, however, other margins of labor supply. For example, the AEP award is independent of the hours that the teacher works beyond a minimum of 20 hours. In a static labor supply framework, without restrictions on hours worked, the pure income effect of the award will reduce hours worked. In practice, teachers may have a coarse choice set on the hours they can work and reducing hours may be unfeasible. Yet teachers can adapt their labor supply by adjusting the number of schools they teach at. There are 1,457 teachers working in more than one school at the time of application. If wages across schools are the same and there are some minimal transportation costs, it must be the case that these teachers cannot get enough working hours in one school. In

[23]This is clearly a sharp regression-discontinuity design and therefore receiving the award is the same than scoring above the cut-off. Indeed, looking at Figure 2.5 is not surprising that estimating equation (2.5.1) using the AEP award variable as a dependent variable, pooling all years or year-by-year, we cannot reject the null hypothesis that the cut-off coefficient is equal to one. The results are available upon request from the authors.

such case, a strong income effect may induce teachers to reduce hours of work mainly by providing incentives to drop second jobs.[24]

In Table 2.6, we present OLS estimates of equation (2.5.1) for the total hours worked and teaching at more than one school recorded two years after applying for an AEP award. We focus on the sample of teachers who are in the school system as we found not evidence of an impact of AEP at the extensive margin.[25][26] We show estimates by experience levels using the brackets designated by AEP to determine the size of the bonus: 0-11 years, 12-21 years and 22 or more years of experience.[27] The odd-columns in the table are the results of estimating equation (2.5.1). In the even columns we add controls for demographics, qualifications, labor outcomes, and main school's characteristics at the time of application.

The estimates are very small and mostly non-statistically significant. For example, looking at the sample of all teachers and excluding additional covariates (first column), we find that receiving the award increases the chance of not working in the school system by 0.0038 percentage points (p-value 0.66), reduces hours of work by 0.6344 hours a week (p-value 0.06) and decreases the likelihood of working in more than one school by 0.0047 percentage points (p-value 0.75). Finally, looking by experience levels we can see a stronger fall in hours worked for more experienced teachers of 1.6 hours a week (p-value 0.02) but no other systematic differences.

Figure 2.8 shows estimates of the parameters of interest separately for each of the nine waves. In general, we cannot reject the null hypotheses that all the coefficients are zero.[28] In Figure 2.9 we explore different time windows for the outcomes of interest (the year previous to the program, the year of application to the program, one year after, two years after and three years after). Looking at Figure 2.9, is reassuring that previous to the application and in the year of application for the award there are no effects.[29] Looking at the one year window or the three year window does not change the initial assessment from Table 2.6.

2.6.2 Between-School Mobility

There is no evidence that the program has locally affected teachers' decisions to leave the profession. But, has the award led to any changes in the way teachers sort themselves between schools? Due to the selective nature of the award process, AEP can provide a signal of ability and those receiving the award may use it to improve the overall deal they get from working in the school system. Hence, are teachers changing schools after receiving an award? Who are those teachers? Where are they moving?

[24] The award may also affect the desirability of taking managerial positions within the school system that take teachers outside the classroom. We find no evidence of such effect. The results are available from the authors upon request.

[25] We have also estimated models including zeros in the dependent variable for those who are not working in the school system. The results are similar and are available from the authors upon request

[26] We also separate the sample between female and male teachers but we find no systematic differences in our results. These estimates are available from the authors upon request.

[27] For each column, we estimate equation 2.5.1 in the sub-sample of teacher within the corresponding age range.

[28] Only the coefficient for *Working at more than one school* in the 2004 wave is statistically different from zero.

[29] Administrative data only starts in 2003. Therefore, the information in the years previous to the application for the 2003 applicants is missing

In Figure 2.10, we look for breaks in teachers' mobility. Teachers receiving the award seem to have higher chances of moving to a new school. The first row of Table 2.7 confirms this insight. Two years after receiving the award, teachers are 0.0447 percentage points (p-value 0.01) more likely to move to a new school (first column, first row). With 12 percent of the teachers changing schools every year, the point estimate implies that the AEP award contributes towards more than a 30 percent boost in mobility.

Can we detect any systematic patterns of between school mobility among those that receive the award? We explore this question in rows (2) to (7) of Table 2.7 where we study mobility in the school system independently of the characteristics of the school the teacher is employed when applying to the program. Teachers with 0 to 11 years of experience are 0.0799 percentage points (p-value 0.08) less likely to teach in private-subsidized schools, without necessarily being more prone to teach in private schools (third column). In contrast, awardees with 12 to 21 years of experience are 0.0242 percentage points (p-value 0.00) more likely to teach in private schools, without being less prone to teach in private-subsidized schools (fifth column). This evidence is consistent with the program contributing to equalize wages for municipal and private-subsidized schools at an earlier stage of the teaching career.[30] Finally, there is some evidence that the receiving the award increases the likelihood of working a rural school but no evidence that, independently of initial conditions, teachers are searching for schools located in municipalities where they could receive higher compensation or in schools with better working conditions.

If teachers use the AEP award as a signal of otherwise hard to observe quality, those who were initially working at under-performing schools are more likely to experience higher mobility. In Table 2.8 we look for heterogeneous effects of the AEP on teachers' mobility two years after application. Because teachers preferences for school may vary, Table 2.8 explores alternative definitions of school quality: good working conditions or high student performance. In each column of Table 2.8 we present the OLS estimates of equation (2.5.1), estimated separately for teachers of *bad* schools (i.e., with characteristics below the median) and *good* schools (i.e., with characteristics above the median) at the time of application. As hypothesized, the mobility effect is only present for teachers that were at *bad* schools at the time of application. AEP awardees teaching at schools with working conditions or student performance below the median at the time they applied for the award are 0.0753 (p-value 0.00) and 0.0966 (p-value 0.02) percentage points more likely to be teaching at a different school two years after. This type of mobility, goes against the spirit of the program and may harm disadvantaged schools.[31][32]

[30]Moreover, we can show (results available from the authors upon request) that the probability of applying to AVDI and receiving the AVDI award 1 year after applying to AEP increases. Indeed, the point estimates suggests that teachers receiving the AEP award are 0.0269 percentage points more likely to apply to AVDI (p-value 0.00). The effect is concentrated among teachers in the first decade of their careers, who are 0.0382 percentage points more likely to become AVDI awardees (p-value 0.00). As AVDI is exclusively available for municipal school teachers, this evidence is consistent with the wage equalization occurring at an earlier stage of the teaching career.

[31]Unfortunately, even in the absence of quitting effects, student test-score information is available only for the 4th and 8th grade so it is no possible to analyze the effect of a teacher/school receiving the award on student outcomes.

[32]We find no evidence of systematic movement towards schools in the top fifteeth percentile (results available upon request).

2.6.3 Robustness Checks

In this section we study how robust are our results. First, we look at whether the absence of statistically significant findings is an artifact of low power. In Table B.1, we present power calculations using the β estimates, standard deviations and sample sizes presented in the odd columns of Tables 2.6 and 2.7. From the ten outcomes we consider in only two of them we have a probability of rejecting the null hypothesis if true below the customary 80 percent level.

Second, we analyze the sensitivity of our estimates to alternative bandwidths. In Figure B.1 and Figure B.2, we provide graphical evidence on the effects of the program using the optimal bandwidth obtained with Imbens and Kalyanaraman (2011)'s criteria and a piece-wise linear polynomial of the score.[33] The results are consistent with those of the full window presented in Figures 2.7 and 2.10.

Third, we estimate the effects of an AEP award using a fully non-parametric specification combined with several bandwith sizes. In Figures B.3 and B.4, we plot the estimated impacts of the program for each bandwidth. Our benchmark findings are consistent with this approach. The only statistically significant effect is the increased likelihood that teachers will switch schools after receiving the award.

Fourth, we replicate Table 2.6 and Table 2.7 clustering both at the score-bin and at the school level following Cameron, Gelbach, and Miller (2011). As it can be seen in Tables B.2 and B.3, there is almost no variation in the standard errors.[34]

2.7 Conclusions

Successful public schools systems can retain the best teachers in their classrooms. We analyze the effects on retention and between school mobility of a program that rewards excellence in pedagogical practice in Chile. Teachers apply voluntarily for the award and those who succeed on a set of assessments receive a six percent annual wage increase for up to ten years.

We use a sharp regression discontinuity design to identify the causal effect of receiving an award for primary and secondary school teachers. Using administrative data over eight cohorts of applicants, our estimates indicate that locally the award does not alter transitions out of the school system. This suggests that teachers marginally failing to receive the award value their jobs more than their outside option.

We observe, however, an increase in mobility within the school system among teachers that receive the award. Some of these mobility patterns are consistent with the award providing a signal of teacher ability. For example, movements are concentrated among teachers working at lower performing schools at the time of application.

We also find evidence that teachers in the first decade of their careers that receive the award move out from private-subsidized schools to municipal schools. These movements

[33]Our benchmark estimation uses the entire window size. As a result, the optimal IK bandwidth is always smaller.

[34]Because we are testing for the effects of obtaining the award in 10 different outcome variables, ideally we should implement some error correction method for multiple testing. However, since we only have one outcome variable whose effect is statistically different from zero, any step-up procedure such as Hochberg (1988)'s correction will give the Bonferroni's standard errors for the statistically significant outcome and make all of the other standard errors even larger. As for the Bonferroni's correction, with an un-adjusted p-value of 0.01 (Table 2.7, first column), the significance of the results depends of the choice of the relevant family of outcome variables.

are consistent with the economic incentives that are present in the system.

In sum, the evidence in these chapter suggests that teachers respond to economic incentives but that the design of the program leaves rents to the teachers that marginally pass the assessment. As our model suggests, there is potential room to capture some of these rents by increasing the difficulty of the assessment. The remaining funds can be used to increase the bonus which *ceteris paribus* may prevent those teacher with the highest outside options from leaving the educational system.

Chapter 3

The Effects of Public Recognition of Teaching Excellence on Peers Voluntary Certification

*written jointly with Samuel Berlinski**

3.1 Introduction

In occupations in which low quality of workers impose a high burden on consumers, testing and pay-for-performance can be used to ensure quality standards and incentivize effort. Angrist and Guryan (2004, 2008), however, find that, in the US, states' licensing requirements increase teachers' wages, but do not improve teacher quality.[1] This means that when testing is compulsory, it can reduce labor supply and increase labor costs, without improving quality. An alternative strategy is to make testing voluntary and provide incentive pay on the basis of certification. Under this scheme there are no barriers to entry, and consumers can still acquire the service from a provider with an established quality standard. Yet the voluntary component comes at the risk of having low number of applicants.

Low application rate to voluntary certification programs can be explain by the opposition to individual incentive schemes. In the case of education, teacher unions worldwide strongly oppose performance-based pay (Lavy, 2007). According to Leigh (2013), in Australia only one four of the primary and secondary school teachers agree with merit pay as an strategy to improve teacher retention; in the UK, by 2000 most teachers disagreed with the performance-related pay; in the US, almost every survey finds that teachers oppose to merit pay. Part of this opposition is based on the fact that incentive schemes can introduce horizontal equity concerns, threaten the bargaining power of the union, or erode collaboration among peers (Lavy, 2007). Jones (2013), for example, finds that perfor-

---

*The views expressed herein are those of the authors and should not be attributed to the Inter-American Development Bank, its Executive Directors, or the governments they represent.

[1]See Kane, Rockoff, and Staiger (2008); Harris and Sass (2009); Hanushek and Rivkin (2010) for further evidence.

mance pay reduces collegiality by decreasing teachers' participation in unpaid cooperative activities, even if participation in paid cooperative activities remains unchanged.

In this chapter we analyze peers effects in the decision to apply for voluntary certification programs. Specifically, we explore to what extent being publicly recognized as an excellence teacher affects peers' future application for the same certification. As in Chapter 2, we address the question in the context of Chilean school teachers, who can voluntary apply for a pedagogical excellence award.

As previously described in section 2.2, *Asignación a la Excelencia Pedagógica* (AEP) is a voluntary certification program designed to reward excellence in teaching practice, both economically and socially. In addition to the certification, awardees are entitled with a 6 percent annual salary increase for up to ten years. Certified teachers are also invited to become mentors of other teachers in the *Red Maestro de Maestros*, and their names are announced in a ceremony with local authorities and media coverage. In spite of these incentives, during the first decade of the program, less than six percent of the eligible teachers applied for it, and six out of every ten schools did not have a single applicant.

One can think that part of the low application rate to AEP is due to the fact that teachers mostly dislike merit-pay schemes that rely on standardized tests (Leigh, 2013). AEP, however, evaluates teachers on the grounds of their pedagogical knowledge and teaching skills; not in terms of student performance. To become certified, teachers must prepare a teaching portfolio and take a written test in their main area of expertise. The results of the two assessments are combined in a final score ranging from 100 to 400, and only applicants with a score of at least 275 receive the certification.

We use the quasi-random variation around the 275 cut-off to identify the effects of being certified on peers' future application for the program using a sharp regression discontinuity design. This identification strategy does not suffer the conventional social-interaction identification problems of reflection, endogenous group formation, or correlated effects. Our findings suggests that being publicly recognized as a teacher of excellence through the AEP certification almost doubles peers' application for the program, without necessarily lowering the quality of the applicants.

The rest of the chapter is organized as follows. In Section 3.2 we provide some background on the Chilean education system and the design of the AEP. In Section 3.3 we describe the data. In Sections 3.4 and 3.5 we present our identification strategy and results. Section 3.6 concludes.

3.2 Background

Asignación a la Excelencia Pedagógica (AEP) is a voluntary certification program available for Chilean school teachers. Although the main features of the program have already been described in 2.2; here we present AEP as part of a battery of pay-for-performance incentives introduced in negotiation with the teachers' union. These additional features are key for the understanding of the context where social interactions of teachers take place. We refer to Mizala and Schneider (2014) for a full description of the negotiation process.

Chilean education system

Since the early 1980s, the government of Chile finances education through a nationwide voucher system. Primary and secondary education are provided by *municipal* schools, *private* schools, and *private-subsidized* schools. *Municipal* schools are non-profit institutions, administered by municipalities, that offer instruction for free and receive a per-student subsidy from the Ministry of Education. *Private* schools are for-profit institutions that charge a tuition and receive no subsidies from the government. *Private-subsidized* schools are for-profit institutions, administered by private corporation, that charge a tuition and receive a per-student subsidy from the Ministry of Education. The per-student subsidy received by private-subsidized schools is the same as the one of municipal schools.[2] We refer to municipal and private-subsidized schools as the voucher school system.

Municipal and private-subsidized schools are the bulk of the education system in Chile. In Figure 2.2 we present the share of students enrolled in municipal and private-subsidized schools from 2004 to 2013. We also present the corresponding share of teachers employed by each provider. Adding the enrollment rate for municipal and private-subsidized schools, we observe that the enrollment in the voucher system is stable at around 93 percent. There have been, however, large changes in composition between municipal and private-subsidized schools. While in 2004 50 percent of the students were enrolled in municipal schools, in 2013 the share of enrollment in municipal schools dropped to 40 percent. This has caused a commensurate shift of teachers away from municipal schools.

Each of the three providers offers teachers different contractual arrangements. While the employment of municipal school teachers follows a *Teacher Statute* negotiated by the teacher's union; the private sectors follows the standard labor law, and the employment of private-subsidized school teachers retains some aspects of the both (Mizala and Romaguera, 2005; Santiago et al., 2013). For instance, the minimum wages, bonuses, and maximum working hours in private-subsidized schools follow the same standards as municipal schools. Teachers that attain the retirement age (60 years for women and 65 for men) can work at private-subsidized schools, but cannot at municipal schools.

Incentive Mechanisms

Along with the voucher system, the goverment implemented battery of performance incentives during the late 1990s and early 2000s. The new scheme was gradually introduced through a negotiation process with the teachers' union, and included both collective and individual incentives.

Collective incentives were introduced first. Since 1996 the National System for Performance Evaluation (*Sistema Nacional de Evaluación del Desempeño*) or SNED, following its Spanish acronym, gathers information about all the municipal and private-subsidized schools in the country. Every two years each school is assigned an aggregate performance index.[3] Schools are grouped by students' socioeconomic status, and the index is used to build a within group ranking. The best schools of each group receive a monetary transfer for the next two years. The transfers is distributed among the staff, and for teachers it represents about a 70 percent of a monthly salary (Mizala and Urquiola, 2013).

[2] See Hsieh and Urquiola (2006) for further description of the Chilean school system.

[3] The index combines information on school's performance at the standardized national examination (SIMCE), repetition and dropout rates, educational activities, parental participation, and overall working conditions.

Voluntary individual incentives were introduced next. Since 2002 teachers of municipal and private-subsidized schools, willing to demonstrate their teaching excellence, can apply for the Pedagogical Excellence Award (*Asignación a la Excelencia Pedagógica*) or AEP, following its Spanish acronym. To become certified teachers must prepare a teaching portfolio and take a written test in their main area of expertise. Applicants with outstanding results are entitled with 6 percent annual salary increase for up to ten years, and are socially recognized in a public ceremony. In addition, AEP certified teachers are also invited to become mentors of other teachers in the Network of Teachers of Teachers (*Red Maestro de Maestros*).

In 2004 the goverment attempted to make teacher testing compulsory. Every 4 years, teachers of municipal schools are required to take a Teaching Evaluation (*Evaluación Docente*). Teachers failing the assessment can retake it up to three times, after which they are no longer allowed to work at municipal or private-subsidized schools. The compulsory character the testing, however, faced significance resistance and by 2006, more than 5,000 teachers refused to submit to evaluation Mizala and Schneider (2014).

Parallel to the compulsory examination, *Asignación Variable al Desempeño Individual* or AVDI, following its Spanish acronym, introduced a second round of voluntary individual incentives. Municipals school teachers with an outstanding Teaching Evaluation can apply for AVDI. To apply for the award, teachers must take a written test in their main area of expertise (the same test as the one required for AEP), and the best teachers are entitled with a 5 to 25 percent increase in their per hour wage rate.

There are no incompatibilities between AEP and AVDI. In fact, it is profitable for teachers to apply for the two awards simultaneously as both require the same knowledge test. Yet not many teachers do so. During the 2003-2011 period, less than 6 percent of the eligible teachers applied for the AEP award, and less 12 percent of the applicants applied for both programs at the same time.

We focus the chapter on the AEP for several reasons. First, it was the first voluntary individual incentive to be introduced. Second, it is available for all the teachers of the voucher school system as opposed to AVDI, which is only available for municipal school teachers. Third, it is the only incentive program with a social recognition component.

Asignación a la Excelencia Pedagógica (AEP)

AEP is available for teachers with two or more years of experience, working at least 20 hours a week at municipal or private-subsidized schools. In addition to the certification, awardees are entitled with a 6 percent annual salary increase for up to ten years, and are invited to become mentors of other teachers in the *Red Maestro de Maestros*.[4][5]

The AEP certification diplomas are presented in local ceremonies with media coverage organized by the Regional Ministerial Secretaries of Education (*Secretarías Regionales Ministeriales de Educación*) or SEREMIS, following its Spanish acronym. In 2003, for instance, the ruling president Ricardo Lagos presided a massive ceremony where certified teachers received their distinctive diplomas (Rodriguez et al., 2015). This social recognition component makes the award particularly salient for the education community.

[4] After 2011, the AEP award period was reduced to four years and the magnitude of the monetary bonus was updated. We restrict our analysis to the period before these changes were implemented.

[5] AEP awardees willing to become members of the Network of Teachers of Teachers undergo a separate application process. Around 40 percent of the AEP awardees become members of the Network at some point after receiving the award.

To become certified, teachers must prepare a teaching portfolio and take a written test in their main area of expertise. In the portfolio teachers demonstrate their teaching practices. This assessment requires a learning plan for the students, an evaluation strategy, a pedagogical reflection, and a recording of a class. In the written test teachers are evaluated on grounds of their knowledge. The results of these two assessments are combined in a final score ranging from 100 to 400.[6] Only teachers with a final score of at least 275 become certified. We identify this cut-off by inspecting the data and, to our knowledge, there is no official document where the threshold is stated.[7]

The entire process of AEP begins in April with the diffusion of the program. Throughout the month, printed material is disseminated among schools, and teachers receive e-mails inviting them to apply (Rodriguez et al., 2015). Teachers can enroll on the program from April to May. In June, once the enrollment is closed, teaching portfolios are distributed. These assessments are submitted in October, and the written examination takes place in November. Teachers finally learn about their scores in March, by the beginning of the school year. Every applicant is privately informed about her final score and her performance in the two assessments. At the same time, the names of certified teachers are publicly announced by the SEREMIS. We present this time line in Figure 2.3.

3.3 Data

In this analysis we use administrative data for all teachers in the school system during 2003 and 2014. Starting in 2003, the Ministry of Education gathers information about the place of work, teaching position, and hours worked for every teacher in Chile. This dataset also contains information on basic demographics, qualifications, and experience of teachers. We complement this information with a separate dataset from the Ministry of Education containing the teaching assignment. These are records of all the subjects taught by a teacher during the school year, dis-aggregated at the school, grade, and class level. We match this data with the scores and certification status of the teachers applying for the AEP applicants from 2002 until 2011. We exclude the waves after 2011 as the monetary incentives and assignment ruled changed starting in 2012.

Sample

Between 2002 and 2010, there were 15,122 applications for the AEP.[8] As some teachers apply for the certification more than once, we restrict the analysis to the 13,568 first time applications. After matching these observations with the administrative data on teaching positions, we are left with 11,015 complete records. For 9,224 of these applicants we also have information about their teaching assignment. Finally, because we are interested in peers' behavior, we restrict the sample to 8,937 applicants working in schools with at least 5 teachers.[9]

[6] For the waves taking place between 2002 and 2011, the final score was a weighted average with 70 percent of the weight given to the portfolio and 30 percent to the written test.

[7] The information was confirmed by the *Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógica* (CPEIP) in internal correspondence.

[8] We do not include the 2011 applicants in the sample as our main outcome variable is the peers' application rate, one period ahead.

[9] Of the 9,224 applicants with complete teaching position and teaching assignment records, 99 percent work at schools with 7 or more teachers.

Definitions of Peers

Our main results focus on the sample of 8,937 first time AEP applicants and their peers. Broadly, we define peers as those teachers working at the same school as the applicant, at time she applied for the program.[10] We refer to these teachers as *reference group I*. We further refine the definition of peers and limit it to the teachers working at the same school and teaching at the same level of instruction –primary or secondary– as the applicant; and to the teachers working at the same school and teaching at, least at, one same grade as the applicant, both at time of application. We refer to these teachers as *reference group II* and *reference group III*. Table 3.1 summarizes these concepts. As before, because we are interested in peers' behavior, we require these reference groups to have at least 2 peers.

Our definition of peers or *reference group* is applicant specific. This implies that if two teachers working at the same school apply for the AEP simultaneously, each of them belongs to other's reference group. Table 3.2 clarifies this idea. For the definition of peers according to reference group I and reference group II, there are 8,937 applicant specific reference groups – one for each applicant–. This number reduces when we define peers according to reference group III, because for some applicants there are less than two teachers teaching at, at least, one same grade. Depending on the definition of peers, reference groups can be form by 148,822 teachers working at the same school as the applicants; 132,286 teachers teaching at the same level of instruction as the applicants; or by 91,590 teachers teaching at the same grades as the applicants. Clearly, these numbers decrease as we refine the definition of peers. Likewise, the average group size decreases. While a typical applicant works with other 26 teachers at the same school, and 23 of them teach at her same level of instruction; only 17 teach at her same grades. In each of these reference groups, there are about two applicants, the applicant and one of her peers.

Descriptive Statistics

In the first column of Table 3.3 we present average characteristics for our sample of AEP applicants in our sample. In the second column we present the same information for the teachers working at schools with at least one applicant, i.e. reference group I. Columns three and four present the analogous exercise for teachers of reference group II and reference group III. In column five we describe the same characteristics, but for all the teachers of the voucher school system during 2003-2010 period.

The average AEP applicant is a 42 years old female with a degree in education and 15 years of experience. She works 38 hours a week at a single school and teaches at three different grades. She is two years younger than her peers (irrespective of how we define peers) and works for two additional hours a week. These difference also hold relative to the average teacher of the voucher system.

Around 70 percent of the applicants are primary school teachers, and 44 percent are secondary school teachers.[11] Half of the applicants teach at private-subsidized schools and the remaining half teach at municipals schools. We highlight the fact that, while 37 percent of the teachers applying for the AEP are receiving the collective incentive award,

[10]For teacher working at multiple schools, we restrict to the school with the largest share of hours worked. However, only 10 percent of teachers works at more than one school.

[11]These rates do not sum up to a 100 as 13 percent of the applicants teach at both levels of instruction.

SNED ; only 30 percent of the teachers of the voucher system receive SNED.

In Table 3.4 we present the application, re-application, and certification rates of AEP. From the 220,236 teachers eligible for the program between 2003 and 2010, less than 6 percent applied. One fourth of the applicants were certified in their first attempt, yet their certification rate barely differs from the one of those who applied more than once.[12] There are significant differences, however, in the certification rate over time. While 42 percent of the 2003 applicants were certified, less than 20 percent did so after 2007.

In Figure 3.3 we provide evidence of the assignment rule of the AEP certification in our sample of applicants. The circles present the mean of a variable that takes the value of 1 if a teacher is certified and 0 otherwise. We plot these means against their corresponding score. As described in Section 3.2, there is a sharp discontinuity: certified applicants have an aggregate score of 275 or more.

3.4 Identification Strategy

Our goal is to measure the effect of being publicly recognized as an excellence teacher through the AEP certification on peers' future application for the program. With this aim, we exploit the quasi-random variation of the certification status around the discontinuity cutt-off. In the absence of manipulation, applicants scoring 274 are not much different from those scoring 275; the same argument applies for their peers.[13] As a result if there are systematic differences in peers' behavior after the certifications are granted, they are attributed to the variation of the certification status.

Regression Discontinuity Design

We implement this strategy through the following regression discontinuity design:

$$R_{it}^{t+k} = \alpha + \boldsymbol{\beta} D_{it} + \gamma_t f(s_{it}) + \delta_w D_{iw} \times f(s_{it}) + \lambda_t + \epsilon_{it}^{t+k}. \tag{3.4.1}$$

For a teacher i who applied for AEP at time t, R_{it}^{t+k} is the share of her peers who applied for the certification at time $t+k$. D_{it} is a variable equal to 1 if i scored at least 275 and 0 otherwise, and s_{it} is i's score centered around the 275 cut-off. The function $f(s_{it})$ is a polynomial of the score that varies at both sides of the cutoff; λ_t is a set of wave fixed effects. Our parameter of interest is β.

We compute the peers' application rate, R_{it}^{t+k}, as

$$R_{it}^{t+k} = \frac{\sum_{j \in N_{t+k}} y_{j,t+k} \;\; g_{ji}^t}{\sum_{j \in N_{t+k}} g_{ji}^t \left(1 - \sum_{\tau=1}^{t} y_{j,\tau}\right)}, \tag{3.4.2}$$

where $y_{j,t+k}$ is a variable that takes the values of 1 if teacher j applied for the certification at $t+k$ and 0 otherwise, N_{t+k} denotes the set of the teachers working for more than 20 hours a week in voucher system schools at time $t+k$, and g_{ji}^t is a variable that takes the value of 1 if teacher j and applicant i where peers at time t. As g_{ji}^t depends on the definition of peers, there is a different outcome variable for each definition in Table 3.1.

[12]The magnitude of the AEP bonus varies at four levels of experience: 0-11 years, 12-21 years, 22-30 years, and 31 plus years. Teachers can apply at most twice within each level of experience.

[13]See Hahn et al. (2001) and Lee (2008) for an interpretation of the regression discontinuity approach as a local randomization.

The numerator of equation (3.4.2) is the number of i's peers that applied for AEP at $t+k$. The denominator is the number of applicant i's peers eligible for the program at time $t+k$, minus the number of peers that already applied for the AEP.[14]

Social Interaction Identification Problems

The identification of peer effect through the regression discontinuity design described in equation (3.4.1) does not suffer the conventional social-interaction identification problems pointed by Manski (1993) (see Dahl, Loken, and Mogstad (2014) for detailed discussion). First, as we concentrate on first time applicants, there are no reflection problems. Once a teacher applies for the certification ($y_{i,t} = 1$), her application decision is deterministic ($y_{i,t+k} = 0$ for all $k > 0$). As a result, after a teacher has applied for the program, her score only affects her peers. Second, notice that the contracts and the teaching assignments are determined at the beginning of the school year, prior to the AEP enrollment. Therefore, reference groups are formed before the announcement of the certifications and endogenous group formation is not at stake. Third, as the certification is quasi-randomly assigned, around the cut-off the AEP award is orthogonal to other regressors and correlated unobservables are not a source of bias.

The third point is particularly relevant for applicants who are peers of other applicants. In such case, the applicant specific reference group includes other applicant(s), and their certification status and score are omitted variables in the specification described in equation (3.4.1). However, they are not a source of bias as, conditional on the score, the certification status of the applicant is uncorrelated with the certification status of her peers. In Appendix C.1 we present an example of this situation.

The presence of reference groups with multiple applicants highlights an additional advantage of our identification strategy. Whenever there are multiple applicants in the same reference group, we have a straight forward criteria to select the relevant certification status and score to control for: the one of the applicant defining the reference group.

In principle, we could take $y_{j,t}$ as our outcome variable, but in this case whether we want to control for the score of the first or the second applicant requires an additional criteria. Moreira (2016) suggests that only the applicant closest to the discontinuity cut-off should be included, as the identification of the parameter β comes from the applicant marginally failing or marginally passing the exam.

As shown in Table 3.2, on average, our reference groups have multiple applicants. Therefore, the results in Section 3.5 will account of the above mentioned correction and limit to the observation of the applicant closet to the threshold. We will also present the results for the subset of applicants whose reference group has no other applicant. In the same spirit, we think about the refinements in the definition of peers of reference group II and III as a natural strategy to exclude other applicants.

Validity of the Regression Discontinuity Design

The basic identifying assumption of the regression discontinuity design is that, around the cut-off, there is no systematic manipulation of the score. There are at least two strategies to test the plausibility of this assumption (Bloom, 2012; Hahn et al., 2001; Imbens and Lemieux, 2008; Lee and Card, 2008; Lee and Lemieux, 2010). First, there should be no

[14]For any time t applicant i, $y_{i,t} = 1$. Therefore, we can define a recursive process up the first application wave. For the time being, we abstract from this additional time dimension.

kinks in the density of the score around the discontinuity. Second, predetermined factors ought to vary smoothly around the 275 cut-off.

In Figure 3.4, we plot the histogram of the AEP score for our sample of 8,937 applicants. In column one of Table 3.5, we present the results for the Frandsen (2014)'s and McCrary (2008)'s continuity test. The first one is a test for variables with discrete support, and second one is a test for continuous variables. As the AEP varies from 100 to 400 in steps of one, Frandsen (2014)'s test provides a more accurate assessment for our identification. The remaining columns of Table 3.5, present the corresponding p-values for each application wave. These tests do not reject the null hypothesis either yearly or pooling all the years together. Overall, this suggests that the estimated densities to the left and to the right of the discontinuity overlap, and we cannot reject the no discontinuity hypothesis.[15]

In Table 3.6 we provide evidence on the continuity of baseline characteristics around the threshold. In columns one and two, we estimate equation (3.4.1) using as outcome variables the applicants' characteristics at the time of application to AEP. In columns three to eight we estimate equation (3.4.1) using as outcome variables the average characteristics of the applicants' reference group at the time of application, for each of our three definitions of peers.[16] The numbers of in parenthesis on the top of each column indicate the order of the piecewise polynomial of the score used in each specification.

Once we condition on the score, there are few statistically significant differences between certified an non-certified applicants; these differences are small in magnitude with respect to the means presented in Table 3.3. This is also true for the average characteristics of the reference groups. With a polynomial of degree one we cannot reject the null hypothesis of continuity for the 12 variables presented, neither for the applicants' characteristics or for the reference group average characteristics.[17] As a visual inspection tool, in Figure 3.5 we plot the p-values of the parameter β when estimating equation (3.4.1) on the applicants' characteristics at the time of application and using a degree one polynomial of the score.[18] We adopt the first degree order polynomial as our benchmark specification as it performs best at eliminating differences in baseline characteristic between certified and non-certified teachers.

[15] As observed in Table 3.2, the number of applicants in the sample slightly reduces when we define peers according to the reference group III. In Figures C.1 and Table C.1 we present the analogous exercise to test for the continuity for this sub-sample of applicants.

[16] As the outcome variable of columns three to eight is the average characteristics of the applicants' reference group, the coefficients for school characteristics are exactly the same as in columns one and two. The reason is that a necessary condition for being a peer is to teach at the same school as the applicant, and that there is only one observation per applicant. In Table C.3, we present the same exercise using outcome variables for columns three to eight peers' characteristics (as opposed to average reference group characteristics). Although peers work at the same school as the applicant, now there is one observation per peer instead of one observation peer applicant. As a result, the coefficients in columns one and two can differ from those of in columns three to eight.

[17] In Table C.2 we present the same exercise for the sub-sample of 8,608 applicants with at least two peers, as defined by reference group III.

[18] In Figures C.2, C.3, and C.4 we depict the analogous p-values for the average characteristics of applicants' reference group.

3.5 Results

After establishing the validity of the regression discontinuity design assumptions, we can now look at the effects of the AEP certification status on peers' future application for the program.

Main results

In principle, the effect of being publicly certified as an excellence teacher can have a negative or positive effect on peers' future behavior. On the one hand, incentive schemes make wages more heterogeneous and can introduce horizontal equity concerns, threaten the bargaining power of the union, or erode collaboration among peers (Baker, Jensen, and Murphy, 1988; Lavy, 2007; Jones, 2013). Thereby, it is possible that the public nature of the AEP certification generates aversion among peers. On the other hand, learning about the certification status of the applicant allow peers to update the expected payoff of applying for the program. For instance, Mizala and Schneider (2014) report that teachers do not apply for AEP because they perceive the assessments are rather difficult and time consuming. But if a colleague is certified, it must be the case that the test is not so terribly difficult and the application for the program should increase.

In Figure 3.6 we summarize the relationship between the AEP score and peers' one period ahead application for the program. In this figure we define peers according to reference group I; that is to say, the teachers working at the same school as the applicant. The circles represent the un-adjusted mean of the peers' application rate within bins of the score. The superimposed lines are fitted values from a piecewise linear polynomial of the score. The visual evidence suggests that peers of certified applicants have a higher application rate.

Table 3.7 confirms the insight of Figure 3.6. The odd columns present the OLS estimates of equation (3.4.1) for the peers' application rate, one, two and three periods ahead. In the even columns we control for baseline average reference group characteristics. Columns one and two present the β coefficients the full sample all applicants. Columns three and four correct for multiple applicants in the same reference group by looking exclusively to the one closets to the discontinuity cut-off. Columns five and six present our most conservative specification, where we only include applicants with no other peer simultaneously applying for the program.

Being publicly recognized as a teacher of excellence through the AEP certification increases peers' one period ahead application for the program by 1.4 percentage points. With an average peers' application rate of 2 percent, our point estimates suggest that the certification increases peers' next year application for the program by 70 percent. Two years after the certification are granted, the effect reduces by half, and three years after it disappears.[19]

As an additional strategy to address the multiple applicants' issue, in Table 3.8 and Table 3.9 we present the analogous exercise defining the definition of peers according to reference group II and reference group III. The intuition is that, as we narrow the definition of peers, we naturally exclude other applicants. The results just described remain unchanged.

[19]The existence of an effect of the certification on peer's application for the program two-periods ahead suggests the presence of *snowball* effects. Our current approach does not address this issue and further investigation is required.

Quality of Future Applicants

Given the low number of applicants for the AEP, these results seem promising. Nevertheless, a plausible concern is that the application rate is increased at the cost of reducing the quality of the applicants. In Figure 3.7 we present the relationship between the AEP score and peers' certification rate one period ahead, conditional on application.[20] There is no visual evidence of breaks around the cut-off. Table 3.10 confirms this intuition. If anything, columns three and four suggest that having a certified peer increases the success rate of applicants in the next year; although the sign of the effect reverts two years after application. In Table 3.11 and Table 3.12 we replicate the exercise using reference group II and reference group III.[21] Even if there are some statistically significant effects, the magnitudes are relatively small compared to the 27 percent average certification rate.

Information Channel

Once we have established the existence of a positive effect of being certified as an excellence teacher on peers' future application for the certification, we would like to explore a potential information channel leading to this effect.

A teacher's decision to apply for AEP depends on her beliefs about the probability of passing the exam, the effort she has to devote to prepare the assessments, and the payment she receives if she succeeds. If her beliefs about the probability of being certified are low, she will not apply for the program. In this sense, the certification status can provide a signal of the relative difficulty of the exam, which the teacher can use to update her believes. If this is the case, the closer the interaction between the certified applicant and the teacher, the more informative the signal should be.

To address this hypothesis we replicate the exercise in Table 3.7, refining the definition of peers. In Table 3.13 we present the OLS estimates of equation (3.4.1) for the one period ahead peers' application rate, for each of the three definitions of peers. We refer to these coefficients as β_I, β_{II}, and β_{III}. In the lower panel of the table we present the p-values of a Wald test comparing the coefficients. In five of our six specifications we reject the equality hypothesis of β_I and β_{III} at a 10 percent significance level. The magnitude of the coefficients suggests that, while the AEP certification increases the schools' one period ahead application rate by 70 percent, it almost doubles the application rate of teachers teaching at, at least, one same grade as the applicant. Therefore, it seems to be the case that being publicly recognized as an excellence teacher affects peers' future behavior by allowing them to updated the believes about the difficulty of the assessments, relative to own quality.

Robustness Checks

We endorse the previous results by making a series of robustness checks. First, in Figure 3.8 we plot the effects of certification on peers' one period ahead application rate separately for each of the eight waves, and we reject the null hypotheses that all the coefficients of are jointly zero. Then, we estimate the effects of the AEP certification on

[20]In Table C.4 we replicate this exercise among all peers, coding as 0 the certification outcome of the peers who did not applied.

[21]In Figure C.6 and Figure C.8, we present the graphical evidence for reference group II and reference group III, one period ahead. In Table C.5 and Table C.6 we replicate Table 3.11 and Table 3.12 coding as 0 the certification outcome of the peers who did not applied.

peers' future application for the program using a fully non-parametric specification and several bandwidth sizes. In Figure 3.9, we plot the estimated effects for each bandwidth. Our benchmark findings are also consistent with this approach. In Figure 3.10, we repeat the exercise for the peers' certification rate, conditional on application. Any evidence of negative effects of the AEP certification status on peers' future certification rate is dismissed.

As an additional exercise, we estimate equation (3.4.1) using fake discontinuity thresholds. In Figure 3.11 and Figure 3.12 we plot the histogram of a series β-coefficients – and its corresponding t-statistics – of estimating equation (3.4.1) for peers' one period ahead application rate, using 200-270 and 280-300 as cutoffs for the allocation of the certification. The dashed lines corresponds to the coefficient in Table 3.7, column one, row one. As expected, our results lie at the right tail of the distributions. Figure 3.13 and Figure 3.14 present the same exercise for peers' one period ahead certification rate, conditional on application.

Finally, as a placebo test, we randomly assign applicants to peers. In Figure 3.15 and Figure 3.16 we plot the histogram of a series β-coefficients – its corresponding t-statistics – of estimating equation (3.4.1) for several random draws of peers. Once we assign applicants to reference groups at random, the certification status no longer affects the peers' behavior

3.6 Conclusions

Voluntary certification allows schools to hire teachers with established quality standards and incentivize performance. Yet it can have low number of applicants if teachers oppose to it. In this chapter, we analyze the presence of peer effects on teachers decision to apply for a certification award in Chile. Teachers apply voluntarily and those who succeed on a set of assessments are publicly recognized as teachers of excellence.

We use a sharp regression discontinuity design to identify the causal effect of being publicly recognized as an excellence teacher on peers' future application for the same certification. We define peers as those teachers who, at the time of application, were teaching at the same school, same level of instruction, or at least one same grade as the applicant. Using administrative data over eight cohorts, we are able to follow the peers of the applicant and their application decision over time.

Our estimates indicate that, locally, being AEP certified almost doubles peers' application for the program, without necessarily lowering the quality of the applicants. This result reinforces the success of the negotiation process between the Chilean government and the teachers' union for the introduction of pay-for-performance. If we had found a negative effect on peers' future participation rate, it would be indicate evidence of a double role of the union negotiating the introduction of individual incentives and sabotaging the program after its implementation. In contrast, our findings highlight the advantage of introducing voluntary individual incentives with the approval of the union, as it prepares the ground for the peer effects to take place.

Some of the patterns we observe are also consistent with the award providing a signal of relative teacher quality. Future applicants can use this information to update their believes about the difficulty of the assessments relative to own quality. For example, while the AEP certification increases the schools' application rate by 70 percent, it doubles the application rate of teachers sharing some grade(s).

The absence of negative peer effects and the information channel are particularly relevant for AEP. Given that from 2003 to 2010, six out of every ten schools had no applicant among their teaching staff, and by 2005, 30 percent of the teachers were not aware of the existence of the program; our results suggests that a profitable strategy to increase application is to advertise heavily in schools where there are no former applicants.

Bibliography

Anna Aizer. The gender wage gap and domestic violence. *The American Economic Review*, 100(4), 2010.

Ingvild Almas, Alex Armand, Orazio Attanasio, and Pedro Carneiro. Measuring and changing control: Women's empowerment and targeted transfers. Working Paper 21717, National Bureau of Economic Research, 2015.

Cesar Alonso-Borrego and Raquel Carrasco. Employment and the risk of domestic violence: Does the breadwinner's gender matter? Working papers, Geary Institute, University College Dublin, 2016.

Dan Anderberg and Helmut Rainer. Economic abuse: A theory of intrahousehold sabotage. *Journal of Public Economics*, 97, 2013.

Dan Anderberg, Helmut Rainer, Jonathan Wadsworth, and Tanya Wilson. Unemployment and domestic violence: Theory and evidence. *The Economic Journal*, 2015. ISSN 1468-0297.

Siwan Anderson and Garance Genicot. Suicide and property rights in india. Working Paper 19978, National Bureau of Economic Research, 2014.

Manuela Angelucci. Love on the rocks: Domestic violence and alcohol abuse in rural mexico. *The B.E. Journal of Economic Analysis & Policy*, 8(1), 2008.

Joshua Angrist and Jonathan Guryan. Teacher testing, teacher education, and teacher characteristics. *American Economic Review*, 94(2):241–246, 2004.

Joshua Angrist and Jonathan Guryan. Does teacher testing raise teacher quality? evidence from state certification requirements. *Economics of Education Review*, 27(5): 483–503, 2008.

Carolina Araya-Ramirez, Sandy Taut, Verónica Santelices-Etchegaray, Jorge Manzi-Astudillo, and Fanisa Mino-Flores. The theory underlying a certification of teaching excellence program in chile. *Revista de Educación*, 2012.

Nava Ashraf. Spousal control and intra-household decision making: An experimental study in the philippines. *American Economic Review*, 99(4), 2009.

Orazio Attanasio and Valerie Lechene. Efficient responses to targeted cash transfers. *Journal of Political Economy*, 122(1), 2014.

George Baker, Michael Jensen, and Kevin Murphy. Compensation and incentives: Practice vs. theory. *The Journal of Finance*, 43(3):593–616, 1988.

Kaushik Basu. Gender and say: a model of household behaviour with endogenously determined balance of power. *The Economic Journal*, 116(511), 2006.

Marianne Bertrand, Emir Kamenica, and Jessica Pan. Gender identity and relative income within households. *The Quarterly Journal of Economics*, 2015.

Francis Bloch and Vijayendra Rao. Terror as a bargaining instrument: A case study of dowry violence in rural india. *American Economic Review*, 92(4), 2002.

Howard Bloom. Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, 5(1):43–82, 2012.

Gustavo Bobonis, Melissa González-Brenes, and Roberto Castro. Public transfers and domestic violence: The roles of private information and spousal control. *American Economic Journal: Economic Policy*, 5(1), 2013.

Audra Bowlus and Shannon Seitz. Domestic violence, employment and divorce. *International Economic Review*, 47(4), 2006.

Martin Browning, Pierre-André Chiappori, and Yoram Weiss. *Economics of the Family*. Cambridge University Press, 2014.

Colin Cameron, Jonah Gelbach, and Douglas Miller. Robust inference with multiway clustering. *Journal of Business and Economic Statistics*, 29(2):238–249, 2011.

David Card and Gordon Dahl. Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1), 2011.

Pierre-André Chiappori and Maurizio Mazzocco. Static and intertemporal household decisions. *Journal of Economic Literature*, 2016.

Yoo-Mi Chin. Male backlash, bargaining, or exposure reduction?: Women's working status and physical spousal violence in india. *Journal of Population Economics*, 25(1), 2012.

Charles Clotfelter, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor. Would higher salaries keep teachers in high-poverty schools? evidence from a policy intervention in north carolina. *Journal of Public Economics*, 92(5-6):1352–1370, 2008.

Jesse Cunha. Testing paternalism: Cash versus in-kind transfers. *American Economic Journal: Applied Economics*, 6(2), April 2014.

Jesse Cunha, Giacomo De Giorgi, and Seema Jayachandran. The price effects of cash versus in-kind transfers. Technical Report 375, Federal Reserve Bank of New York, 2015.

Janet Currie and Firouz Gahvari. Transfers in cash and in-kind: Theory meets the data. *Journal of Economic Literature*, 46(2), 2008.

Gordon B. Dahl, Katrine V. Loken, and Magne Mogstad. Peer effects in program participation. *American Economic Review*, 104(7):2049–74, 2014.

Giacomo De Giorgi, Michele Pellizzari, and Silvia Redaelli. Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2), April 2010.

Angus Deaton. Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48(2):424–55, June 2010.

Angus Deaton and Nancy Cartwright. Understanding and misunderstanding randomized controlled trials. Working Paper Series 22595, National Bureau of Economic Research, 2016.

Matthias Doepke and Michele Tertilt. Does female empowerment promote economic development? Discussion Papers 8441, CEPR, 2011.

Peter Dolton and Wilbert Van der Klaauw. Leaving teaching in the uk: A duration analysis. *The Economic Journal*, 105(429):431–444, 1995.

Peter Dolton and Wilbert Van der Klaauw. The turnover of teachers: A competing risks explanation. *The Review of Economics and Statistics*, 81(3):543–550, 1999.

Esther Duflo. Women empowerment and economic development. *Journal of Economic Literature*, 50(4), 2012.

Mary Ellsberg and Lori Heise. Researching violence against women: A practical guide for researchers and activists. World health organization, Washington, D.C., 2005.

Mukesh Eswaran and Nisha Malhotra. Domestic violence and women's autonomy in developing countries: Theory and evidence. *Canadian Journal of Economics*, 44(4), 2011.

Torberg Falch. Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment. *American Economic Review*, 101(3):460–65, 2011.

Amy Farmer and Jill Tiefenthaler. An economic analysis of domestic violence. *Review of Social Economy*, 55(3), 1997.

Brigham R. Frandsen. Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. Manuscript, 2014.

Richard Gelles. Violence in the family: A review of research in the seventies. *Journal of Marriage and Family*, 42(4), 1980.

Luis Guasch and Andrew Weiss. Wages as sorting mechanisms in competitive markets with asymmetric information: A theory of testing. *The Review of Economic Studies*, 47(4):653–664, 1980.

Luis Guasch and Andrew Weiss. Self-selection in the labor market. *The American Economic Review*, 71(3):275–284, 1981.

Jinyong Hahn, Petra Todd, and Wilbert Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

Eric Hanushek and Steven Rivkin. The quality and distribution of teachers under the no child left behind act. *Journal of Economic Perspectives*, 24(3):133–50, 2010.

Eric Hanushek, John Kain, and Steven Rivkin. Why public schools lose teachers. *The Journal of Human Resources*, 39(2):326–354, 2004.

Douglas Harris and Tim Sass. The effects of nbpts-certified teachers on student achievement. *Journal of Policy Analysis and Management*, 28(1):55–80, 2009.

James J. Heckman. Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature*, 48(2):356–98, June 2010.

Melissa Hidrobo and Lia Fernald. Cash transfers and domestic violence. *Journal of Health Economics*, 32(1), 2013.

Melissa Hidrobo, John Hoddinott, Amber Peterman, Amy Margolies, and Vanessa Moreira. Cash, food, or vouchers? evidence from a randomized experiment in northern ecuador. *Journal of Development Economics*, 107, 2014.

Melissa Hidrobo, Amber Peterman, and Lori Heise. The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in northern ecuador. *American Economic Journal: Applied Economics*, 2016.

Yosef Hochberg. A sharper bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4):800–802, 1988.

Chang-Tai Hsieh and Miguel Urquiola. The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of Public Economics*, 90(8-9):1477–1503, 2006.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 2011.

Guido Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635, 2008.

Instituto Nacional de Estadística y Censos. Encuesta nacional de relaciones familiares y violencia de genero contra las mujeres, 2011.

Murat Iyigun and Randall Walsh. Endogenous gender power, household labor supply and the demographic transition. *Journal of Development Economics*, 82(1):138–155, 2007.

Michael Jones. Teacher behavior under performance pay incentives. *Economics of Education Review*, 37:148–164, 2013.

Thomas Kane, Jonah Rockoff, and Douglas Staiger. What does certification tell us about teacher effectiveness? evidence from new york city. *Economics of Education Review*, 27(6):615–631, 2008.

Morris Kleiner. Occupational licensing. *Journal of Economic Perspectives*, 14(4):189–202, 2000.

Victor Lavy. Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy*, 110:1286–1317, 2002.

Victor Lavy. Using performance-based pay to improve the quality of teachers. *The Future of Children*, 17(1):87–109, 2007.

Victor Lavy. Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, 99(5):1979–2011, 2009.

David Lee. Randomized experiments from non-random selection in u.s. house elections. *Journal of Econometrics*, 142(2):675–697, 2008.

David Lee and David Card. Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674, 2008.

David Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, 2010.

Andrew Leigh. The economics and politics of teacher merit pay. *CESifo Economic Studies*, 59(1), 2013.

Charles Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–42, 1993.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

David McKenzie. Beyond baseline and follow-up: The case for more t in experiments. *Journal of Development Economics*, 99(2), 2012.

Alejandra Mizala and Pilar Romaguera. *Incentives to Improve Teaching: lessons from Latin America*, chapter Teacher's Salary Structure and Incentives in Chile, pages 130–150. The World Bank, 2005.

Alejandra Mizala and Ben Ross Schneider. Negotiating education reform: Teacher evaluations and incentives in chile (1990-2010). *Governance*, 27(1):87–109, 2014.

Alejandra Mizala and Miguel Urquiola. School markets: The impact of information approximating schools' effectiveness. *Journal of Development Economics*, 103:313–335, 2013.

Diana Moreira. Success spills over: How awards affect winners' and peers' performance in brazil. 2016.

Karthik Muralidharan and Venkatesh Sundararaman. Teacher performance pay: Experimental evidence from india. *Journal of Political Economy*, 119(1):39–77, 2011.

OECD. Education at a glance 2013, 2013.

Canice Prendergast. The provision of incentives in firms. *Journal of Economic Literature*, 37(1):7–63, 1999.

Beatriz Rodriguez, Jorge Manzi, Claudia Peirano, Roberto Gonzalez, and David Bravo. *Reconocimiento al merito docente AEP 2002-2014*. Centro UC Medicion-MIDE, 2015.

Paulo Santiago, Francisco Benavides, Charlotte Danielson, Laura Goe, and Deborah Nusche. Teacher evaluation in chile. Technical report, OECD, 2013.

Peter Schochet. Statistical power for regression discontinuity designs in education evaluations. *Journal of Educational and Behavioral Statistics*, 34:238–266, 2009.

WHO Department of Reproductive Health and Research, London School of Hygiene and Tropical Medicine, and South African Medical Research Council. Global and regional estimates of violence against women: Prevalence and health effects of intimate partner violence and non- partner sexual violence. Wold health organization, Geneva, Switzerland, 2013.

Tables and Figures for Chapter 1

Figures

Figure 1.1: In-Kind versus Cash Transfers

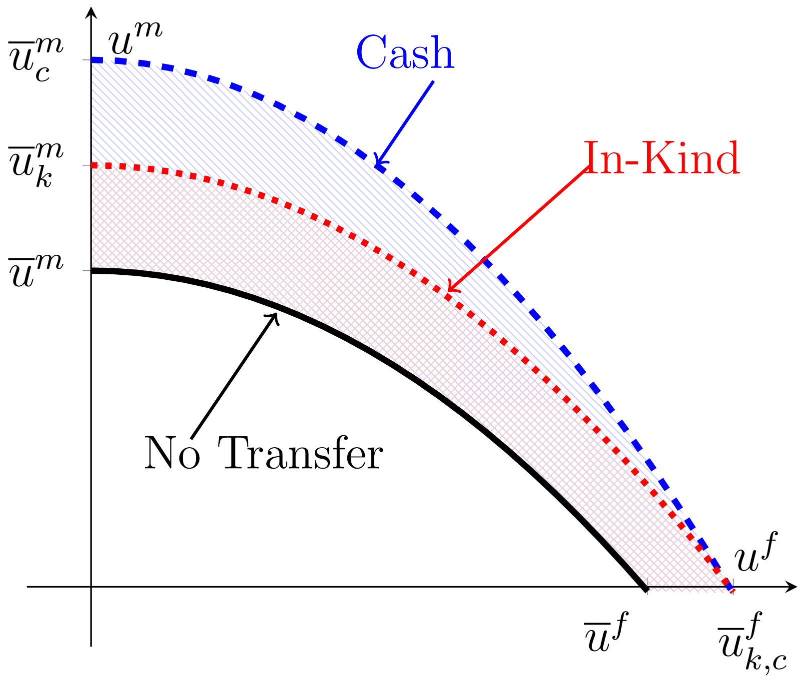
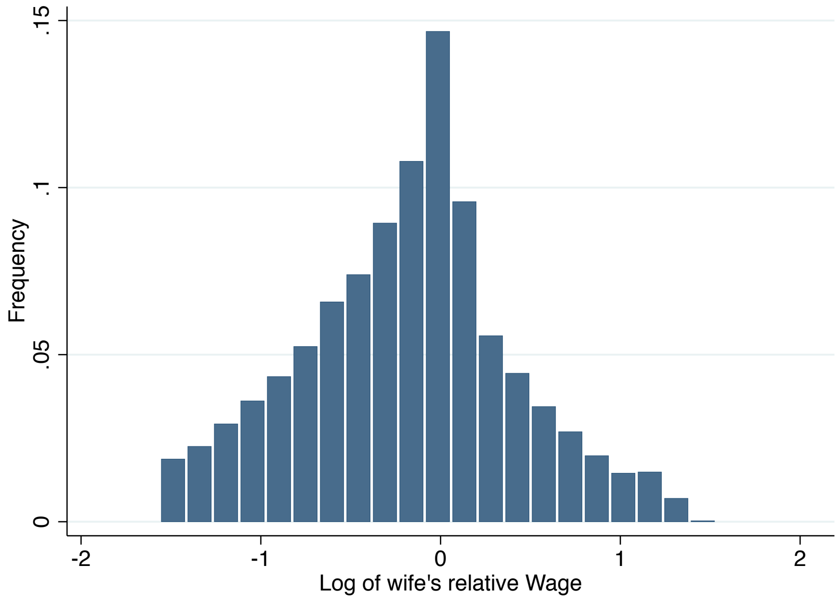
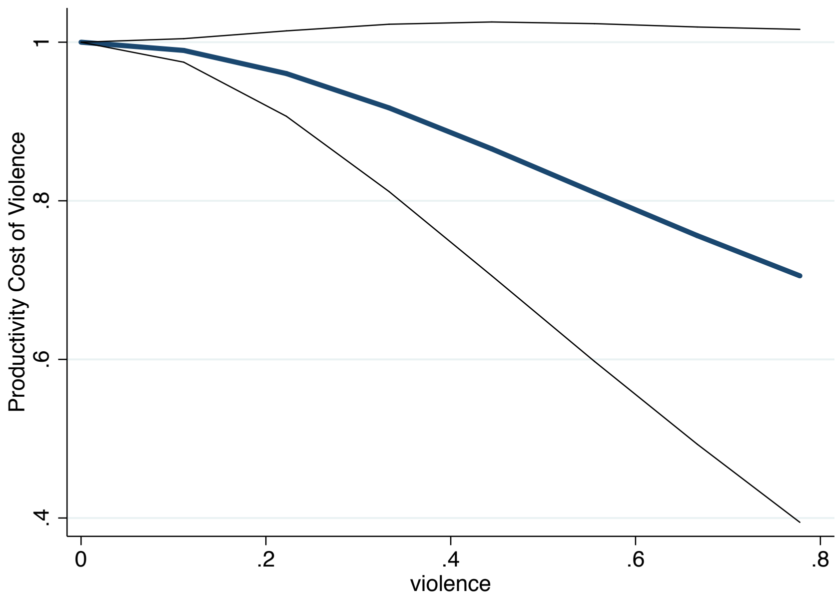


Figure 1.2: Distribution of Female Relative Wage



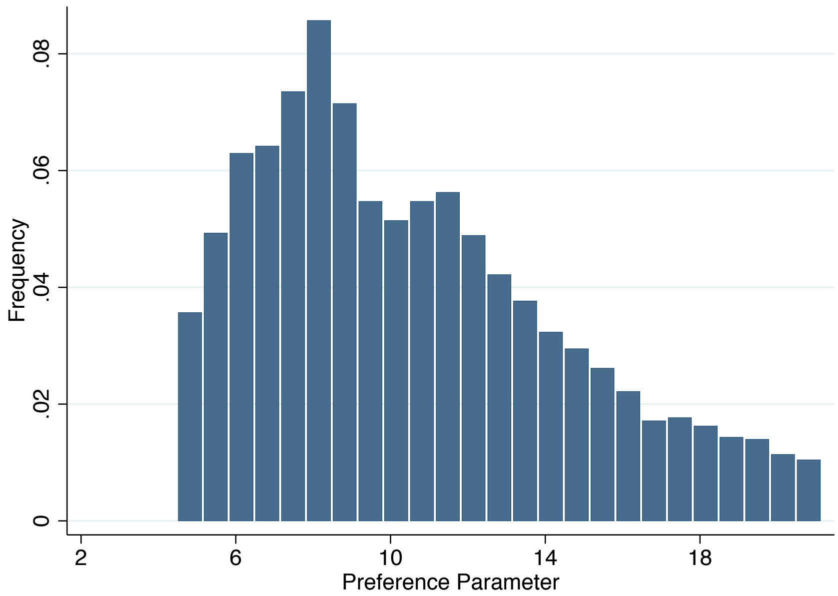
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Distribution of female relative per hour wage rate, in logarithms. The distribution comes from a bootstrap in which, for each sample draw we estimate the Heckman Two-Step procedure, split the predicted wages in bins of equal size, and compute the median and the frequency from each bin. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Figure 1.3: Productivity Cost of Violence



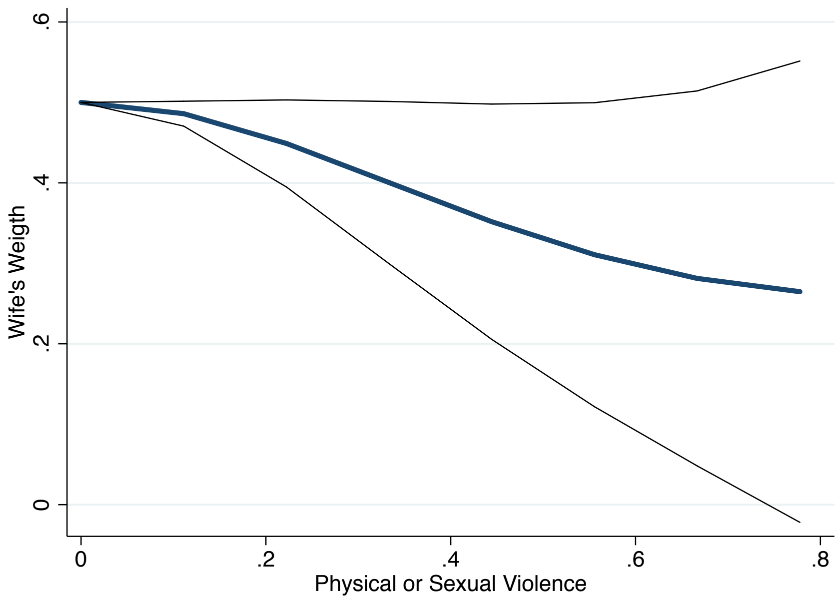
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Bootstrapped confidence intervals. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Figure 1.4: Distribution of Disagreement in Preference Parameter



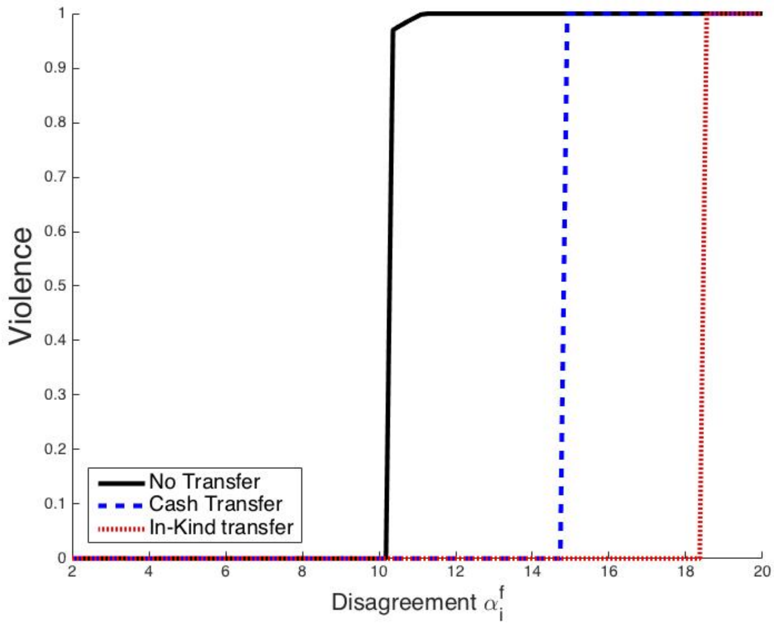
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: The distribution comes from a bootstrap in which, for each sample draw we estimate Equation (1.5.5), split the predicted α_i^fs in bins of equal size, and compute the median and the frequency from each bin. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Figure 1.5: Violence and Balance of Power



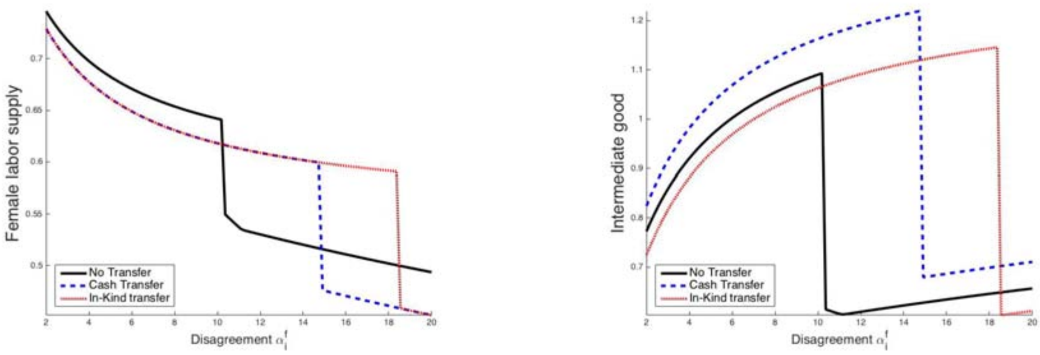
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Bootstrapped confidence intervals. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Figure 1.6: Effect of In-kind versus Cash Transfers on Violence



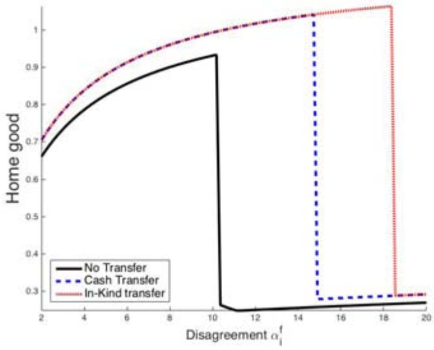
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: The black solid line represents the no transfer scenario, the red dotted line represents an in-kind transfer, and the blue dashed line represents a cash transfer. Transfers are equivalent to 10% of the male income. Relative wages and no-violence female relative weight set to $w_f = 0.5$ and $\mu(0, \widetilde{\omega}_f) = 0.5$, respectively.

Figure 1.7: Effect of In-kind versus Cash Transfers in the Inputs of Home Production



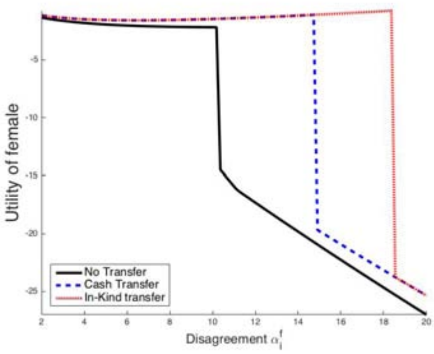
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: The black solid line represents the no transfer scenario, the red dotted line represents an in-kind transfer, and the blue dashed line represents a cash transfer. The left panel depicts the effect of In-kind versus Cash in the female labor supply. The right panel depicts the effect of In-kind versus Cash in the market inputs demand. Transfers are equivalent to 10% of the male income. Relative wages and no-violence female relative weight set to $w_f = 0.5$ and $\mu(0, \widetilde{\omega}_f) = 0.5$, respectively.

Figure 1.8: Effect of In-kind versus Cash Transfers on the Home Good and the market Good



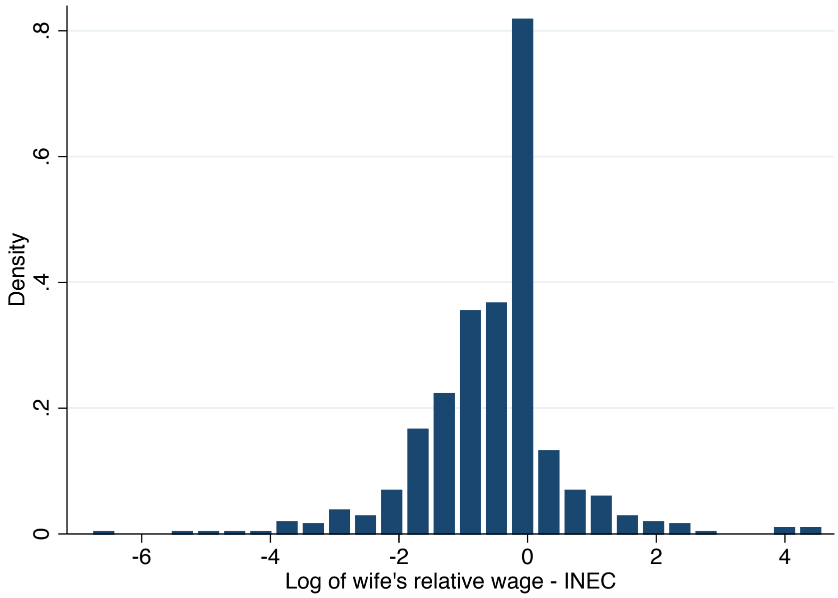
Source: *Food, Cash or Voucher* (World Food Programme).
Notes: The black solid line represents the no transfer scenario, the red dotted line represents an in-kind transfer, and the blue dashed line represents a cash transfer. The left panel depicts the effect of In-kind versus Cash in the home good. The right panel depicts the effect of In-kind versus Cash in the market good. Transfers are equivalent to 10% of the male income. Relative wages and no-violence female relative weight set to $w_f = 0.5$ and $\mu(0, \widetilde{\omega}_f) = 0.5$, respectively.

Figure 1.9: *Food, Cash or Voucher* (World Food Programme)



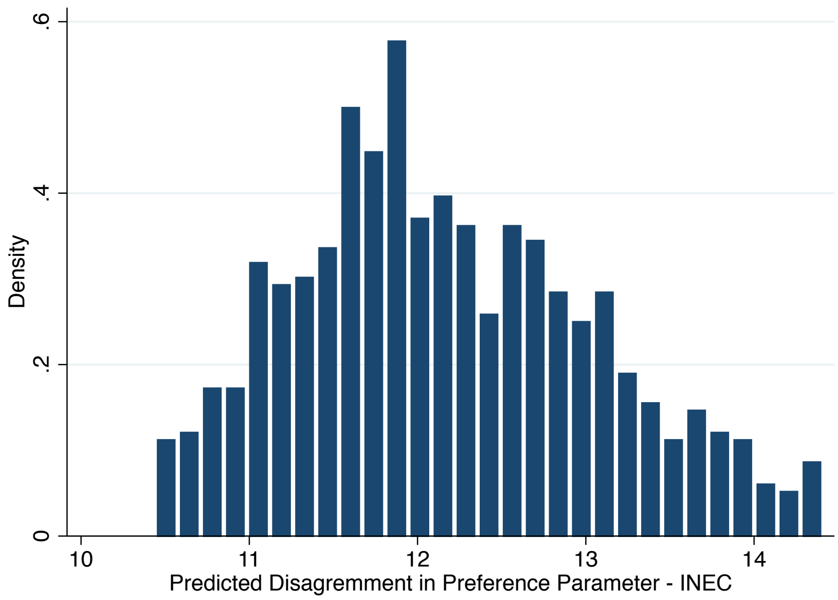
Source: Effect of In-kind versus Cash Transfers on the Utility of the Female and the Male.
Notes: The left panel depicts the effect of In-kind versus Cash transfers on the utility of the female. The right panel depicts the effect of In-kind versus Cash on the utility of the male. Transfers are equivalent to 10% of the male income. Relative wages and no-violence female relative weight set to $w_f = 0.5$ and $\mu(0, \widetilde{\omega}_f) = 0.5$, respectively.

Figure 1.10: Distribution of Female Relative Wage
Bono de Desarollo Humano



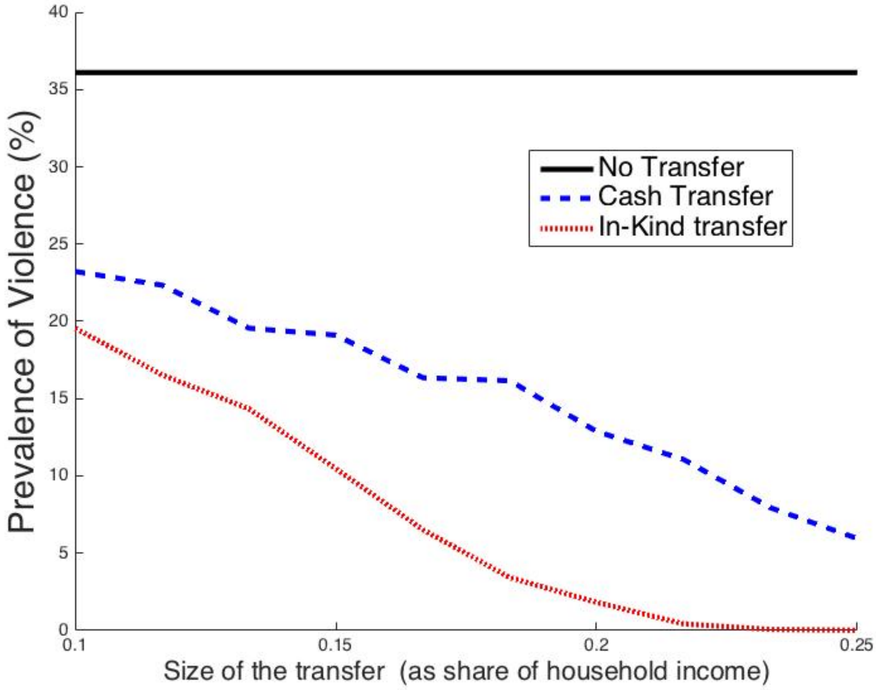
Source: *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres.*
Notes: Empirical distribution of female relative per hour wage rate, in logarithms.

Figure 1.11: Distribution of Predicted Disagreement in Preferences
Bono de Desarollo Humano



Source: *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres.*
Notes: For each of the households in the sub-sample of the *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres*, we predict α_i^f using the observable household characteristics and the coefficients in Table 1.8.

Figure 1.12: Prevalence of Violence for Different Size of Transfers



Source: *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres.*
Notes: The black solid line represents the no transfer scenario, the red dotted line represents an in-kind transfer, and the blue dashed line represents a cash transfer. Transfers as share of average household income.

Tables

Table 1.1: Descriptive Statistics

	All	Control	Treatment		p-value	
			In-Kind	Cash	Control vs. Treatment	In-kind vs. Cash
Panel A. Demographics						
No. of household members	5.37	5.58	5.26	5.37	0.01	0.52
Male head of household	0.97	0.97	0.97	0.98	0.41	0.33
Married couple	0.42	0.42	0.43	0.41	0.72	0.79
No. children form 0 to 5	0.75	0.72	0.78	0.73	0.29	0.41
No. children from 6 to 14	0.92	1.02	0.85	0.92	0.04	0.48
Male age	38.67	39.27	38.54	38.21	0.32	0.72
Female Age	34.81	35.29	34.44	34.98	0.38	0.60
Couple age difference	3.35	3.48	3.58	2.76	0.72	0.07
Male education years	8.03	7.81	8.17	8.01	0.40	0.70
Female education years	8.02	7.75	8.22	7.96	0.32	0.51
Female more educated than male	0.18	0.20	0.17	0.19	0.30	0.44
Panel B. Intimate Partner Violence						
Any type of violence	0.29	0.27	0.32	0.28	0.74	0.25
Physical Violence	0.15	0.12	0.17	0.15	0.82	0.84
Physical or sexual violence	0.16	0.12	0.18	0.16	0.82	0.75
Sexual violence	0.03	0.02	0.04	0.04	0.63	0.88
Emotional violence	0.26	0.24	0.28	0.24	0.80	0.23
Panel C. Variables for the Estimation						
Index of physical or sexual violence	0.25	0.26	0.25	0.25	0.62	0.64
Household day expenses in food	3.96	3.88	3.94	4.09	0.22	0.09
Household income a day	14.00	14.87	13.65	13.69	0.19	0.92
Female employed	0.32	0.31	0.32	0.34	0.57	0.69
Female labor income a day	6.55	7.36	6.17	6.41	0.06	0.95
Female hours of work a day	5.21	5.68	4.93	5.25	0.10	0.30
Female per hour wage	1.74	1.83	1.70	1.71	0.12	0.99
Female hours of household work a day	7.30	7.52	7.22	7.18	0.57	0.80
Male employed	0.96	0.96	0.96	0.97	0.52	0.60
Male labor income a day	12.40	13.14	12.22	11.92	0.28	0.78
Male hours of work a day	6.91	7.04	6.75	7.08	0.74	0.54
Male per hour wage	1.98	2.06	1.99	1.88	0.16	0.67

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: All variables at baseline. Columns (5) and (6) present the p-value of a Wald test, with clustered standard errors at the cluster level.

Table 1.2: Impact of *Food, Cash or Voucher* on Domestic Violence

	Full Sample Mean at baseline =0.16				Working Females Mean at baseline 0=.17			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any transfer	-0.061*	-0.062*			-0.097*	-0.121**		
	(0.033)	(0.032)			(0.050)	(0.051)		
Cash			-0.052	-0.051			-0.066	-0.094*
			(0.037)	(0.035)			(0.057)	(0.056)
In-Kind			-0.066*	-0.068**			-0.114**	-0.136**
			(0.035)	(0.033)			(0.052)	(0.054)
p-value: In-Kind vs. Cash			0.57	0.50			0.24	0.30
Clusters	145	145	145	145	128	128	128	128
N	1,230	1,230	1,230	1,230	395	395	395	395

Source: *Food, Cash, or Voucher* (World Food Programme).
Notes: Clustered standard errors at the cluster level in parenthesis. Columns (2) and (4) control for number of members in the household, sex of the head of the household, marriage or cohabitation, number of children aged 0 to 5, number of children aged 6 to 14, female's age, partner's age, age difference, partner's years of education, female's years of education, female more educated than her partner, female's employment status, female's labor income a day, female's hours of work a day, female's hours of household work a day, partner's employment status a day, partner's labor income a day, partner's hours of work a day, and household food consumption a day, all at baseline.
Indicates statistical significance at 10%.
* Indicates statistical significance at 5%.
** Indicates statistical significance at 1%.

Table 1.3: Estimated Parameters

Parameter	Estimates
θ	0.86
	(0.02)
β_2	-0.853
	(1.24)
δ	3.05
	(0.32)

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Bootstrapped standard errors in parenthesis. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Table 1.4: Prevalence of Violence

	Prevalence of Violence
No Transfer	17.63 %
Food, Cash, or Voucher	8.23%

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Simulations based in the parameters of Table 1.3, the estimated distribution of α_i^fs depicted in Figure 1.4 and the empirical distribution of female relative wage presented in Figure 1.2.

Table 1.5: Predicted Rates of Violence under Different Transfer Regimes

	Prevalence of Violence
No Transfer	17.63 %
Cash transfers (only)	9.86 %
In-kind transfers (only)	7.41 %

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Simulations based in the parameters of Table 1.3, the estimated distribution of α_i^fs depicted in Figure 1.4 and the empirical distribution of female relative wage presented in Figure 1.2.

Table 1.6: Predicted Rates of Violence under Different Transfer Regimes
National Level

	Prevalence of Violence
No Transfer	36.10%
Cash transfers (only)	23.23%
In-kind transfers (only)	19.40%

Source: *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres* (INEC).
Notes: Simulations based in the parameters of Table 1.3, the predicted distribution of α_i^fs depicted in Figure 1.11 and the empirical distribution of female relative wage presented in Figure 1.10.

Table 1.7: Descriptive Statistics of the Beneficiaries of *Bono de Desarollo Humano*

Variable	Mean
No. of household members	4.91
	(2.05)
Male head of household	0.97
	(0.17)
Married couple	0.64
	(0.48)
No. children form 0 to 5	0.73
	(0.90)
No. children from 6 to 14	1.37
	(1.30)
Male age	44.76
	(15.24)
Female Age	40.97
	(13.42)
Couple age difference	3.78
	(8.19)
Male education years	4.23
	(2.19)
Female education years	4.08
	(2.26)
Female more educated than male	0.22
	(0.41)
Female employed	0.32
	(0.47)
Female labor income a day	5.12
	(17.82)
Female hours of work a day	8.50
	(3.05)
Female hours of household work a day	8.51
	(3.05)
Male employed	0.91
	(0.29)
Male labor income a day	8.34
	(7.56)
Male hours of work a day	8.49
	(3.04)
Physical Violence	0.37
	(0.48)

Source: *Encuesta Nacional sobre Relaciones Familiares y Violencia de Género contra las Mujeres* (INEC).
Notes: Standard deviations in parenthesis.

Table 1.8: Household Observable Characteristics Predicting Disagreement in Preferences

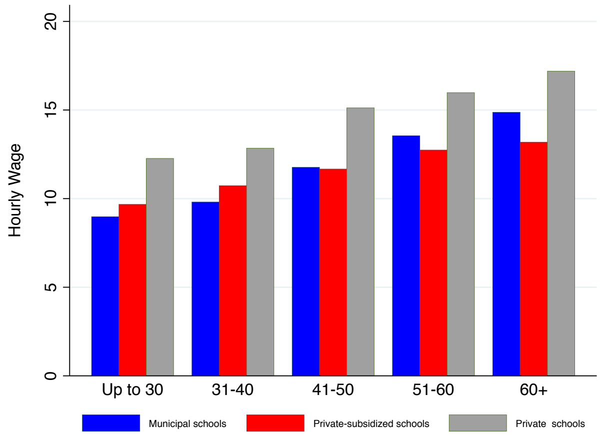
Observable Characteristic	α_i^f
No. of household members	0.53
	(0.34)
Male head of household	3.26
	(2.45)
Married couple	0.16
	(0.83)
No. children form 0 to 5	-0.09
	(0.54)
No. children from 6 to 14	-0.42
	(0.48)
Female age	0.04
	(0.04)
Couple age difference	0.06
	(0.06)
Male education years	0.02
	(0.14)
Female education years	-0.09
	(0.16)
Female more educated than male	-0.07
	(1.13)
Female employed	0.92
	(1.14)
Female labor income a day	0.21
	(0.16)
Female hours of work a day	-0.04
	(0.19)
Female hours of household work a day	-0.08
	(0.11)
Male employed	1.85
	(3.19)
Male labor income a day	0.11
	(0.08)
Male hours of work a day	0.08
	(0.16)

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Bootstrapped standard errors in parenthesis. For each sample draw we estimate Equation (1.5.5) and regress it on household observable characteristics. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Tables and Figures for Chapter 2

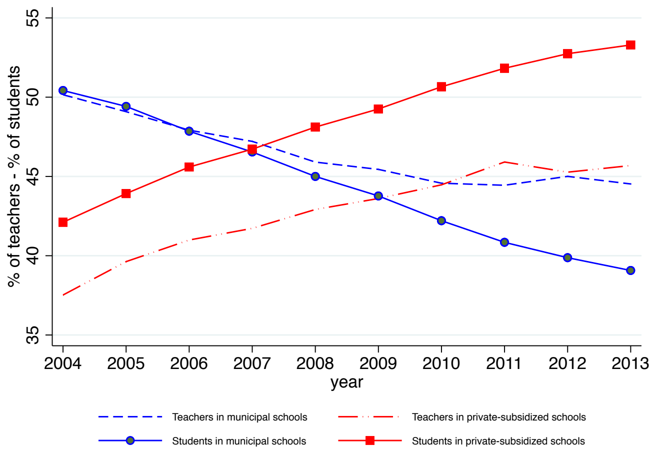
Figures

Figure 2.1: Average hourly wage by age group and type of school in 2005.



Source: *Encuesta Longitudinal Docente 2005: Análisis y Principales Resultados.*
Notes: Wages in USD of 2005

Figure 2.2: Student enrollment and teachers employed in municipal and private-subsidized schools



Source: Own calculations based on data from the Ministry of Education (Chile)

Figure 2.3: Timeline

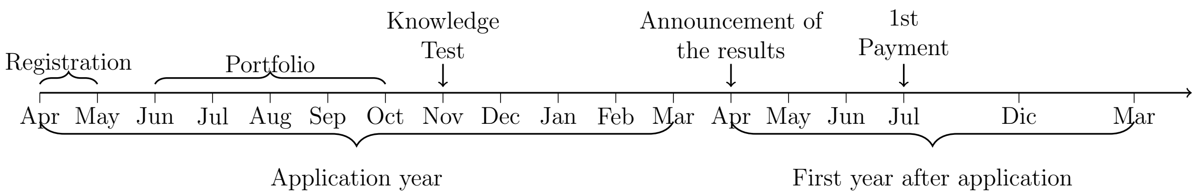


Figure 2.4: Flowchart for AEP sample

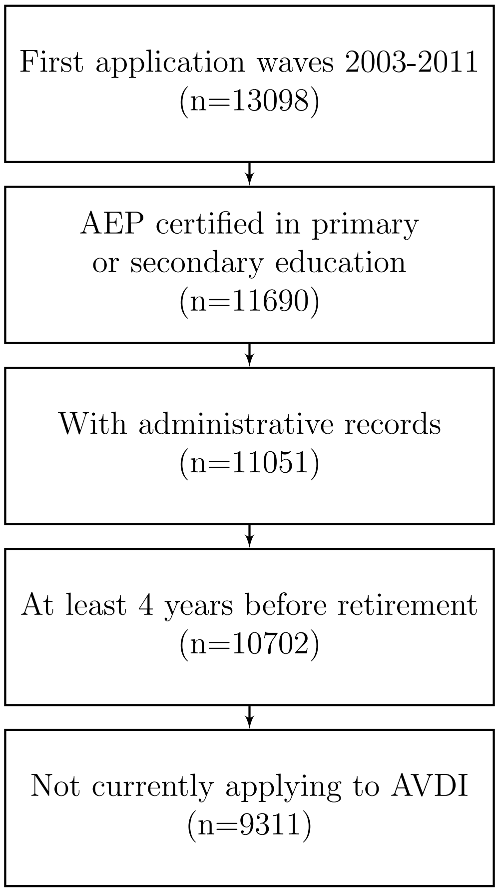
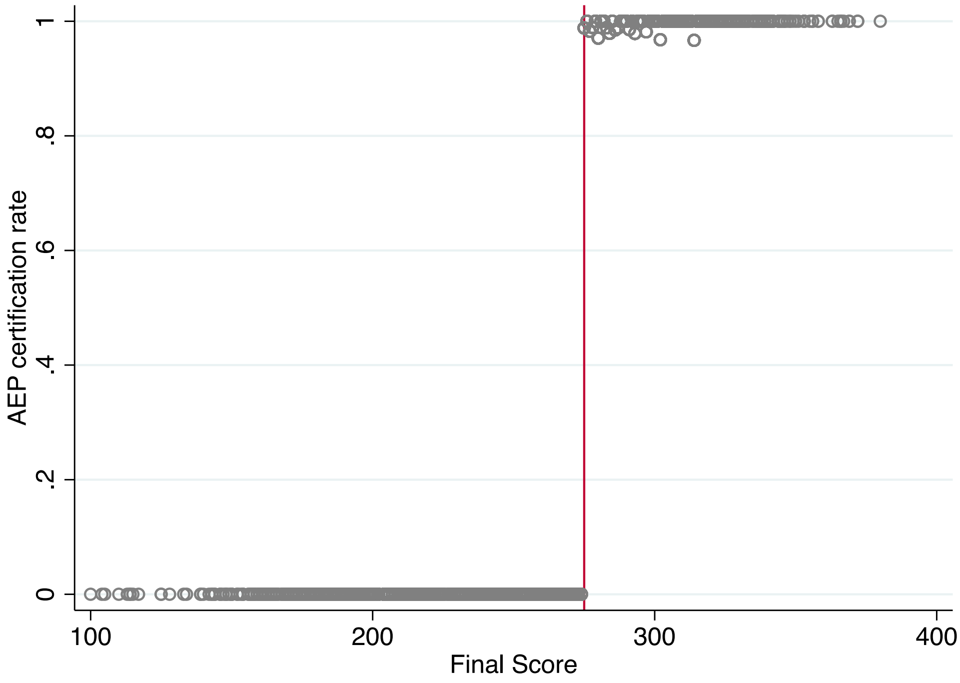
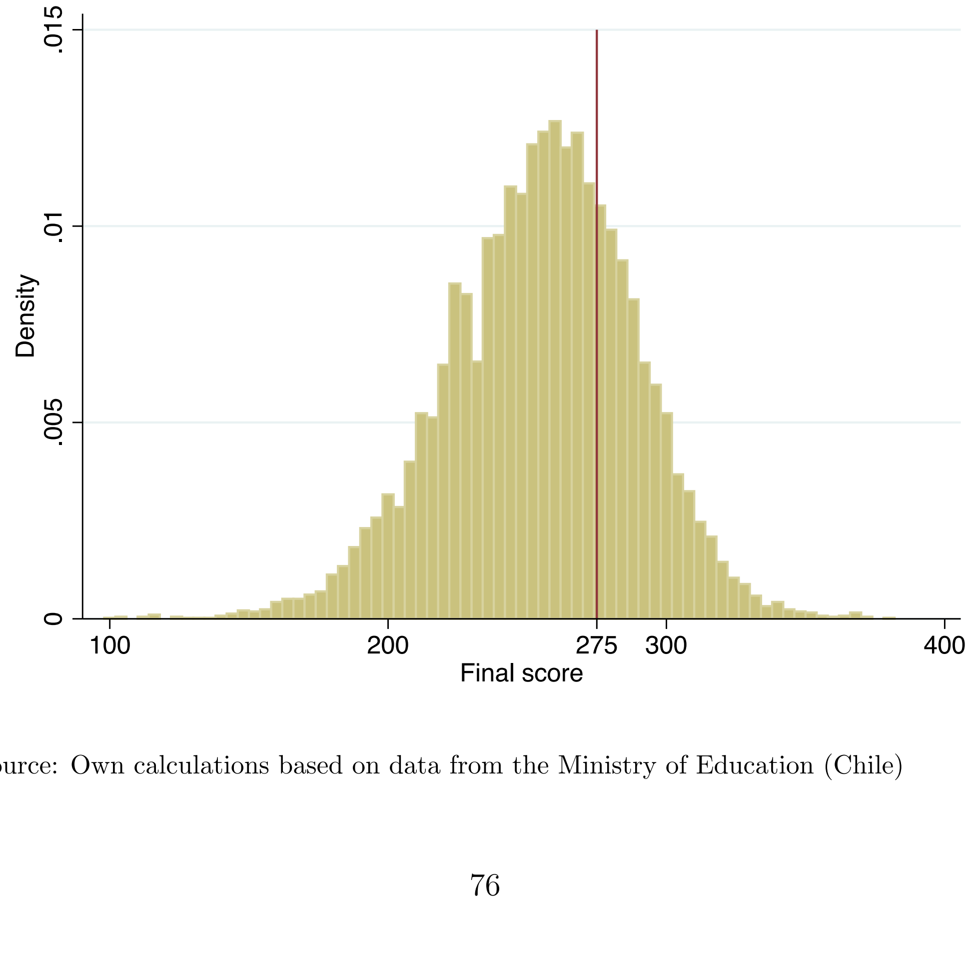


Figure 2.5: AEP Assignment Rule



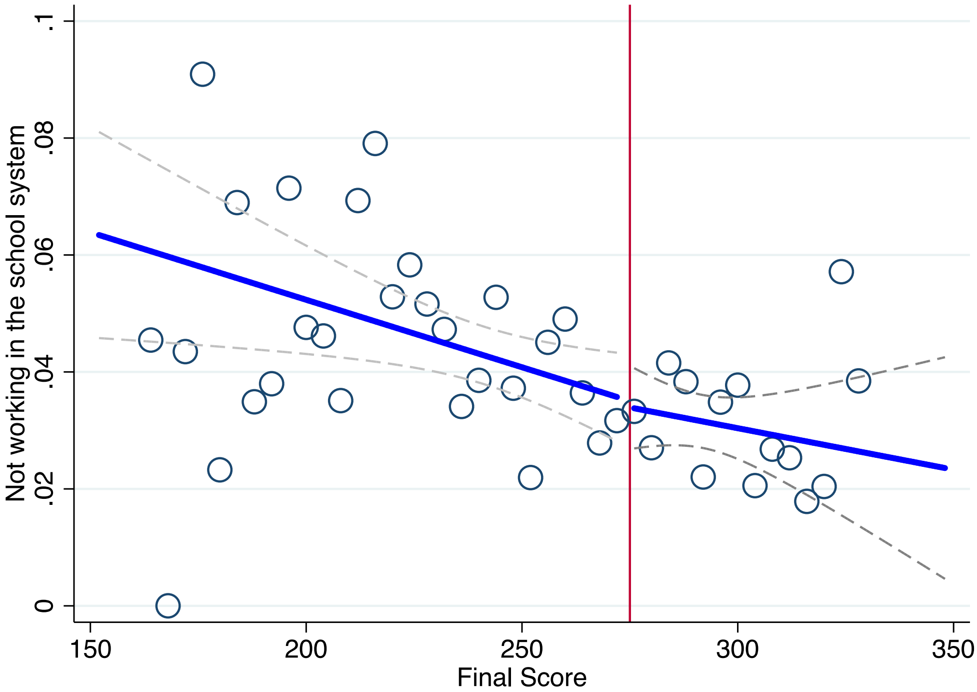
Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Circles represent the proportion of applicants passing the exam within each final score cell.

Figure 2.6: Distribution of the AEP Final Score



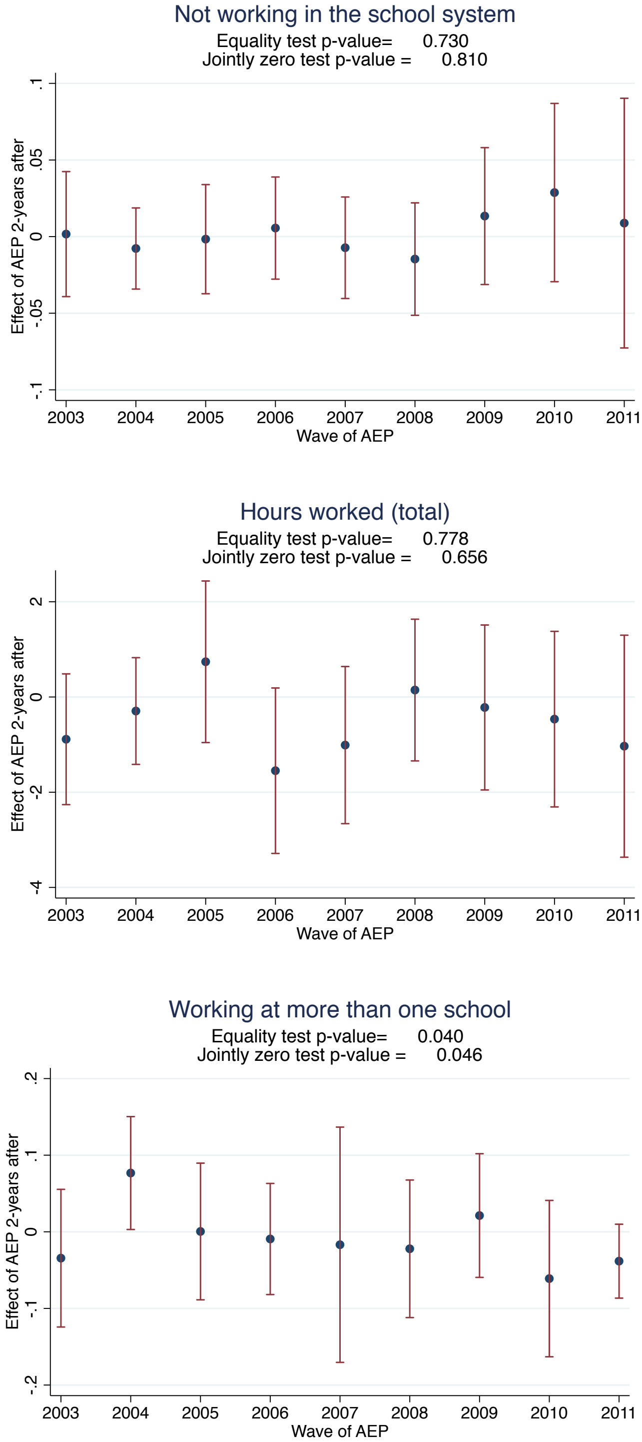
Source: Own calculations based on data from the Ministry of Education (Chile)

Figure 2.7: AEP effects on Retention and Labor Supply



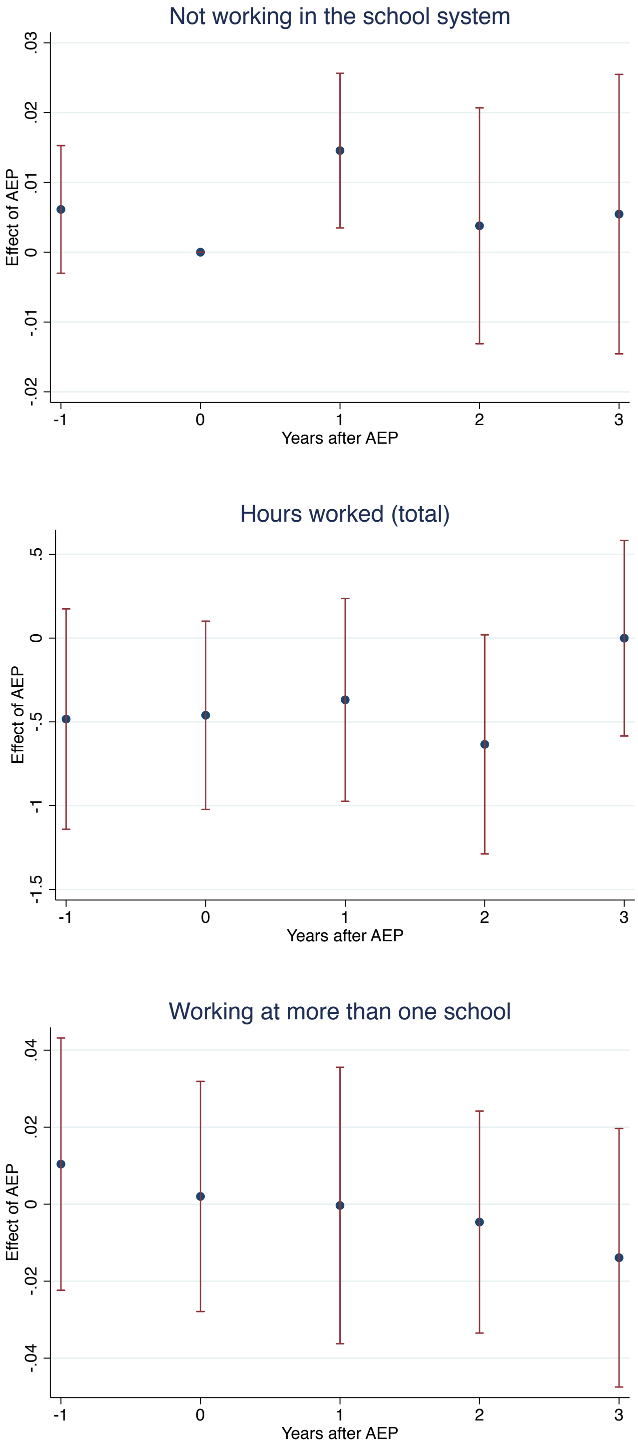
Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: The circles represent mean of the outcome variable within bins of size 4 of the final score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Doted lines represent the confidence intervals, for errors clustered at the final score cell level.

Figure 2.8: AEP effects on Retention and Labor Supply by Application Wave



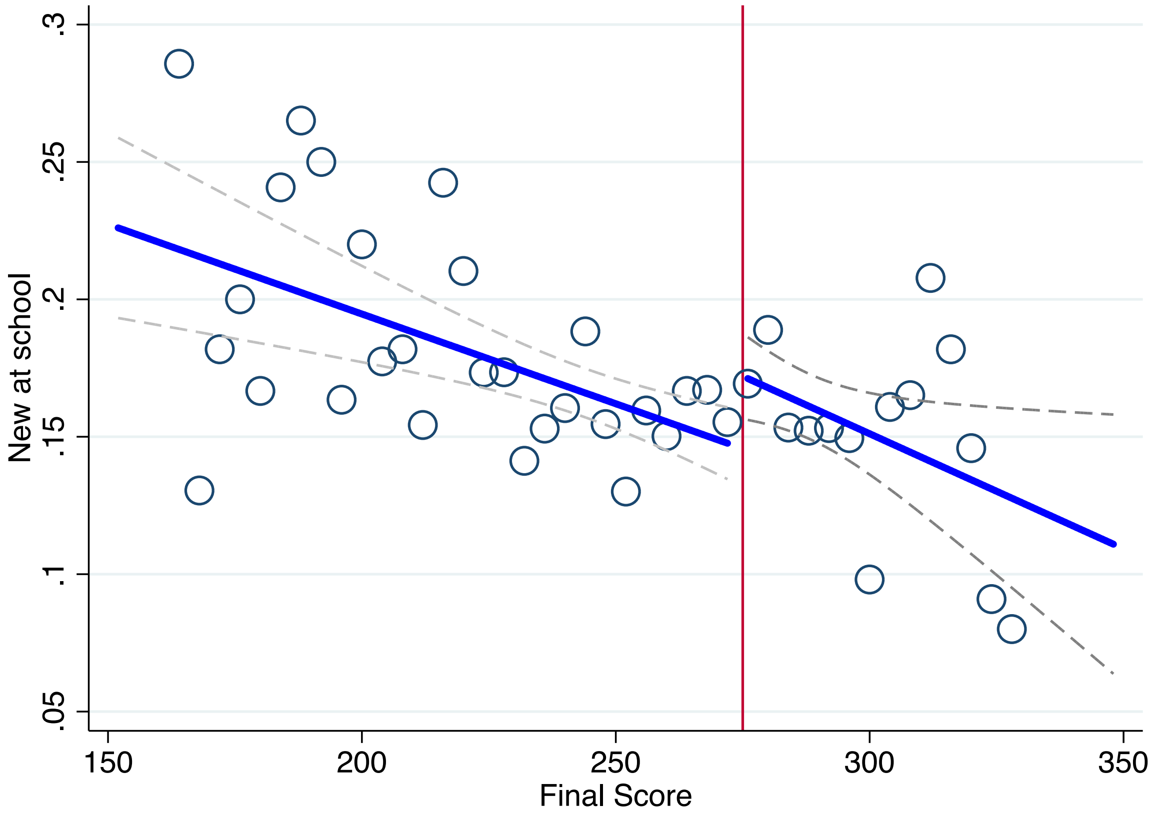
Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Each point represent the estimates of equation (2.5.1) for each application wave separately, 2 years after application. The red lines represent the 95 confidence intervals. Robust standard errors, adjusted for clustering in final score cells.

Figure 2.9: AEP effects on Retention and Labor Supply over Time



Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Each point represent the estimates of equation (2.5.1), t years after the program. The red lines represent the 95 confidence intervals. Robust standard errors, adjusted for clustering in final score cells. Dependent variable at $t = -1$ for 2003 applicants coded as missing.

Figure 2.10: AEP effects on Between-School Mobility



Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: The circles represent mean of the outcome variable within bins of size 4 of the final score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Doted lines represent the confidence intervals, for errors clustered at the final score cell level.

Tables

Table 2.1: Turnover Rates

	No of teachers (2003)	Percentage of Teachers (2005)		
Experience	Baseline	Not teaching	Change of school	Change of commune
0-11 years	38,993	18	15	8
11-12 years	35,002	8	9	4
22+ years	66,647	12	6	2
All	140,642	12	9	4

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: 2003 baseline year. Measures two years after, 2005.

Table 2.2: Proportion of Applicants Receiving the AEP Award over Time

	All	2003	2004	2005	2006	2007	2008	2009	2010	2011
	Panel A Final Sample									
AEP certification rate (%)	28	44	33	38	29	22	22	21	21	24
Compliance with allocation rule (%)	100	100	100	100	100	99	100	100	100	100
N	9,311	745	1,307	885	1,550	1,133	918	1,109	871	793
	Panel B First time applicants									
AEP certification rate (%)	26	44	32	34	28	20	19	18	18	21
Compliance with the 275 allocation rule (%)	100	100	100	100	100	99	100	100	100	100
N	13,098	935	1,561	1,658	1,988	1,483	1,494	1,597	1,286	1,096

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Standard deviation in parenthesis. Data for teachers' applying to waves 2003-2011.

Table 2.3: Test for Continuity of the Final Score

	All	2003	2004	2005	2006	2007	2008	2009	2010	2011
McCrary test p-value	0.864	0.412	0.081	0.651	0.973	0.692	0.973	0.993	0.388	0.337
Frandsen Discrete test p-value	0.363	0.339	0.251	0.250	0.892	0.870	0.397	0.880	0.277	0.856

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: McCrary (2008) test at the 275 cut-off, using a bandwidth of 30 and bin size 1.

Table 2.4: Descriptive Statistics

	2003-2014		AEP applicants	
	Voucher System Teachers	AEP Applicants	At time of application	2-years after
Male	0.291	0.278	0.300	0.299
	(0.454)	(0.448)	(0.458)	(0.458)
Age	44.131	43.906	41.196	43.306
	(11.879)	(9.695)	(8.780)	(8.768)
Degree in education	0.947	0.981	0.970	0.986
	(0.224)	(0.136)	(0.171)	(0.116)
Years of experience	17.531	17.880	14.718	17.212
	(12.560)	(10.364)	(9.118)	(9.112)
Not working in the school system				0.040
				(0.196)
Hours worked (total)	34.727	36.186	35.918	36.581
	(8.770)	(7.256)	(7.031)	(7.224)
Main job: primary school teacher	0.754	0.721	0.596	0.564
	(0.431)	(0.448)	(0.491)	(0.496)
Working at more than one school	0.103	0.132	0.156	0.136
	(0.305)	(0.338)	(0.363)	(0.342)
In a managerial job	0.104	0.058	0.027	0.065
	(0.306)	(0.234)	(0.161)	(0.246)
AEP applicant (ever)	0.065			
	(0.246)			
Currently applying to AEP				0.021
				(0.143)
Receiving AEP	0.012	0.194		0.279
	(0.111)	(0.395)		(0.448)
AVDI applicant (ever)	0.169	0.418	0.303	0.303
	(0.374)	(0.493)	(0.460)	(0.460)
Currently applying to AVDI				0.058
				(0.233)
Receiving AVDI	0.026	0.097	0.047	0.057
	(0.160)	(0.295)	(0.212)	(0.231)
New at school	0.121	0.103		0.166
	(0.326)	(0.304)		(0.373)
Private-subsidized school	0.402	0.456	0.550	0.519
	(0.490)	(0.498)	(0.497)	(0.500)
Private school	0.117	0.019		0.012
	(0.322)	(0.138)		(0.107)
Working conditions (top-50 school)	0.404	0.417	0.426	0.432
	(0.491)	(0.493)	(0.495)	(0.495)
Student performance (top-50 school)	0.629	0.693	0.695	0.695
	(0.483)	(0.461)	(0.460)	(0.461)
SNED awarded school	0.325	0.376	0.369	0.404
	(0.468)	(0.484)	(0.483)	(0.491)
Change of municipality	0.067	0.056		0.101
	(0.251)	(0.230)		(0.301)
Rural school	0.143	0.131	0.115	0.113
	(0.350)	(0.337)	(0.319)	(0.316)
In municipality with zone allowance	0.419	0.461	0.453	0.455
	(0.493)	(0.498)	(0.498)	(0.498)
N	1,576,800	98,287	9,311	9,311

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Standard deviation in parenthesis. For the 2003-2014 period, *New at school* stands for whether or not the teacher was teaching at that particular school in the previous year. For the AEP applicants 2 years after application, *New at school* is a dummy taking the value of 1 if the school is different from the school at time of application. Except *Not working in the school system*, the dependent variables for teachers not working in the school system 2 years after application are coded as missing. From the full set of observations in the 2003-2014 period, 16 percent are from teachers who applied to the at some time during the period. Yet, only 6 percent of the ever eligible candidates applied to the AEP.

Table 2.5: Balance at Baseline AEP

AEP	Degree of polynomial			
Dependent Variable	(1)	(2)	(3)	(4)
Male	0.009	0.023	0.015	0.042
	(0.017)	(0.021)	(0.024)	(0.029)
Age	-0.135	-0.316	-0.237	-0.172
	(0.282)	(0.420)	(0.564)	(0.725)
Degree in education	0.003	0.008	0.007	0.017
	(0.006)	(0.007)	(0.010)	(0.011)
Years of experience	-0.333	-0.733	-0.748	-0.811
	(0.319)	(0.476)	(0.616)	(0.776)
Hours worked (total)	-0.165	-0.461	-0.922***	-0.479
	(0.226)	(0.285)	(0.334)	(0.365)
Working at more than one school	0.004	0.002	0.010	-0.030
	(0.012)	(0.015)	(0.019)	(0.026)
In a managerial job	-0.003	-0.009	-0.013	-0.013
	(0.006)	(0.007)	(0.008)	(0.008)
Main job: primary school teacher	0.005	-0.012	-0.004	0.029
	(0.014)	(0.018)	(0.021)	(0.024)
Receiving AVDI	0.008	-0.010	-0.013	-0.003
	(0.009)	(0.012)	(0.015)	(0.016)
Private-subsidized school	-0.027	-0.014	-0.019	-0.010
	(0.023)	(0.033)	(0.039)	(0.045)
Working conditions (top-50 school)	-0.037**	-0.005	0.024	0.040
	(0.018)	(0.026)	(0.033)	(0.040)
Student performance (top-50 school)	-0.009	0.012	-0.005	0.014
	(0.019)	(0.026)	(0.033)	(0.038)
SNED awarded school	0.009	0.017	0.010	0.015
	(0.017)	(0.022)	(0.027)	(0.031)
Rural school	0.020*	0.019	0.029	0.038*
	(0.012)	(0.015)	(0.018)	(0.021)
In municipality with zone allowance	-0.011	-0.013	0.002	0.048
	(0.019)	(0.023)	(0.027)	(0.031)
Wald test p-value	0.5089	0.1734	0.0087	0.0004

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AEP data for teachers' applying to AEP waves 2003-2011, at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Robust standard errors,adjusted for clustering in final score cells, in parenthesis. Column numbers indicate the order of the polynomial on the score centered around 275.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Table 2.6: AEP effects on Retention and Labor Supply

	All Teachers		0-11 years		12-21 years		22 +years	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not working in the school system	0.0038	0.0036	-0.0018	0.0063	0.0078	0.0083	-0.0061	-0.0014
	(0.0086)	(0.0096)	(0.0167)	(0.0175)	(0.0179)	(0.0188)	(0.0098)	(0.0098)
N	9,311	9,311	3,756	3,756	2,872	2,872	2,683	2,683
Clusters	230	230	187	187	198	198	206	206
Hours worked (total)	-0.6344*	-0.5327*	-0.3629	-0.3012	-0.2493	-0.4344	-1.5792**	-1.1180*
	(0.3317)	(0.3218)	(0.5235)	(0.5758)	(0.5375)	(0.4566)	(0.6775)	(0.6235)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205
Working at more than one school	-0.0047	-0.0032	-0.0041	-0.0054	-0.0048	-0.0010	0.0159	0.0200
	(0.0146)	(0.0141)	(0.0224)	(0.0223)	(0.0260)	(0.0247)	(0.0278)	(0.0291)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AEP data for teachers' applying to AEP waves 2003-2011, 2 years after application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Odd-columns present the estimates of equation (2.5.1). Even columns present the estimates of equation (2.5.1) and add controls interacted with wave fixed effects. Controls include gender, age, degree in education, years of experience, teaching at a single school, hours worked, receiving AVDI, rural school, private-subsidized school, working conditions (top-50 school), student performance (top-50 school), SNED awarding school and education level of the main job: primary; all at time of application and excluding the outcome variable at time of application. Robust standard errors, adjusted for clustering in final score cells, in parenthesis.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Table 2.7: AEP effects on Between-School Mobility

	All Teachers		0-11 years		12-21 years		22 +years	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New at school	0.0447***	0.0434**	0.0470	0.0530	0.0296	0.0338	0.0394	0.0346
	(0.0168)	(0.0179)	(0.0314)	(0.0343)	(0.0342)	(0.0341)	(0.0266)	(0.0293)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205
Private-subsidized school	-0.0350	-0.0424*	-0.0799*	-0.0803**	0.0438	0.0321	-0.0848**	-0.0738*
	(0.0329)	(0.0227)	(0.0453)	(0.0362)	(0.0402)	(0.0305)	(0.0428)	(0.0398)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205
Private school	0.0006	0.0005	0.0197	0.0184	-0.0242***	-0.0268***	-0.0034	-0.0037
	(0.0061)	(0.0059)	(0.0131)	(0.0139)	(0.0069)	(0.0077)	(0.0030)	(0.0030)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205
In municipality with zone allowance	-0.0153	-0.0092	-0.0110	-0.0016	0.0488	0.0492	-0.1000**	-0.0753*
	(0.0239)	(0.0252)	(0.0364)	(0.0378)	(0.0482)	(0.0487)	(0.0431)	(0.0452)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205
Rural school	0.0245*	0.0219*	0.0292	0.0275	0.0301	0.0319	0.0222	0.0124
	(0.0138)	(0.0130)	(0.0217)	(0.0217)	(0.0254)	(0.0274)	(0.0254)	(0.0242)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Clusters	229	229	187	187	197	197	205	205
Working conditions (top-50 school)	-0.0180	-0.0245	-0.0197	-0.0254	0.0179	-0.0029	-0.0522	-0.0397
	(0.0271)	(0.0248)	(0.0386)	(0.0403)	(0.0422)	(0.0425)	(0.0387)	(0.0366)
N	8,831	8,831	3,469	3,469	2,741	2,741	2,621	2,621
Clusters	229	229	185	185	197	197	205	205
Student performance (top-50 school)	0.0032	-0.0047	0.0084	-0.0019	0.0764*	0.0605	-0.0954*	-0.0749
	(0.0287)	(0.0279)	(0.0350)	(0.0352)	(0.0409)	(0.0406)	(0.0552)	(0.0481)
N	8,831	8,831	3,469	3,469	2,741	2,741	2,621	2,621
Clusters	229	229	185	185	197	197	205	205

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AEP data for teachers' applying to AEP waves 2003-2011, 2 years after application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Odd-columns present the estimates of equation (2.5.1). Even columns present the estimates of equation (2.5.1) and add controls interacted with wave fixed effects. Controls include gender, age, degree in education, years of experience, teaching at a single school, hours worked, receiving AVDI, rural school, private-subsidized school, working conditions (top-50 school), student performance (top-50 school), SNED awarding school and education level of the main job: primary; all at time of application and excluding the outcome variable at time of application. Robust standard errors, adjusted for clustering in final score cells, in parenthesis.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Table 2.8: AEP Heterogenous Effect on Between-School Mobility

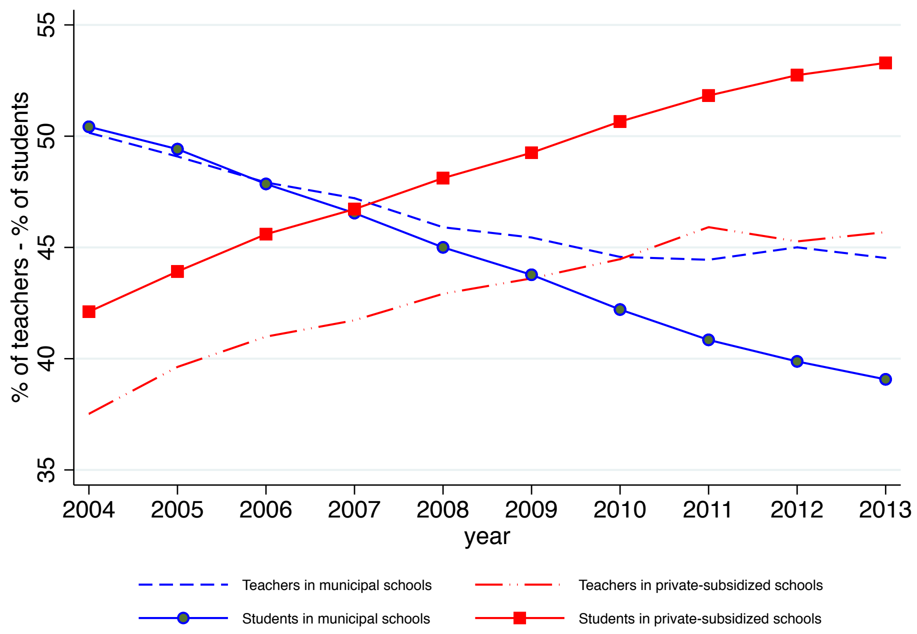
	All Teachers	0-11 years	12-21 years	22 +years
By Working Conditions				
At good school when applying	-0.0009	0.0067	-0.0312	-0.0010
	(0.0349)	(0.0476)	(0.0588)	(0.0391)
At bad school when applying	0.0753***	0.0617	0.0927**	0.0655
	(0.0208)	(0.0488)	(0.0436)	(0.0402)
By Student Performance				
At good school when applying	0.0256	0.0468	-0.0240	0.0496
	(0.0211)	(0.0421)	(0.0269)	(0.0334)
At bad school when applying	0.0966**	0.1078	0.2010**	0.0341
	(0.0421)	(0.0801)	(0.0952)	(0.0525)

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AEP data for teachers' applying to AEP waves 2003-2011, 2 years after application. A school is coded as good if its characteristic listed in the corresponding Panel is above the median. A school is coded as bad if its characteristic listed in the corresponding Panel is below the median. Separate equations estimated for each good and bad schools. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Tables and Figures for Chapter 3

Figures

Figure 3.1: Student enrollment and teacher supply across municipal and private-subsidized schools.



Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Private sector as residual.

Figure 3.2: Timeline

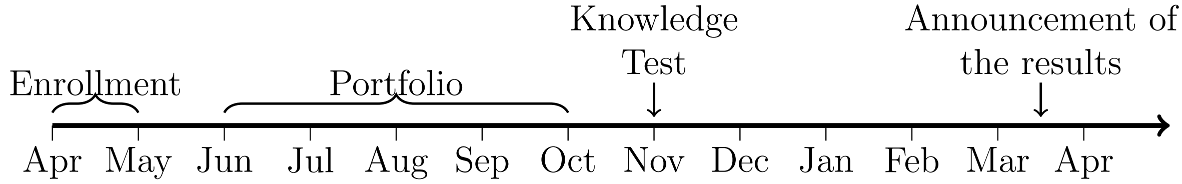
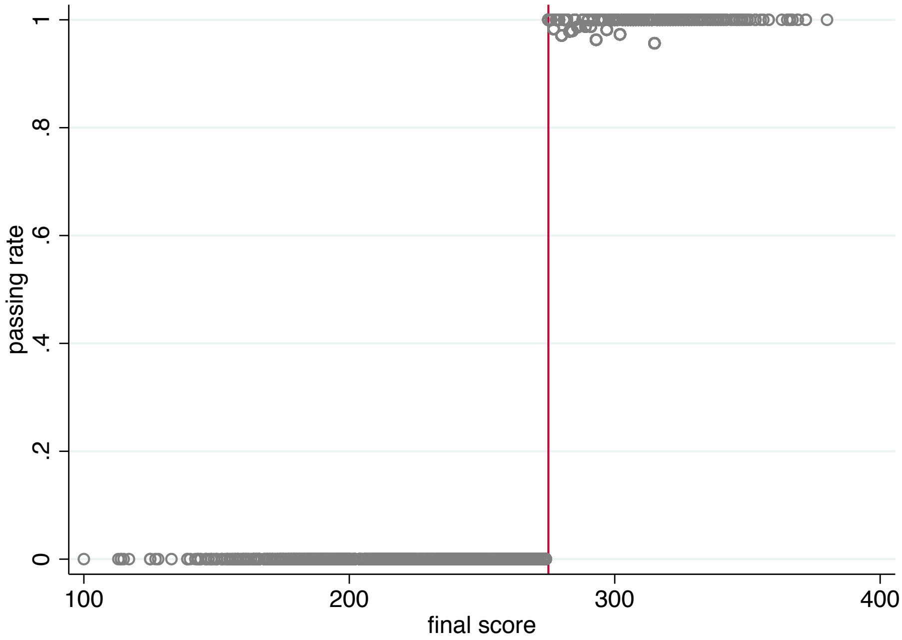
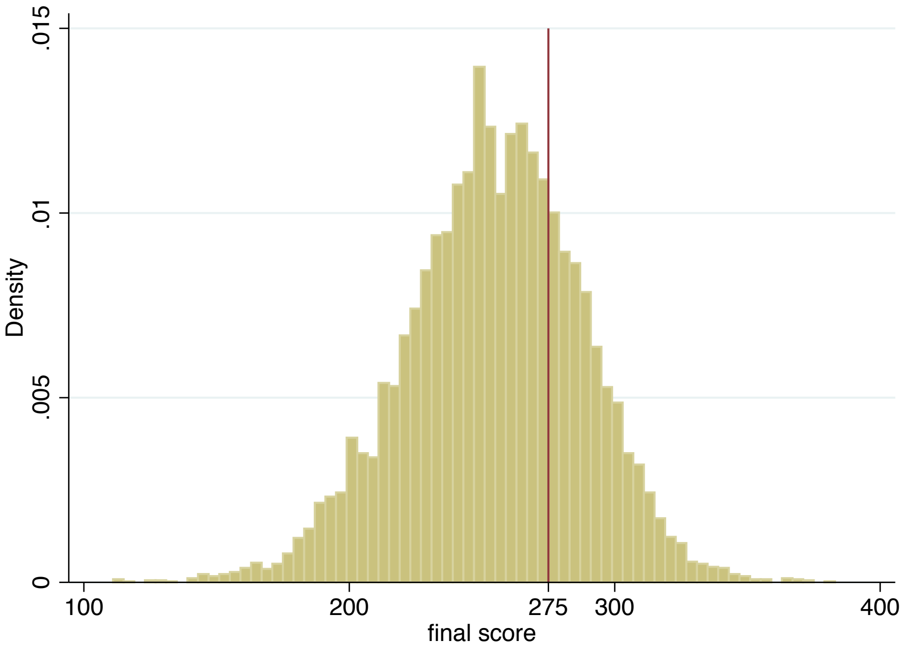


Figure 3.3: AEP assignment rule



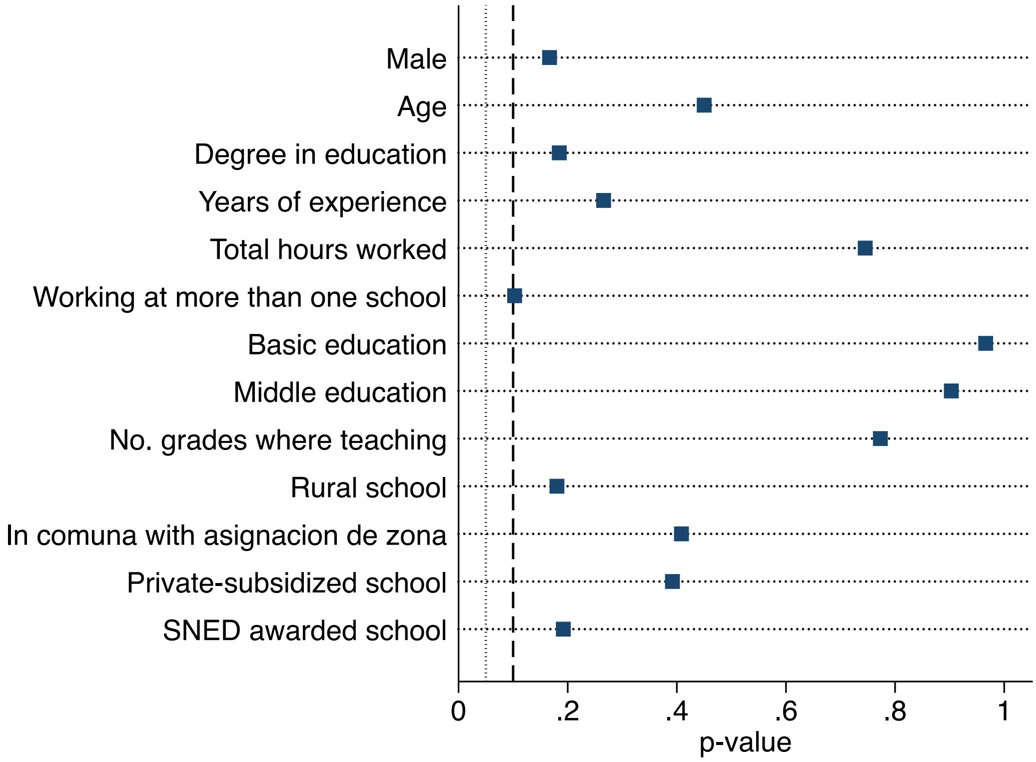
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Circles represent the share of applicants passing the exam within each score cell.

Figure 3.4: Distribution of the AEP score



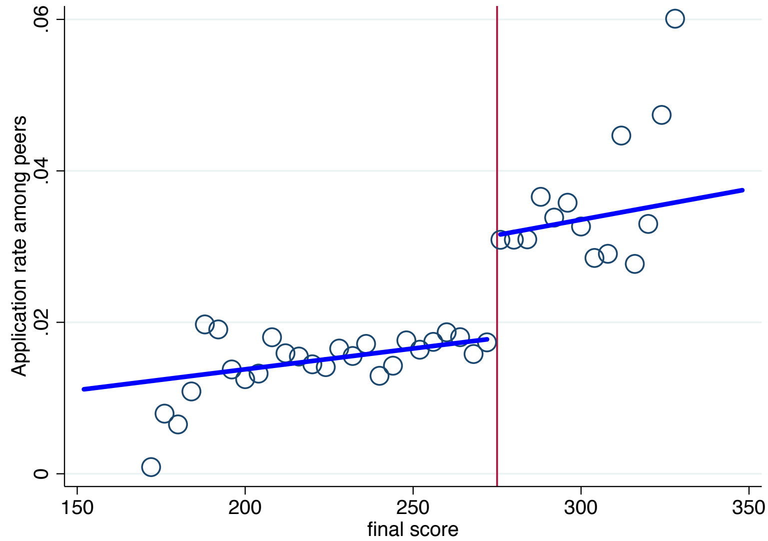
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Full sample of 8,937 applicants.

Figure 3.5: Balance of applicants' characteristics at baseline



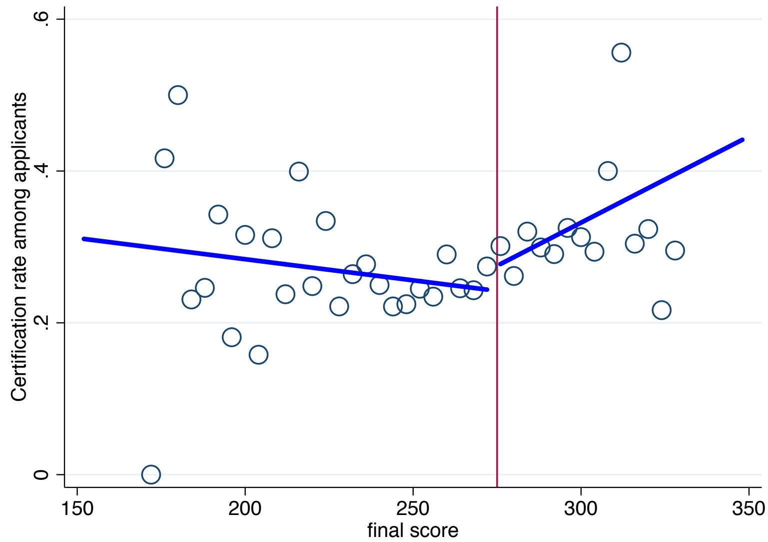
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Polynomial of order 1. The dashed line indicates statistical significance at 10%. The dotted line indicates statistical significance at 5%.

Figure 3.6: Effect of AEP certification on peers' one period ahead application rate (Reference group I)



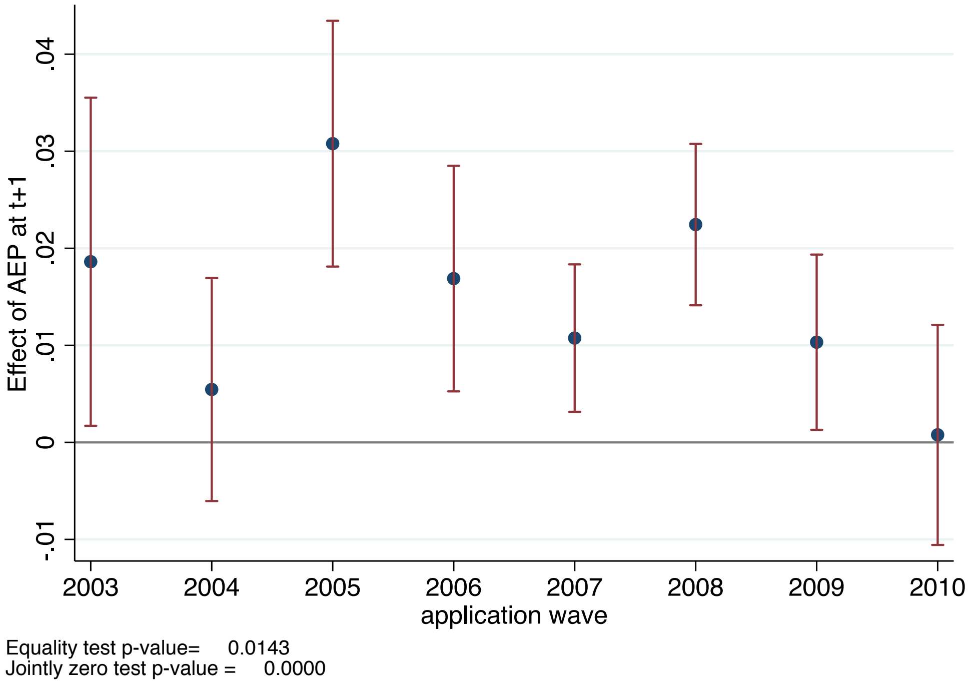
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: The circles represent mean of the outcome variable within bins of size 4 of the score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Errors clustered at the score cell level. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.7: Effect of AEP certification on peers' one period ahead certification rate (Reference group I)



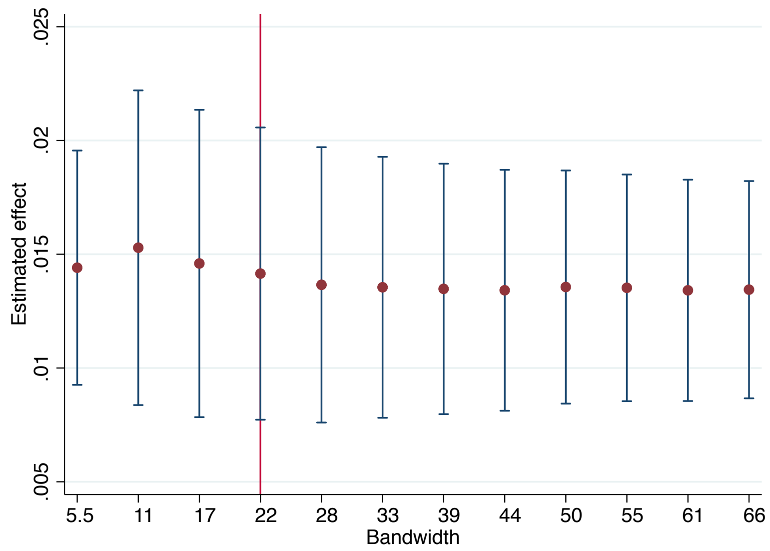
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. The circles represent mean of the outcome variable within bins of size 4 of the score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Errors clustered at the score cell level. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.8: Effect of AEP certification on peers' one period ahead application rate, by application wave (Reference group I)



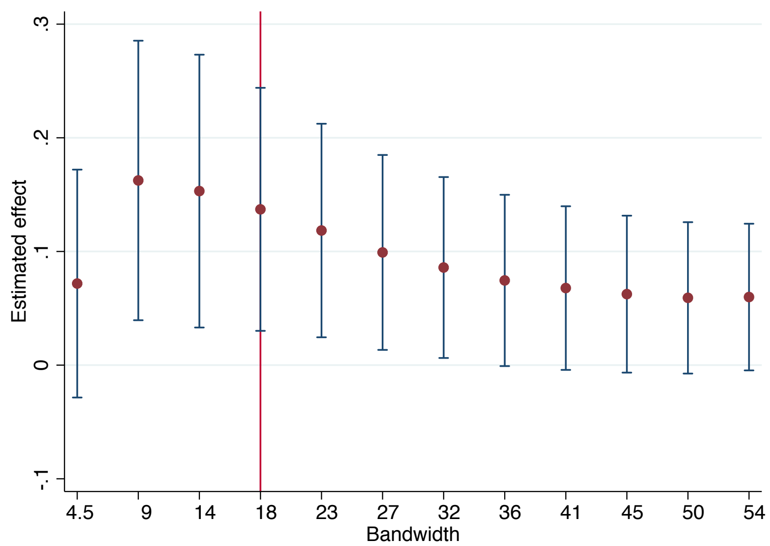
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Each point represent the estimates of equation (3.4.1) for each application wave separately. The red lines represent the 95 confidence intervals. Robust standard errors, adjusted for clustering in final score cells. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.9: Effect of AEP certification on peers' one period ahead application rate, alternative bandwidths (Reference group I)



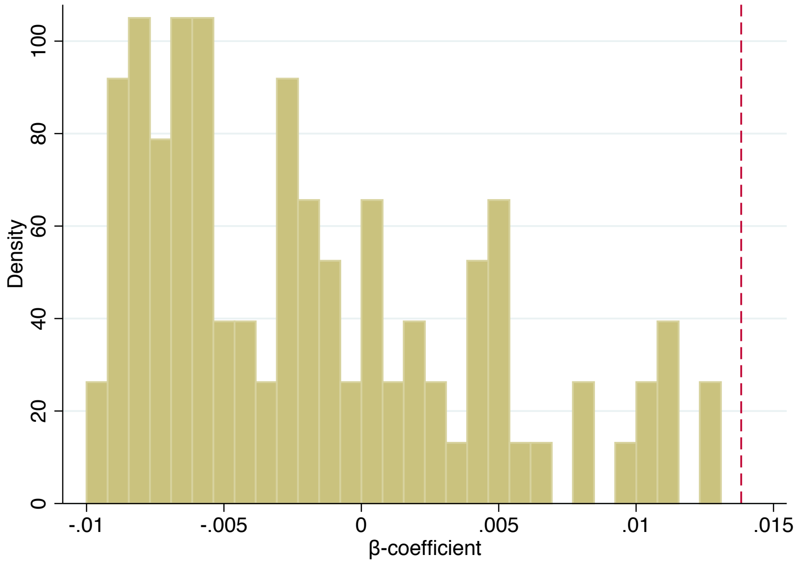
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Each point represent the point estimate of a fully non-parametric specification for each application bandwidth separately. The blue lines represent the 95 confidence intervals. The red line indicates the optimal bandwidth following Imbens and Kalyanaraman (2011). Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.10: Effect of AEP certification on peers' one period ahead certification rate, alternative bandwidths (Reference group I)



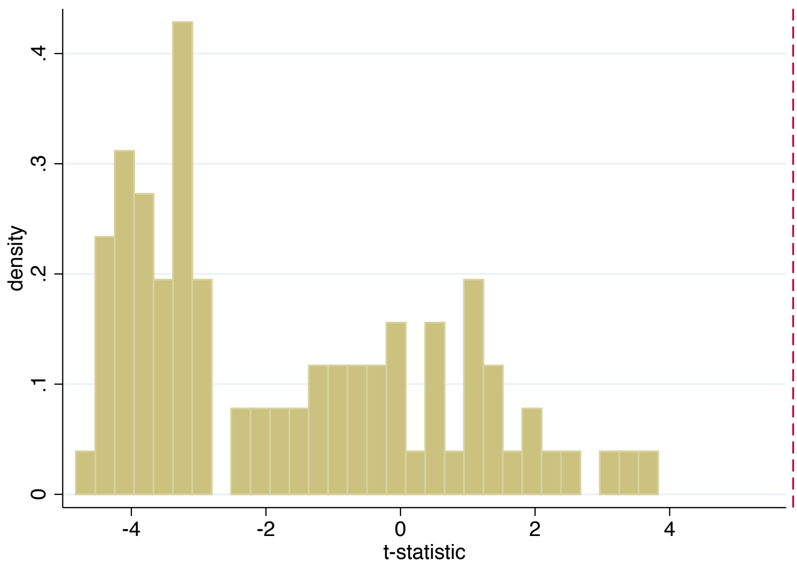
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. Each point represent the point estimate of a fully non-parametric specification for each application bandwidth separately. The blue lines represent the 95 confidence intervals. The red line indicates the optimal bandwidth following Imbens and Kalyanaraman (2011). Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.11: Distribution of the estimated β-coefficient of peers' one period ahead application rate, for fake cutoffs (Reference group I)



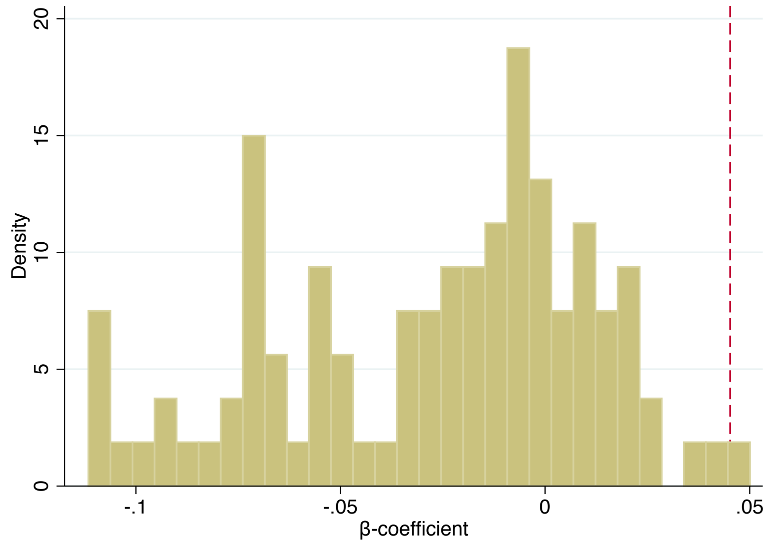
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Histogram of the estimated β-coefficient of equation (3.4.1) using 200-270 and 280-300 as discontinuity cut-offs. The dashed line corresponds to the β-coefficient of the true discontinuity cut-off (275). Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.12: Distribution of the t-statistics of the estimated β-coefficient of peers' one period ahead application rate, for fake cutoffs (Reference group I)



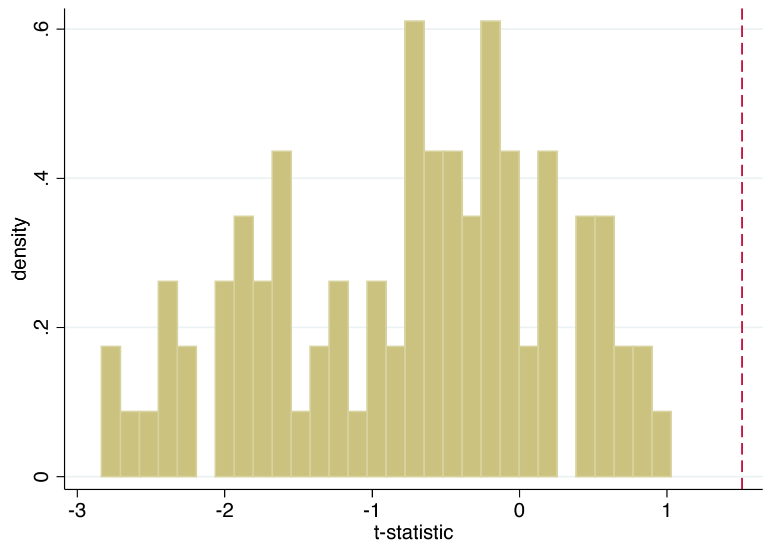
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Histogram of the associated t-statistics of the estimated β-coefficient of equation (3.4.1) using 200-270 and 280-300 as discontinuity cut-offs. The dashed line corresponds to the t-statistics of the β-coefficient of the true discontinuity cut-off (275). Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.13: Distribution of the estimated β-coefficient of peers' one period ahead certification rate, for fake cutoffs (Reference group I)



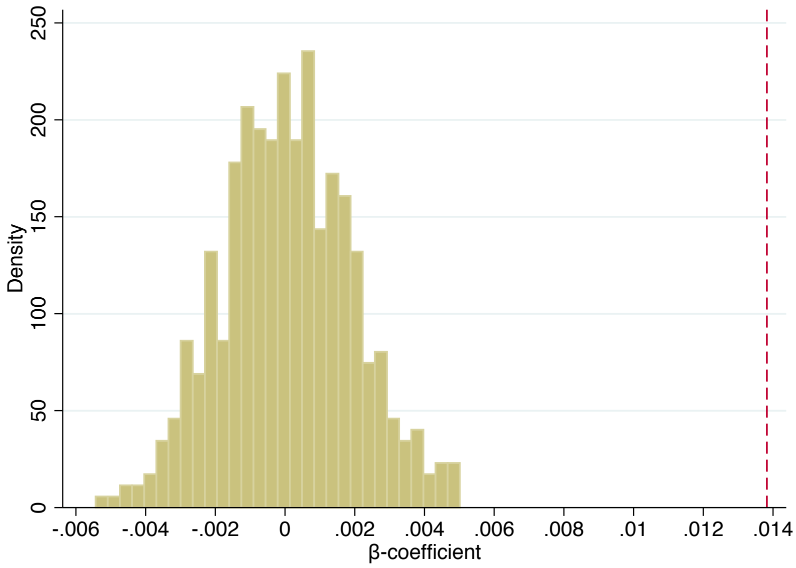
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. Histogram of the estimated β-coefficient of equation (3.4.1) using 200-270 and 280-300 as discontinuity cut-offs. The dashed line corresponds to the β-coefficient of the true discontinuity cut-off (275). Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.14: Distribution of the t-statistics of the estimated β-coefficient of peers' one period ahead certification rate, for fake cutoffs (Reference group I)



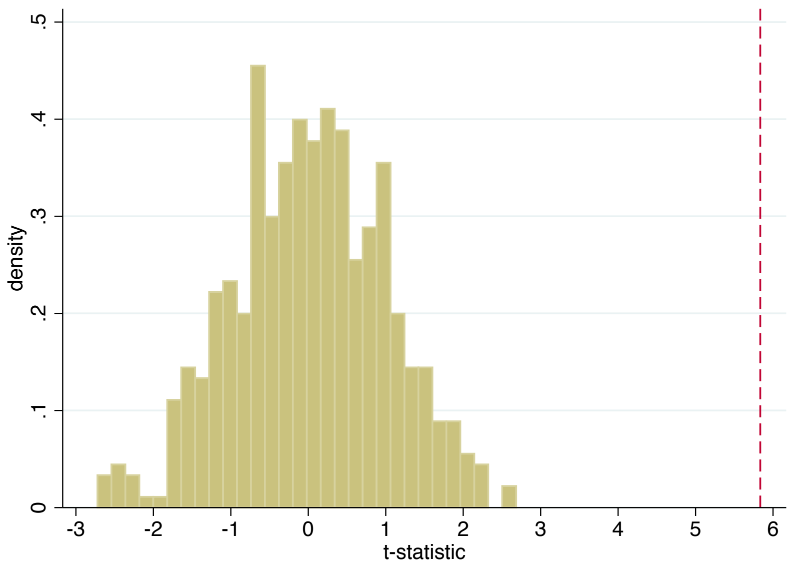
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. Histogram of the associated t-statistics of the estimated β-coefficient of equation (3.4.1) using 200-270 and 280-300 as discontinuity cut-offs. The dashed line corresponds to the t-statistics of the β-coefficient of the true discontinuity cut-off (275). Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.15: Distribution of the estimated β-coefficient of peers' one period ahead application rate, for random reference groups (Reference group I)



Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Histogram of the estimated β-coefficient of equation (3.4.1) for 500 random draws of reference groups. The dashed line corresponds to the β-coefficient of the true reference group. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Figure 3.16: Distribution of the t-statistics of the estimated β-coefficient of peers' one period ahead application rate, for random reference groups (Reference group I)



Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. Histogram of the estimated β-coefficient of equation (3.4.1) for 500 random draws of reference groups. The dashed line corresponds to the β-coefficient of the true reference group. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application.

Tables

Table 3.1: Definition of reference groups

	Same as the applicant			Minimum
	school	level of instruction	grade	No. of peers
Reference group I	X	0	0	4
Reference group II	X	X	0	2
Reference group III	X	X	X	2

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: All definitions at the time of application. Grade, means teaching at, at least, one same grade as the applicant.

Table 3.2: Reference groups

	No. of applicant-specific reference groups	No. of unique peer observations	Average group size	No. applicants per reference group
Reference group I	8,937	148,822	27.09	1.948
Reference group II	8,937	132,286	23.99	1.875
Reference group III	8,608	91,590	17.77	1.666

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Data for teachers' applying to waves 2003-2011.

Table 3.3: Descriptive statistics

	Applicants	Reference Group I (school)	Reference Group II (primary-secondary)	Reference Group III (grade)	Voucher system teachers
Male	0.30	0.33	0.33	0.38	0.31
	(0.46)	(0.47)	(0.47)	(0.49)	(0.46)
Age	42.17	44.45	44.67	44.49	45.12
	(9.17)	(10.98)	(10.96)	(10.90)	(11.06)
Degree in education	0.97	0.91	0.91	0.90	0.92
	(0.17)	(0.29)	(0.28)	(0.30)	(0.27)
Years of experience	15.58	17.48	17.71	17.34	18.06
	(9.67)	(11.71)	(11.73)	(11.57)	(12.04)
Total hours worked	38.49	36.86	36.89	37.19	36.59
	(8.20)	(9.49)	(9.42)	(9.65)	(9.58)
Working at more than one school	0.15	0.14	0.14	0.16	0.13
	(0.36)	(0.35)	(0.35)	(0.37)	(0.33)
Basic education	0.69	0.62	0.63	0.51	0.73
	(0.46)	(0.49)	(0.48)	(0.50)	(0.44)
Middle education	0.44	0.51	0.51	0.65	0.36
	(0.50)	(0.50)	(0.50)	(0.48)	(0.48)
No. grades where teaching	3.38	3.02	3.07	3.50	3.22
	(1.85)	(1.79)	(1.83)	(1.82)	(1.93)
Rural school	0.10	0.05	0.05	0.05	0.16
	(0.31)	(0.22)	(0.23)	(0.22)	(0.37)
In comuna with asignacion de zona	0.46	0.45	0.46	0.46	0.45
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Private-subsidized school	0.49	0.49	0.46	0.45	0.44
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
SNED awarded school	0.37	0.38	0.38	0.38	0.31
	(0.48)	(0.49)	(0.49)	(0.48)	(0.46)
	8,937	148,822	132,286	91,590	1,210,081

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Standard deviation in parenthesis. The data on primary or secondary education, and the number of grades where teaching of column three comes from a sub-sample of 903,678 teachers.

Table 3.4: Application, re-application, and certification rate to AEP

	Ever applicants	Waves 2003	2004	2005	2006	2007	2008	2009	2010
Application rate	5.39	0.65	1.09	1.16	1.39	1.03	1.04	1.09	0.87
Repetition rate	12.02	12.38	7.89	11.91	12.16	12.83	10.78	12.26	13.74
Certification rate	24.98	42.01	32.23	33.71	28.05	20.49	19.35	17.97	17.67
Retakers' certification rate	27.98	54.21	30.71	40.09	29.85	20.75	16.48	15.00	15.35
Sample certification rate	26.96	42.64	32.73	35.11	27.40	21.28	24.92	20.35	19.98
N. applicants	11,521	757	1,483	1,605	1,936	1,440	1,457	1,575	1,268
N. retakers	1,529	107	127	217	268	212	176	220	202
N. sample	8,937	591	1,317	903	1,555	1,292	971	1,302	1,006

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Data for teachers' applying to waves 2003-2010.

Table 3.5: Test for continuity of the AEP final score

	All	2003	2004	2005	2006	2007	2008	2009	2010
Frandsen Discrete test p-value	0.781	0.522	0.289	0.508	0.946	0.875	0.621	0.970	0.929
McCrary test p-value	0.351	0.120	0.250	0.715	0.691	0.363	0.529	0.695	0.535

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: McCrary (2008) test at the 275 cut-off, using the optimal bandwidth of 20 and bin size 1.

Table 3.6: Balance of applicants and average reference group characteristics at baseline

	Applicants		Reference Group I (school)		Reference Group II (primary-secondary)		Reference Group III (grade)	
Dependent Variable	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Male	0.018	0.028	0.001	0.005	0.007	0.015	0.013*	0.023**
	(0.016)	(0.021)	(0.006)	(0.008)	(0.007)	(0.009)	(0.008)	(0.012)
Age	-0.227	-0.431	-0.120	-0.258	-0.159	-0.317	-0.270	-0.480*
	(0.316)	(0.492)	(0.185)	(0.249)	(0.190)	(0.254)	(0.213)	(0.290)
Degree in education	0.006	0.008	0.000	-0.001	-0.002	-0.006	-0.003	-0.006
	(0.005)	(0.008)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
Years of experience	-0.418	-0.654	-0.152	-0.124	-0.196	-0.205	-0.288	-0.412
	(0.365)	(0.549)	(0.193)	(0.262)	(0.202)	(0.274)	(0.215)	(0.296)
Total hours worked	0.072	-0.294	-0.166	-0.139	-0.139	-0.133	-0.024	-0.121
	(0.310)	(0.407)	(0.152)	(0.201)	(0.153)	(0.204)	(0.174)	(0.220)
Working at more than one school	0.019	0.018	-0.002	0.005	-0.001	0.005	0.004	0.011
	(0.013)	(0.017)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
Basic education	0.001	-0.009	0.019	0.023	0.010	0.006	0.006	0.000
	(0.015)	(0.018)	(0.012)	(0.015)	(0.012)	(0.014)	(0.012)	(0.014)
Middle education	0.005	0.003	-0.008	-0.015	0.004	0.006	0.007	0.013
	(0.016)	(0.020)	(0.013)	(0.018)	(0.014)	(0.018)	(0.014)	(0.018)
No. grades where teaching	0.004	-0.025	0.045	0.017	0.055	0.030	0.013	0.008
	(0.060)	(0.083)	(0.031)	(0.048)	(0.034)	(0.052)	(0.044)	(0.059)
Rural school	0.015	0.024*	0.015	0.024*	0.015	0.024*	0.015	0.018
	(0.011)	(0.013)	(0.011)	(0.013)	(0.011)	(0.013)	(0.010)	(0.012)
Private-subsidized school	0.023	0.018	0.023	0.018	0.023	0.018	0.024	0.025
	(0.024)	(0.034)	(0.024)	(0.034)	(0.024)	(0.034)	(0.023)	(0.034)
SNED awarded school	0.020	0.024	0.020	0.024	0.020	0.024	0.019	0.027
	(0.015)	(0.019)	(0.015)	(0.019)	(0.015)	(0.019)	(0.015)	(0.019)

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Columns one and two present the characteristics of the 2003-2010 AEP applicants at the time of application. Columns three to eight, present the average characteristics of the applicants' reference group at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Robust standard errors, adjusted for clustering at score cells, in parenthesis. Column numbers in parenthesis indicate the order of the polynomial on the score centered around 275.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.7: Effect of AEP certification on peers' future application rate (Reference group I)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Application rate among peers	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0138***	0.0138***	0.0153***	0.0149***	0.0141***	0.0137***
	(0.0024)	(0.0024)	(0.0026)	(0.0026)	(0.0031)	(0.0031)
N	8,934	8,934	6,384	6,384	4,645	4,645
Clusters	224	224	207	207	205	205
t+2	0.0075***	0.0075***	0.0078***	0.0075***	0.0086***	0.0082***
	(0.0017)	(0.0017)	(0.0019)	(0.0019)	(0.0023)	(0.0022)
N	8,935	8,935	6,385	6,385	4,646	4,646
Clusters	224	224	207	207	205	205
t+3	0.0007	0.0006	0.0013	0.0011	0.0021	0.0020
	(0.0016)	(0.0016)	(0.0017)	(0.0017)	(0.0019)	(0.0018)
N	8,933	8,933	6,383	6,383	4,644	4,644
Clusters	224	224	207	207	205	205
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Polynomial of order 1. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis. Sample sizes reduces whenever an applicant has no eligible peers.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.8: Effect of AEP certification on peers' future application rate (Reference Group II)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Application rate among peers	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0146***	0.0145***	0.0156***	0.0153***	0.0150***	0.0148***
	(0.0022)	(0.0023)	(0.0025)	(0.0025)	(0.0031)	(0.0032)
N	8,934	8,934	6,513	6,513	4,831	4,831
Clusters	224	224	208	208	206	206
t+2	0.0072***	0.0070***	0.0077***	0.0074***	0.0083***	0.0081***
	(0.0019)	(0.0019)	(0.0020)	(0.0020)	(0.0022)	(0.0022)
N	8,935	8,935	6,514	6,514	4,832	4,832
Clusters	224	224	208	208	206	206
t+3	0.0004	0.0003	0.0007	0.0005	0.0021	0.0020
	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0020)	(0.0020)
N	8,933	8,933	6,512	6,512	4,830	4,830
Clusters	224	224	208	208	206	206
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Polynomial of order 1. Definition of peers according to reference group II, i.e. teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis. Sample sizes reduces whenever an applicant has no eligible peers.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.9: Effect of AEP certification on peers' future application rate (Reference Group III)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Application rate among peers	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0192***	0.0190***	0.0201***	0.0196***	0.0190***	0.0186***
	(0.0028)	(0.0029)	(0.0032)	(0.0033)	(0.0039)	(0.0040)
N	8,582	8,582	6,676	6,676	5,275	5,275
Clusters	222	222	209	209	207	207
t+2	0.0079***	0.0078***	0.0083***	0.0081***	0.0098***	0.0095***
	(0.0026)	(0.0026)	(0.0028)	(0.0028)	(0.0031)	(0.0031)
N	8,574	8,574	6,668	6,668	5,267	5,267
Clusters	222	222	209	209	207	207
t+3	0.0002	-0.0000	0.0015	0.0010	0.0035	0.0029
	(0.0020)	(0.0020)	(0.0021)	(0.0020)	(0.0024)	(0.0023)
N	8,561	8,561	6,659	6,659	5,260	5,260
Clusters	222	222	209	209	207	207
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Polynomial of order 1. Definition of peers according to reference group III, i.e. teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis. Sample sizes reduces whenever an applicant has no eligible peers.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.10: Effect of AEP certification on peers' future certification rate (Reference group I)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Certification rate among applicants	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0452	0.0430	0.0774**	0.0733**	0.0499	0.0363
	(0.0300)	(0.0303)	(0.0304)	(0.0316)	(0.0388)	(0.0407)
N	2,388	2,388	1,500	1,500	952	952
Clusters	179	179	160	160	157	157
t+2	-0.0499	-0.0479	-0.0521	-0.0507	-0.0778	-0.0809
	(0.0333)	(0.0318)	(0.0396)	(0.0384)	(0.0546)	(0.0562)
N	1,867	1,867	1,201	1,201	771	771
Clusters	176	176	151	151	148	148
t+3	-0.0140	-0.0174	0.0026	-0.0037	-0.0327	-0.0272
	(0.0292)	(0.0295)	(0.0340)	(0.0346)	(0.0523)	(0.0527)
N	1,489	1,489	922	922	586	586
Clusters	177	177	150	150	148	148
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. OLS regression. Polynomial of order 1. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.11: Effect of AEP Certification on peers' future certification rate (Reference Group II)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Certification rate among applicants	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0231	0.0211	0.0503	0.0465	0.0183	0.0046
	(0.0329)	(0.0334)	(0.0353)	(0.0368)	(0.0415)	(0.0442)
N	2,199	2,199	1,418	1,418	923	923
Clusters	176	176	157	157	155	155
t+2	-0.0645**	-0.0629**	-0.0740*	-0.0722*	-0.0800	-0.0812
	(0.0325)	(0.0306)	(0.0403)	(0.0390)	(0.0562)	(0.0578)
N	1,717	1,717	1,136	1,136	748	748
Clusters	176	176	149	149	147	147
t+3	-0.0170	-0.0145	0.0110	0.0094	-0.0194	-0.0127
	(0.0331)	(0.0325)	(0.0365)	(0.0371)	(0.0512)	(0.0534)
N	1,362	1,362	864	864	562	562
Clusters	172	172	146	146	143	143
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. OLS regression. Polynomial of order 1. Definition of peers according to reference group II, i.e. teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.12: Effect of AEP Certification on peers' future certification rate (Reference Group III)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Certification rate among applicants	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0225	0.0176	0.0470	0.0422	0.0118	0.0116
	(0.0447)	(0.0470)	(0.0504)	(0.0540)	(0.0554)	(0.0595)
N	1,688	1,688	1,150	1,150	793	793
Clusters	166	166	146	146	142	142
t+2	-0.0415	-0.0514*	-0.0455	-0.0556	-0.0566	-0.0761
	(0.0327)	(0.0310)	(0.0397)	(0.0368)	(0.0529)	(0.0509)
N	1,350	1,350	927	927	646	646
Clusters	167	167	146	146	144	144
t+3	-0.0151	-0.0073	0.0269	0.0293	0.0068	0.0313
	(0.0359)	(0.0356)	(0.0380)	(0.0373)	(0.0580)	(0.0596)
N	1,053	1,053	704	704	489	489
Clusters	166	166	137	137	132	132
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. OLS regression. Polynomial of order 1. Definition of peers according to reference group III, i.e. teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table 3.13: Comparison of the Effect of AEP certification on peers' one period ahead application rate by reference group

Dependent variable:	All		Closest to cut-off		Only one applicant	
Application rate among peers	(1)	(2)	(3)	(4)	(5)	(6)
Reference Group I (school)	0.0138***	0.0138***	0.0153***	0.0149***	0.0141***	0.0137***
	(0.0024)	(0.0024)	(0.0026)	(0.0026)	(0.0031)	(0.0031)
Reference group II (primary-secondary)	0.0146***	0.0145***	0.0156***	0.0153***	0.0150***	0.0148***
	(0.0022)	(0.0023)	(0.0025)	(0.0025)	(0.0031)	(0.0032)
Reference group III (grade)	0.0192***	0.0190***	0.0201***	0.0196***	0.0190***	0.0186***
	(0.0028)	(0.0029)	(0.0032)	(0.0033)	(0.0039)	(0.0040)
p-value $\beta_I = \beta_{II}$	0.220	0.237	0.608	0.568	0.217	0.133
p-value $\beta_{II} = \beta_{III}$	0.024	0.034	0.047	0.068	0.163	0.217
p-value $\beta_I = \beta_{III}$	0.019	0.028	0.060	0.081	0.101	0.128
p-value $\beta_I = \beta_{II} = \beta_{III}$	0.063	0.090	0.137	0.188	0.189	0.178
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Polynomial of order 1. Reference group I are teachers working at the same school as the applicant at the time of application. Reference group II are teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application. Reference group III are teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Appendix A

Chapter 1

A.1 A Model with Private Goods and Disutility of Violence

In this appendix, we solve the problem of the household with private goods and disutility of violence. Call q_s the private good of spouse s. As before, we normalize the prices of both goods to 1. The problem of the household is

$$\max_{q_f, q_m, l_f, d, v} \mu(v, \widetilde{\omega}_f) u^f(q^f, Q, v) + (1 - \mu(v, \widetilde{\omega}_f))\, u^m\,(q_m, Q)\,,$$

subject to

$$q_f + q_m + d = e^{\gamma(v)} l_f w_f + w_m + t_c,$$

and

$$Q = e^{\gamma(v)} F\,(d + t_k, (1 - l_f))\,.$$

Notice that now v enters directly into the utility of the female.

The optimality conditions of this problem are given by

$$\begin{aligned}
\frac{\partial Q}{\partial d}\left[\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial Q} + (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial Q}\right] &= \frac{1}{2}\left[\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial q^f} + (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q^m}\right],\\
\frac{\partial Q}{\partial 1 - l_f}\left[\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial Q} + (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial Q}\right] &= \frac{1}{2}\left[\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial q^f} + (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q^m}\right] e^{\gamma(v)} w_f,\\
\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial q^f} &= (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q^m},
\end{aligned}$$

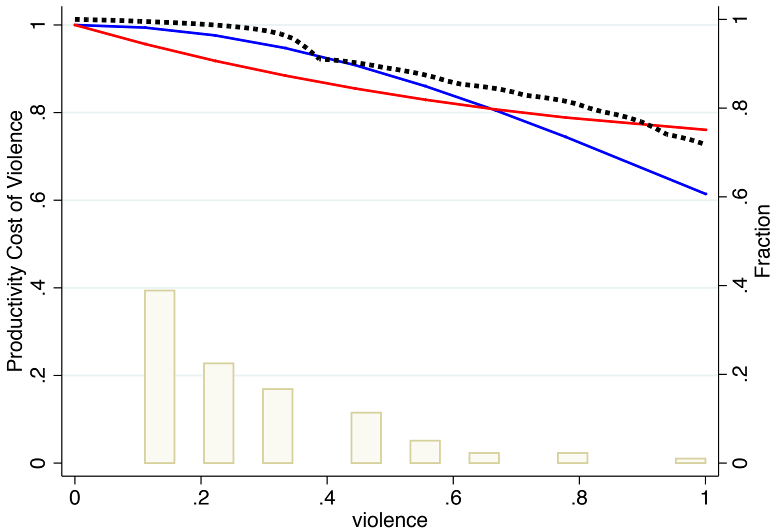
and

$$\begin{aligned}
\frac{\partial \mu(v, \widetilde{\omega}_f)}{\partial v}\Delta u_f^m &= \left[\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial Q} + (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial Q}\right] Q\frac{\partial \gamma(v)}{\partial v}\\
&+ \frac{1}{2}\left[\mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial q^f} + (1 - \mu(v, \widetilde{\omega}_f))\,\frac{\partial u^m}{\partial q^m}\right] e^{\gamma(v)} l_f w_f \frac{\partial \gamma(v)}{\partial v} + \mu(v, \widetilde{\omega}_f)\frac{\partial u^f}{\partial v},
\end{aligned}$$

where $\Delta u_f^m = u^m(Q, q) - u^f(Q, q)$. As in the model with only public goods and no dis-utility of violence, the level effect Δu_f^m generates differential gains of violence under different transfer regimes. Yet as the female experience an additional (direct) dis-utility from being abused, violence is more costly and less likely to occur (regardless of the transfer regime). As for the private goods, they implicitly decrease the cost of violence.

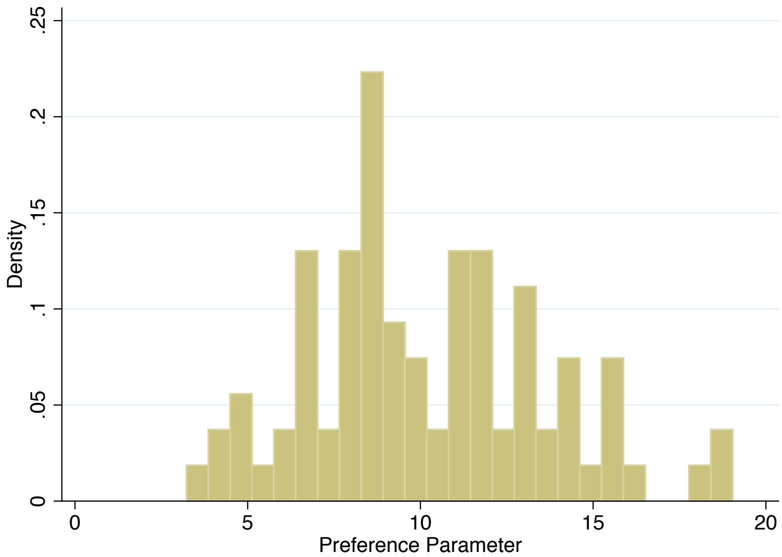
A.2 Figures

Figure A.1: Productivity Cost of Violence



Source: *Food, Cash or Voucher* (World Food Programme).
Notes: The bars represent a histogram of the intensity of violence, among the violent households. The black dotted line corresponds to a local polynomial regression of the LHS of Equation (1.5.3) on violence. The red dashed line depicts the predicted $\widehat{\gamma}(v)$ from column (1) in Table A.4. The blue solid line depicts the predicted $\widehat{\gamma}(v)$ from column (2) in Table A.4. Since the concavity of $\gamma(v)$ is required for the existence of interior solutions (see Section 1.2), column (2) in Table A.4 is the most preferred specification.

Figure A.2: Disagreement in Preference Parameter



Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Distribution of the α_i^f. The α_i^fs were estimated through Equation (1.5.5)

A.3 Tables

Table A.1: Sample Selection Criteria

	N
Not aged 15-70 (at baseline)	105
Not at union (at baseline)	764
No woman head of household or spouse (at baseline)	49
Not alone at time on interview (at baseline)	20
No woman head of household or spouse (at follow-up)	135
Not alone at time on interview (at follow-up)	39
Different respondent (at follow-up)	15
Same women, at union at baseline, aged 15-70, alone at time of interviews	1,230
Sample for structural estimation	1,210

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: The structural estimation sample further requires complete information on complete information on violence, labor income, food expenditure, hours worked and hours devoted to household work, both at baseline and follow-up.

Table A.2: Female Wages

	Log Wages	Selection
Age	0.0317*	0.0030
	(0.0165)	(0.0273)
Age, squared	-0.0003	0.0001
	(0.0002)	(0.0003)
Female's education years	0.0325**	0.0011
	(0.0158)	(0.0199)
Female with secondary education or more	0.0755	0.0169
	(0.1081)	(0.1617)
Female's hours of work a day	-0.0319	0.4806***
	(0.1758)	(0.0491)
Female's hours of work a day, squared	-0.0067	-0.0316***
	(0.0119)	(0.0040)
Carchi	-0.1489*	0.0252
	(0.0784)	(0.1243)
Married couple		0.2006*
		(0.1081)
No. children form 0 to 5		-0.1253
		(0.0910)
No. children from 6 to 14		-0.0409
		(0.0541)
Constant	-0.6354	-0.7589
	(0.9177)	(0.5322)
Lambda	0.90	
Clusters	141	
N	922	

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Heckman Two-Step procedure. Estimation on the sample of females working in the labor market. Selection on violence. Exclusion restrictions: marital status and number of children from 0 to 5. Bootstrapped standard errors in parenthesis. Bootstrapped sample at the household level, taking with cluster (cluster) and strata (province and treatment arm).

Table A.3: Polynomials of v for $\gamma(v)$

	Adjusted R2		R2	
Polynomial (p)	1-p	p	1-p	p
1	0.0003	0.0003	0.0015	0.0015
2	-0.0008	-0.0001	0.0015	0.0011
3	-0.0020	-0.0005	0.0015	0.0007
4	-0.0029	-0.0007	0.0018	0.0004
5	-0.0040	-0.0009	0.0019	0.0003
6	-0.0043	-0.0010	0.0027	0.0002
7	-0.0054	-0.0011	0.0028	0.0001
8	-0.0054	-0.0011	0.0029	0.0001
9	-0.0053	-0.0011	0.0029	0.0000
10	-0.0053	-0.0011	0.0029	0.0000

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Selection of the polynomial of v to estimate $\gamma(v)$ through Equation (1.5.3).

Table A.4: Productivity Cost of Violence

	LHS Equation 1.5.3	
	(1)	(2)
v	-0.416	
	(0.701)	
v^2	0.142	-0.487
	(1.214)	(0.478)
β_0	1.870***	1.861***
	(0.051)	(0.047)
Adjusted R^2	-0.0008	-0.0001
Clusters	139	139
N	859	859

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Clustered standard errors in parenthesis. Estimation of Equation (1.5.3), according to the best-fit polynomial in Table A.3.
* Indicates statistical significance at 10%.
* Indicates statistical significance at 5%.
** Indicates statistical significance at 1%.

Table A.5: Effect of Violence on Female Weight

	LHS Equation 1.5.6
c	3.270***
	(0.025)
Clusters	119
N	393

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Clustered standard errors in parenthesis. Estimation of Equation (1.5.6).
* Indicates statistical significance at 10%.
* Indicates statistical significance at 5%.
** Indicates statistical significance at 1%.

Table A.6: Household Observable Characteristics Predicting Disagreement in Preferences

Observable Characteristic	α_i^f
No. of household members	1.132***
	(0.281)
Male head of household	2.148
	(1.479)
Married couple	0.154
	(0.735)
No. children form 0 to 5	-0.692
	(0.471)
No. children from 6 to 14	-0.711
	(0.440)
Female	-0.011
	(0.035)
Couple age difference	0.100
	(0.065)
Male education years	-0.089
	(0.102)
Female education years	-0.098
	(0.142)
Female more educated than male	0.257
	(1.063)
Female employed	1.007
	(1.043)
Female labor income a day	0.107
	(0.113)
Female hours of work a day	0.024
	(0.130)
Female hours of household work a day	-0.046
	(0.086)
Male employed	1.171
	(2.020)
Male labor income a day	0.073
	(0.053)
Male hours of work a day	0.416***
	(0.153)

Source: *Food, Cash or Voucher* (World Food Programme).
Notes: Clustered standard at the cluster level in parenthesis.
* Indicates statistical significance at 10%.
* Indicates statistical significance at 5%.
** Indicates statistical significance at 1%.

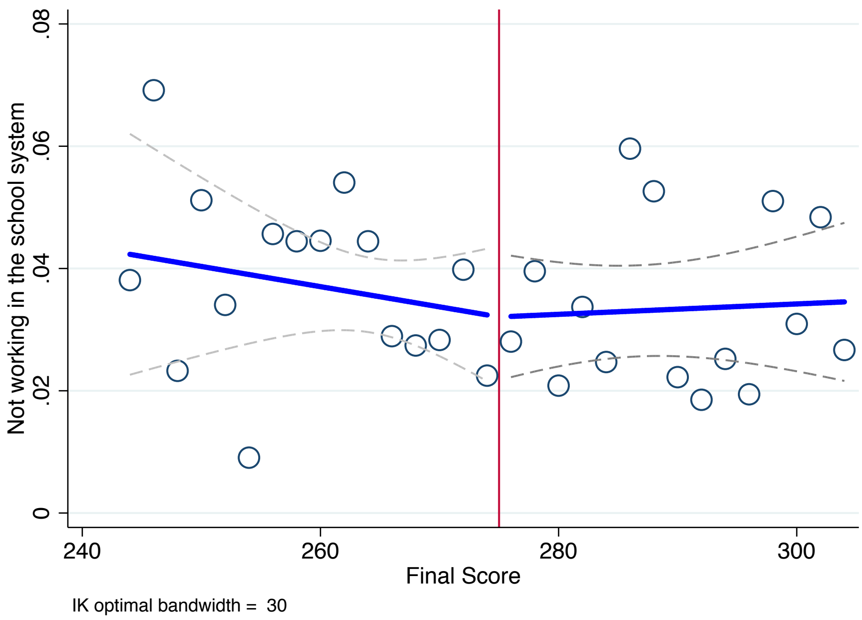
Appendix B

Chapter 2

B.1 AEP

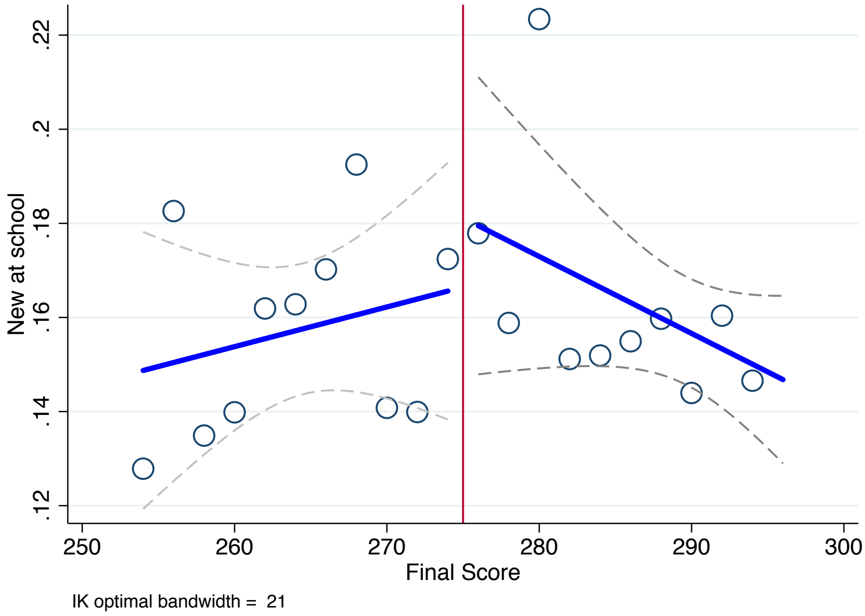
B.1.1 Figures

Figure B.1: AEP effects on Retention and Labor Supply



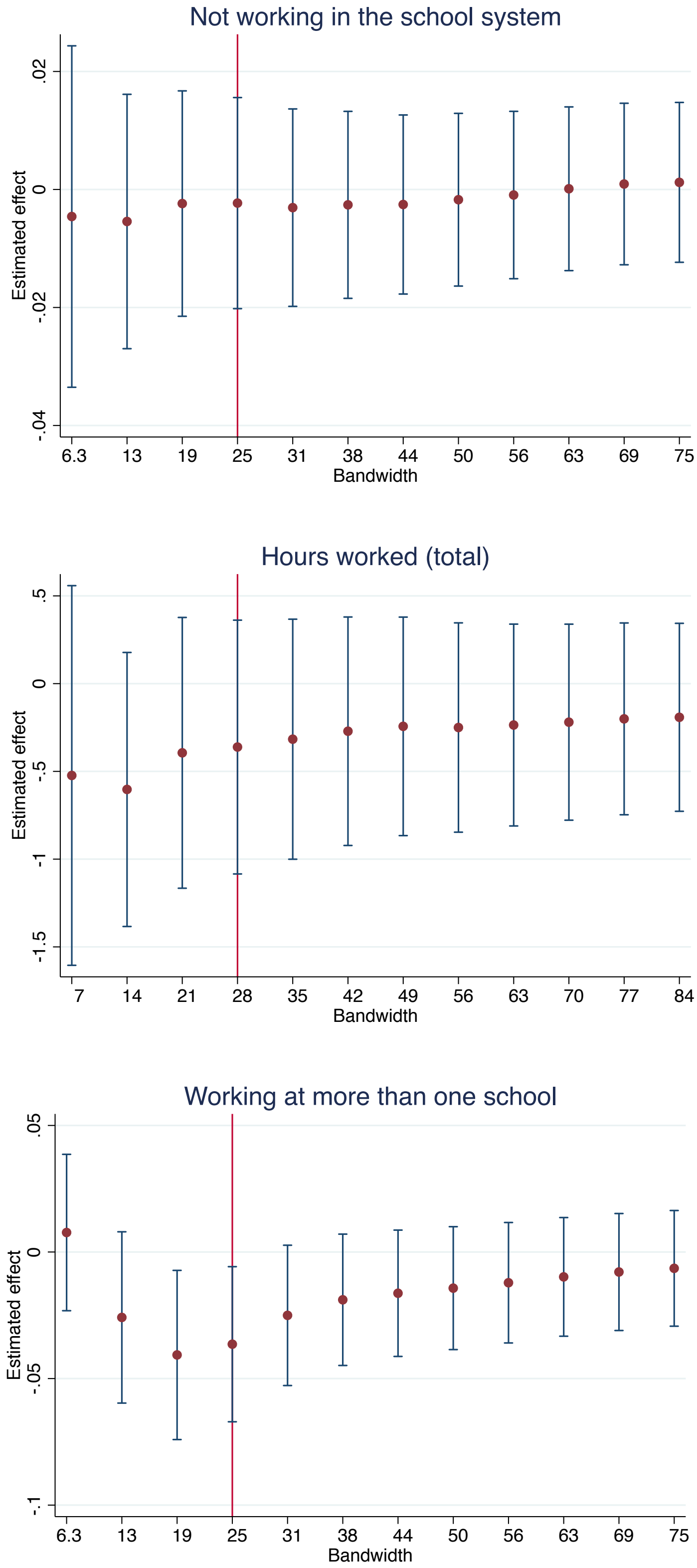
Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: The circles represent mean of the outcome variable within bins of size 4 of the final score. Solid lines show fitted values of a piecewise linear polynomial. Doted lines represent the confidence intervals, for errors clustered at the final score cell level. Imbens and Kalyanaraman (2011)'s optimal bandwidth.

Figure B.2: AEP effects on Between School-Mobility



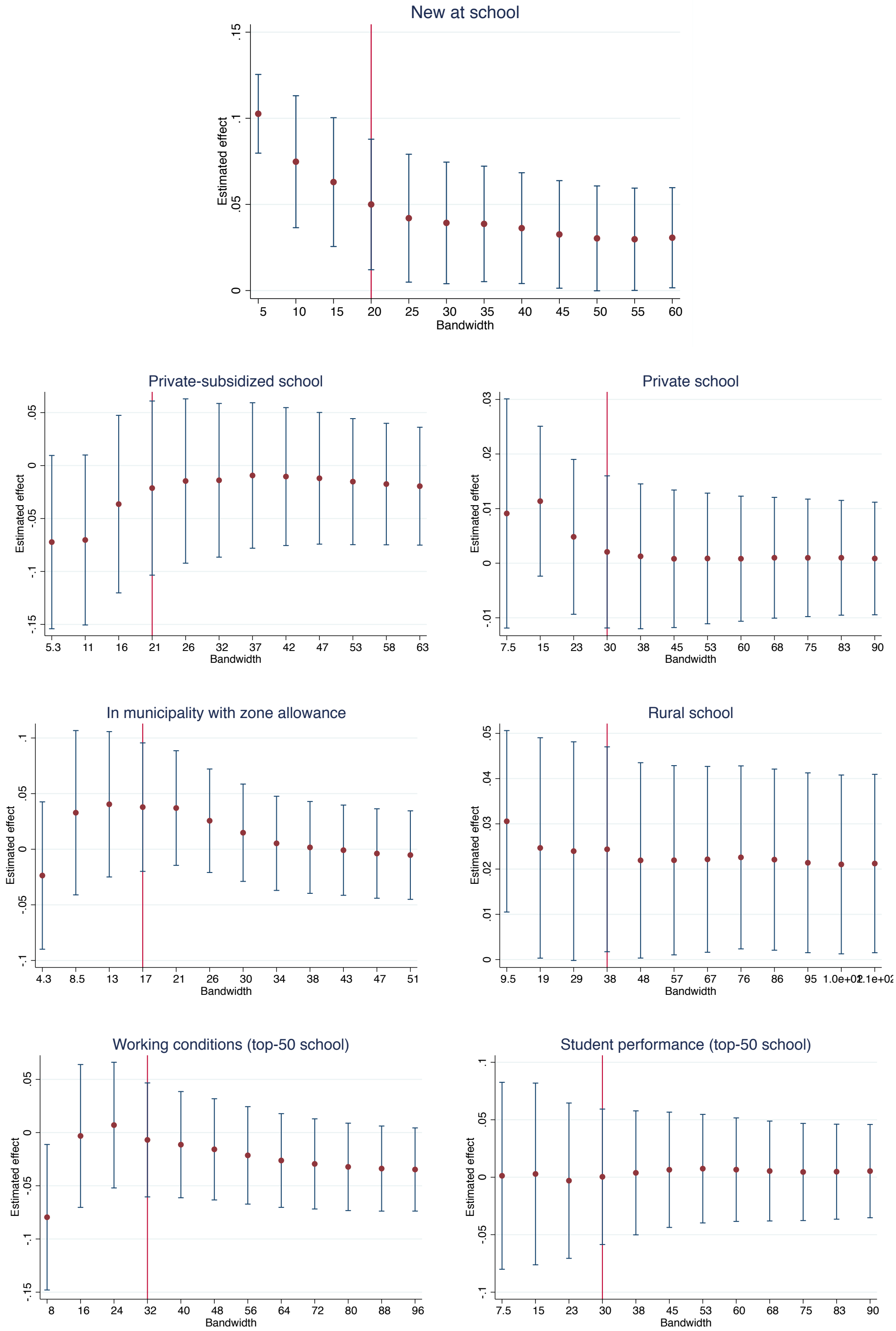
Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: The circles represent mean of the outcome variable within bins of size 4 of the final score. Solid lines show fitted values of a piecewise linear polynomial. Doted lines represent the confidence intervals, for errors clustered at the final score cell level. Imbens and Kalyanaraman (2011)'s optimal bandwidth.

Figure B.3: AEP effects on Turnover and Labor Supply by Application Wave



Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Each point represent the point estimate of a fully non-parametric specification for each application bandwidth separately, 2 years after application. The blue lines represent the 95 confidence intervals. The red line indicates the optimal bandwidth following Imbens and Kalyanaraman (2011).

Figure B.4: AEP effects on Between School Mobility by Application Wave



Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Each point represent the point estimate of a fully non-parametric specification for each application bandwidth separately, 2 years after application. The blue lines represent the 95 confidence intervals. The red line indicates the optimal bandwidt following Imbens and Kalyanaraman (2011).

B.1.2 Tables

Table B.1: Power Calculations

AEP	All Teachers	Years of experience		
Dependent variable		0-11	11-21	21+
Panel A Labor Supply				
Not working in the school system	0.943	0.980	0.937	0.995
Hours worked (total)	0.998	0.996	0.992	1.000
Working at more than one school	0.990	0.984	0.984	0.918
Panel B Between-School Mobility				
New at school	0.436	0.682	0.865	0.687
Private-subsidized school	0.997	1.000	0.811	1.000
Private school	0.896	0.681	1.000	0.999
In municipality with zone allowance	0.994	0.988	0.831	1.000
Rural school	0.714	0.734	0.784	0.864
Working conditions (top-50 school)	0.997	0.993	0.939	0.999
Student performance (top-50 school)	0.955	0.957	0.541	1.000

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Power calculations accounting for clustering at the cell level (see Schochet (2009)).

Table B.2: Two-way clustering: AEP effects on Retention and Labor Supply

	All Teachers		0-11 years		12-21 years		22 +years	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not working in the school system	0.0038	0.0036	-0.0018	0.0063	0.0078	0.0083	-0.0061	-0.0014
	(0.0085)	(0.0098)	(0.0166)	(0.0179)	(0.0178)	(0.0201)	(0.0098)	(0.0101)
N	9,311	9,311	3,756	3,756	2,872	2,872	2,683	2,683
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Hours worked (total)	-0.6344*	-0.5327	-0.3629	-0.3012	-0.2493	-0.4344	-1.5792**	-1.1180*
	(0.3315)	(0.3243)	(0.5303)	(0.5983)	(0.5391)	(0.4852)	(0.6828)	(0.6530)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Working at more than one school	-0.0047	-0.0032	-0.0041	-0.0054	-0.0048	-0.0010	0.0159	0.0200
	(0.0146)	(0.0141)	(0.0222)	(0.0229)	(0.0259)	(0.0256)	(0.0294)	(0.0315)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AEP data for teachers' applying to AEP waves 2003-2011, 2 years after application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Odd-columns present the estimates of equation (2.5.1). Even columns present the estimates of equation (2.5.1) and add controls interacted with wave fixed effects. Controls include gender, age, degree in education, years of experience, teaching at a single school, hours worked, receiving AVDI, rural school, private-subsidized school, working conditions (top-50 school), student performance (top-50 school), SNED awarding school and education level of the main job: primary; all at time of application and excluding the outcome variable at time of application. Robust standard errors, adjusted for clustering in final score cells and school level following Cameron et al. (2011)Cameron et al. (2011), in parenthesis.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Table B.3: Two-way clustering: AEP effects on Between-School Mobility

	All Teachers		0-11 years		12-21 years		22 +years	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New at school	0.0447***	0.0434**	0.0470	0.0530	0.0296	0.0338	0.0394	0.0346
	(0.0172)	(0.0185)	(0.0318)	(0.0364)	(0.0344)	(0.0353)	(0.0262)	(0.0296)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Private-subsidized school	-0.0350	-0.0424*	-0.0799*	-0.0803**	0.0438	0.0321	-0.0848*	-0.0738*
	(0.0330)	(0.0230)	(0.0455)	(0.0367)	(0.0400)	(0.0316)	(0.0435)	(0.0414)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Private school	0.0006	0.0005	0.0197	0.0184	-0.0242***	-0.0268***	-0.0034	-0.0037
	(0.0063)	(0.0062)	(0.0135)	(0.0148)	(0.0069)	(0.0079)	(0.0030)	(0.0031)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
In municipality with zone allowance	-0.0153	-0.0092	-0.0110	-0.0016	0.0488	0.0492	-0.1000**	-0.0753
	(0.0243)	(0.0255)	(0.0377)	(0.0390)	(0.0482)	(0.0493)	(0.0445)	(0.0468)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Rural school	0.0245*	0.0219	0.0292	0.0275	0.0301	0.0319	0.0222	0.0124
	(0.0145)	(0.0140)	(0.0217)	(0.0228)	(0.0263)	(0.0285)	(0.0255)	(0.0258)
N	8,937	8,937	3,543	3,543	2,768	2,768	2,626	2,626
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Working conditions (top-50 school)	-0.0180	-0.0245	-0.0197	-0.0254	0.0179	-0.0029	-0.0522	-0.0397
	(0.0276)	(0.0254)	(0.0396)	(0.0418)	(0.0419)	(0.0426)	(0.0398)	(0.0370)
N	8,831	8,831	3,469	3,469	2,741	2,741	2,621	2,621
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725
Student performance (top-50 school)	0.0032	-0.0047	0.0084	-0.0019	0.0764*	0.0605	-0.0954*	-0.0749
	(0.0292)	(0.0285)	(0.0357)	(0.0364)	(0.0408)	(0.0412)	(0.0563)	(0.0497)
N	8,831	8,831	3,469	3,469	2,741	2,741	2,621	2,621
Score clusters	230	230	187	187	198	198	206	206
School clusters	3,894	3,894	2,235	2,235	1,859	1,859	1,725	1,725

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AEP data for teachers' applying to AEP waves 2003-2011, 2 years after application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Odd-columns present the estimates of equation (2.5.1). Even columns present the estimates of equation (2.5.1) and add controls interacted with wave fixed effects. Controls include gender, age, degree in education, years of experience, teaching at a single school, hours worked, receiving AVDI, rural school, private-subsidized school, working conditions (top-50 school), student performance (top-50 school), SNED awarding school and education level of the main job: primary; all at time of application and excluding the outcome variable at time of application. Robust standard errors, adjusted for clustering in final score cells and school level following Cameron et al. (2011), in parenthesis.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

B.2 AVDI

In 2004 the Ministry of Education implemented a compulsory examination for municipal school teachers. Every 4 years, teachers of municipal schools are assessed through a written examination (*Evaluación Docente* (EV)). Municipal school teachers with an outstanding evaluation (EV) can apply to a performance award: *Asignación Variable al Desempeño Individual* or AVDI (following its Spanish acronym). For this purpose, teachers must take the same knowledge test than for AEP (no portfolio is required). The results of these tests are combined and those who score above a threshold of 275 receive a monetary annual compensation that lasts between two to four years, depending on when they are required to re-take the EV. For the average teacher, the AVDI award would be equivalent to a 6 to 10 percent increase of her monthly pay.

We use the administrative we already described in Section 2.3. Figure B.5 presents a sample flowchart. We start with the 31,237 teachers that applied for the first time for an AVDI award between 2004 and 2011. Further, we restrict to individuals who applied for the award in primary or secondary education. We match this data with administrative records and restrict our analysis to individuals that at the time of application are at least four years away from the retirement age (i.e., 56 for females and 61 for males). We focus on the sample of 23,868 not currently applying to AEP.

We start by showing that the assignment rule was strictly enforced. In Figure B.6, we plot the mean of a variable that takes the value of 1 if an individual has an AVDI award and 0 otherwise, for each possible score cell (circles). There is clearly a sharp discontinuity. Those who obtained the award have an aggregate score of 275 or more. In Table B.4, we present the awardee rates by year. We divide the data in two samples, Panel A has the 23,870 teachers from our benchmark sample and Panel B has the 31,237 first time applicants. The table confirms the information on the graph: compliance with the allocation rule is 100 percent, regardless of the application wave or sample. Focusing on Panel A, 31 percent of the teachers that apply for AVDI obtained it.

In the first column of Table B.6, we present average information for all employed teachers in the Municipal School System during the 2004-2014 period. In the second column, we present the same information but only for those who have applied to AVDI during the 2004-2014 window. Beginning with basic demographic and qualification variables, we observe that over the 2004-2014 period, the average Chilean teacher in a municipal school is a 47 years old woman with a degree in education and 21 years of teaching experience, working 35 hours a week. Around 89 percent of the teachers work at a single school, 80 percent work as primary school teachers, and 11 percent hold a managerial position. Every year, 11 percent of the teachers change schools and 4 percent move to a different municipality. Around 49 percent of the teachers work in municipalities considered as isolated and are monetarily compensated with an allowance. Around 30 percent of the teachers work in schools ranked in the top 50 percentile in terms of working conditions and 47 percent in schools ranked in the top 50 percentile in terms of student performance.

In the third and fourth column of Table B.6, we describe the sample at the first time of application to AVDI and two-years after. Two-years after applying to AVDI, 3 percent of the teachers are not employed in the school system, 0.2 percent work in a private school, 6 percent change from municipality, and 14 percent moved to a different school from the one they were at when applying.

We exploit the sharp discontinuity in the allocation of the award for teachers with 275 points or more in the aggregate evaluation score to estimate the causal impact of an

AVDI award. Like in section 2.5 we implement the regression discontinuity design using equation 2.5.1.

In Figure B.7, we plot the histogram of the final score for the pooled sample of applicants. In column one of Table B.5, we present the results of testing for a discontinuity using the McCrary (2008) test and Frandsen (2014)'s approach for variables with discrete support. Table B.5 also presents the McCrary (2008) and Frandsen (2014)'s p-values for each AVDI wave. Figure B.7 shows a clear spike in the final score before the cut-off of 275. Not surprisingly, we reject the no discontinuity hypothesis in several years for both tests.

In Table B.7 we provide evidence on the continuity of baseline characteristics around the threshold. We estimate equation (2.5.1) using as outcome variables the characteristics of the teachers and their schools, at time of application to AVDI. The number of the column in this table indicates the order of the piece-wise polynomial of the score used in each specification. Unlike the case of AEP, there are systematic differences in the baseline variables and we tend to reject the null hypothesis of continuity for the 14 variables using the joint (Wald) test.

We have no explanation of why there might be manipulation of the data at the left of the cut-off (i.e., this stops teachers from receiving the award). We use a second degree polynomial as the benchmark specification and we also control for baseline variables interacted with wave fixed effects. The latter results are preferred. However, the causal interpretation should be considered with the appropriate caveats.

First, we look at the effect of receiving an AVDI award on teacher retention. In Table B.8, we present OLS estimates of equation 2.5.1 in the odd columns. In the even columns we add controls for demographics, qualifications, labor outcomes, and main school's characteristics at the time of application. We show estimates by three experience levels: 0-11 years, 12-21 years and 22 or more years of experience. The estimates show a small positive and statistically significant on total hours worked. Second, in B.9 we look at the pattern of mobility between schools which could have been caused by the program. We find no systematic evidence that the receiving an AVDI award affected between school mobility.

Finally, we also explore the effect of the receiving to both AEP and AVDI award (results available upon request from the authors). On average, awardees both programs will receive a 12 percent increase in their salary, yet not even this wages increase seems to alter teachers' behavior at the extensive margin.

B.2.1 Figures

Figure B.5: Flowchart for AVDI sample

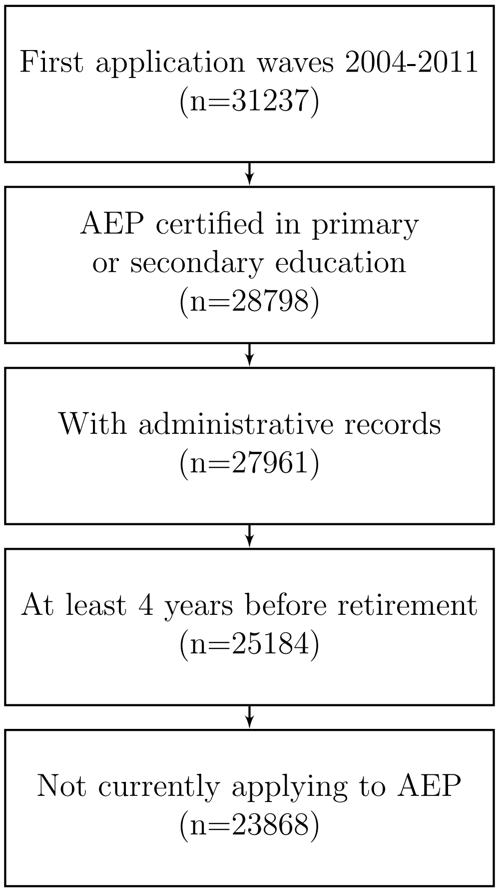
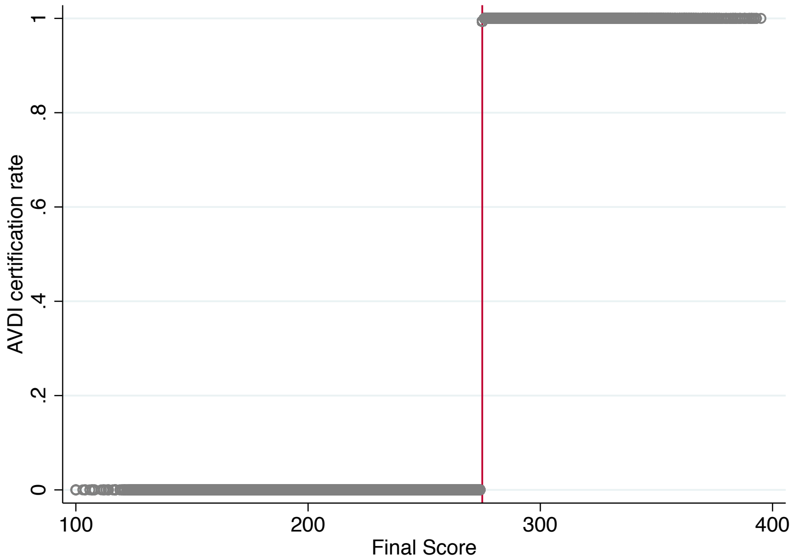
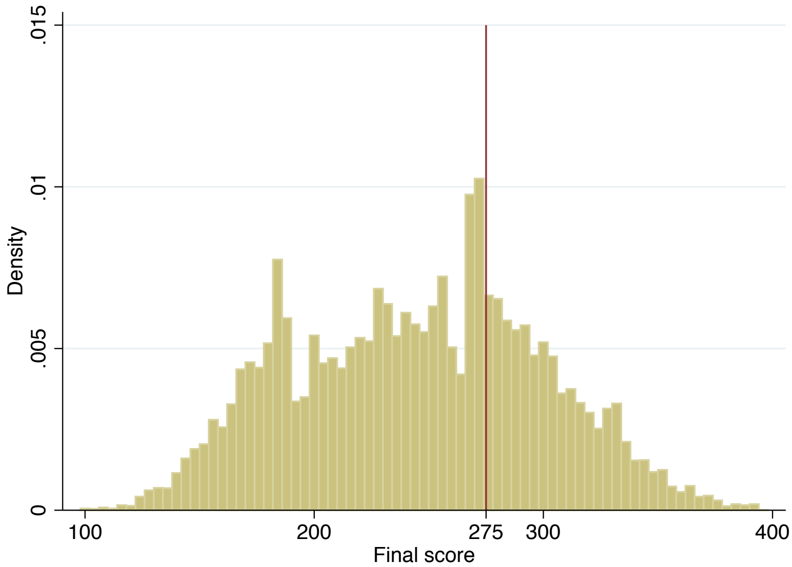


Figure B.6: AVDI Assignment Rule



Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Circles represent the proportion of applicants passing the exam within each final score cell.

Figure B.7: Distribution of the AVDI Final Score



Source: Own calculations based on data from the Ministry of Education (Chile)

B.2.2 Tables

Table B.4: Proportions of Applicants Receiving the AVDI Award over Time

	All	2004	2005	2006	2007	2008	2009	2010	2011
	Panel A Final Sample								
AVDI certification rate (%)	31	28	29	23	30	33	33	34	32
Compliance with allocation rule (%)	100	100	100	100	100	100	100	100	100
N	23,868	918	703	2,601	5,146	3,671	3,136	4,624	3,069
	Panel B First time applicants								
AVDI certification rate (%)	30	27	28	22	29	32	30	33	31
Compliance with the 275 allocation rule (%)	100	100	100	100	100	100	100	100	100
N	31,237	1,191	859	3,240	6,486	4,348	5,375	6,153	3,585

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Notes: Standard deviation in parenthesis. Data for teachers' applying to waves 2004-2011.

Table B.5: Test for Continuity of the Final Score

	All	2004	2005	2006	2007	2008	2009	2010	2011
McCrary test p-value	0.000	0.796	0.823	0.789	0.587	0.630	0.000	0.000	0.000
Frandsen Discrete test p-value	0.559	0.307	0.286	0.135	0.017	0.111	0.587	0.688	0.034

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: McCrary (2008) test at the 275 cut-off, using a bandwidth of 30 and bin size 1.

Table B.6: Descriptive Statistics

	2004-2014		AVDI applicants	
	Voucher System Teachers	AVDI Applicants	At time of application	2-years after
Male	0.300	0.270	0.288	0.286
	(0.458)	(0.444)	(0.453)	(0.452)
Age	47.526	47.148	44.531	46.665
	(11.302)	(9.625)	(9.373)	(9.318)
Degree in education	0.961	0.983	0.987	0.996
	(0.194)	(0.131)	(0.114)	(0.062)
Years of experience	21.146	21.272	18.243	20.769
	(12.750)	(11.126)	(10.634)	(10.505)
Not working in the school system				0.027
				(0.162)
Hours worked (total)	35.301	35.803	35.504	36.414
	(7.802)	(6.733)	(6.710)	(6.651)
Main job: primary school teacher	0.807	0.824	0.752	0.726
	(0.395)	(0.381)	(0.432)	(0.446)
Working at more than one school	0.114	0.117	0.124	0.108
	(0.318)	(0.321)	(0.329)	(0.310)
In a managerial job	0.112	0.055	0.041	0.072
	(0.316)	(0.228)	(0.197)	(0.259)
AVDI applicant (ever)	0.334			
	(0.472)			
Currently applying to AVDI				0.004
				(0.060)
Receiving AVDI	0.053	0.156		0.311
	(0.224)	(0.363)		(0.463)
AEP applicant (ever)	0.070	0.159	0.110	0.110
	(0.256)	(0.365)	(0.313)	(0.313)
Currently applying to AEP				0.011
				(0.107)
Receiving AEP	0.014	0.034	0.030	0.037
	(0.117)	(0.181)	(0.171)	(0.188)
New at school	0.112	0.084		0.147
	(0.315)	(0.278)		(0.354)
Private-subsidized school	0.000	0.039	0.014	0.027
	(0.000)	(0.194)	(0.119)	(0.163)
Private school	0.000	0.006		0.002
	(0.000)	(0.077)		(0.046)
Working conditions (top-50 school)	0.301	0.322	0.312	0.323
	(0.459)	(0.467)	(0.463)	(0.468)
Student performance (top-50 school)	0.466	0.526	0.499	0.517
	(0.499)	(0.499)	(0.500)	(0.500)
SNED awarded school	0.305	0.336	0.348	0.353
	(0.460)	(0.472)	(0.476)	(0.478)
Change of municipality	0.043	0.030		0.062
	(0.202)	(0.171)		(0.241)
Rural school	0.237	0.260	0.257	0.248
	(0.425)	(0.439)	(0.437)	(0.432)
In municipality with zone allowance	0.492	0.499	0.498	0.499
	(0.500)	(0.500)	(0.500)	(0.500)
N	757,831	259,434	23,868	23,868

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: Standard deviation in parenthesis. For the 2003-2014 period, *New at school* stands for whether or not the teacher was teaching at that particular school in the previous year. For the AEP applicants 2 years after application, *New at school* is a dummy taking the value of 1 if the school is different from the school at time of application. Except *Not working in the school system*, the dependent variables for teachers not working in the school system 2 years after application are coded as missing.

Table B.7: Balance at Baseline AVDI

AVDI	Degree of polynomial			
Dependent Variable	(1)	(2)	(3)	(4)
Male	0.021*	0.022*	0.024	0.023
	(0.012)	(0.013)	(0.016)	(0.018)
Age	-0.326	-0.157	-0.058	0.395
	(0.255)	(0.354)	(0.432)	(0.492)
Degree in education	0.006**	0.005	0.009**	0.011*
	(0.003)	(0.003)	(0.004)	(0.006)
Years of experience	-0.144	-0.124	0.061	0.287
	(0.265)	(0.365)	(0.436)	(0.505)
Hours worked (total)	0.032	0.229	0.300	0.314
	(0.157)	(0.199)	(0.245)	(0.306)
Working at more than one school	-0.001	0.007	-0.001	0.006
	(0.009)	(0.012)	(0.014)	(0.018)
In a managerial job	0.009**	0.002	0.007	0.008
	(0.005)	(0.007)	(0.008)	(0.009)
Main job: primary school teacher	0.045***	0.010	-0.005	-0.006
	(0.011)	(0.014)	(0.019)	(0.023)
Receiving AEP	0.012**	-0.004	-0.003	-0.003
	(0.005)	(0.006)	(0.007)	(0.008)
Working conditions (top-50 school)	0.016	0.028**	0.016	0.034*
	(0.010)	(0.012)	(0.015)	(0.018)
Student performance (top-50 school)	-0.020*	-0.003	-0.011	-0.018
	(0.011)	(0.014)	(0.018)	(0.020)
SNED awarded school	-0.013	-0.007	-0.018	-0.022
	(0.014)	(0.020)	(0.027)	(0.032)
Rural school	0.029***	0.023	0.019	0.017
	(0.011)	(0.016)	(0.021)	(0.026)
In municipality with zone allowance	0.011	0.001	0.009	0.026
	(0.013)	(0.018)	(0.023)	(0.027)
Wald test p-value	0.0027	0.0366	0.4410	0.1002

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AVDI data for teachers' applying to AVDI waves 2004-2011, at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Robust standard errors, adjusted for clustering in final score cells, in parenthesis. Column numbers indicate the order of the polynomial on the score centered around 275.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Table B.8: AVDI effects on Retention and Labor Supply

	All Teachers		0-11 years		12-21 years		22 +years	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not working in the school system	0.0002	0.0011	0.0033	0.0066	-0.0134	-0.0130	0.0024	0.0037
	(0.0044)	(0.0043)	(0.0119)	(0.0115)	(0.0113)	(0.0109)	(0.0051)	(0.0053)
N	23,868	23,868	7,368	7,368	5,557	5,557	10,943	10,943
Clusters	288	288	272	272	272	272	283	283
Hours worked (total)	0.4070**	0.4356***	0.5958	0.5131	0.6114	0.5918	0.1911	0.2446
	(0.1706)	(0.1459)	(0.3711)	(0.3480)	(0.4236)	(0.3837)	(0.3106)	(0.2496)
N	23,228	23,228	6,994	6,994	5,423	5,423	10,811	10,811
Clusters	288	288	272	272	272	272	283	283
Working at more than one school	0.0013	0.0020	-0.0172	-0.0208	0.0056	0.0042	0.0133	0.0180
	(0.0098)	(0.0098)	(0.0171)	(0.0176)	(0.0199)	(0.0204)	(0.0162)	(0.0160)
N	23,228	23,228	6,994	6,994	5,423	5,423	10,811	10,811
Clusters	288	288	272	272	272	272	283	283

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AVDI data for teachers' applying to AVDI waves 2004-2011, 2 years after application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Odd-columns present the estimates of equation (2.5.1). Even columns present the estimates of equation (2.5.1) and add controls interacted with wave fixed effects. Controls include gender, age, degree in education, years of experience, teaching at a single school, hours worked, receiving AEP, rural school, private-subsidized school, working conditions (top-50 school), student performance (top-50 school), SNED awarding school and education level of the main job: primary; all at time of application and excluding the outcome variable at time of application. Robust standard errors, adjusted for clustering in final score cells, in parenthesis.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Table B.9: AVDI effects on Between-School Mobility

	All Teachers		0-11 years		12-21 years		22 +years	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New at school	-0.0013	-0.0043	-0.0112	-0.0178	0.0155	0.0168	-0.0021	-0.0037
	(0.0111)	(0.0110)	(0.0214)	(0.0210)	(0.0282)	(0.0275)	(0.0131)	(0.0124)
N	23,228	23,228	6,994	6,994	5,423	5,423	10,811	10,811
Clusters	288	288	272	272	272	272	283	283
Private-subsidized school	-0.0065	-0.0068	-0.0169	-0.0206	0.0001	0.0002	-0.0018	-0.0016
	(0.0054)	(0.0054)	(0.0130)	(0.0134)	(0.0082)	(0.0089)	(0.0052)	(0.0051)
N	23,228	23,228	6,994	6,994	5,423	5,423	10,811	10,811
Clusters	288	288	272	272	272	272	283	283
Private school	0.0001	0.0002	-0.0016	-0.0015	0.0049	0.0051	-0.0011	-0.0011
	(0.0017)	(0.0018)	(0.0041)	(0.0042)	(0.0037)	(0.0037)	(0.0008)	(0.0008)
N	23,228	23,228	6,994	6,994	5,423	5,423	10,811	10,811
Clusters	288	288	272	272	272	272	283	283
In municipality with zone allowance	-0.0002	-0.0037	0.0190	0.0127	-0.0445	-0.0525	0.0000	0.0049
	(0.0181)	(0.0182)	(0.0300)	(0.0311)	(0.0322)	(0.0332)	(0.0253)	(0.0254)
N	23,224	23,224	6,994	6,994	5,422	5,422	10,808	10,808
Clusters	288	288	272	272	272	272	283	283
Rural school	0.0275**	0.0214*	0.0221	0.0219	0.0301	0.0224	0.0323	0.0159
	(0.0138)	(0.0129)	(0.0242)	(0.0248)	(0.0292)	(0.0268)	(0.0201)	(0.0186)
N	23,228	23,228	6,994	6,994	5,423	5,423	10,811	10,811
Clusters	288	288	272	272	272	272	283	283
Working conditions (top-50 school)	0.0079	0.0062	-0.0406*	-0.0343	0.0833***	0.0799***	0.0003	-0.0036
	(0.0108)	(0.0108)	(0.0229)	(0.0234)	(0.0283)	(0.0282)	(0.0198)	(0.0192)
N	23,151	23,151	6,943	6,943	5,406	5,406	10,802	10,802
Clusters	288	288	272	272	272	272	283	283
Student performance (top-50 school)	0.0098	0.0121	-0.0053	0.0276	0.0173	0.0162	0.0210	-0.0014
	(0.0139)	(0.0133)	(0.0238)	(0.0238)	(0.0310)	(0.0312)	(0.0210)	(0.0181)
N	23,151	23,151	6,943	6,943	5,406	5,406	10,802	10,802
Clusters	288	288	272	272	272	272	283	283

Source: Own calculations based on data from the Ministry of Education (Chile)
Notes: OLS regression. Dependent variable for teachers not working in the school system coded as missing. AVDI data for teachers' applying to AVDI waves 2004-2011, 2 years after application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piece-wise polynomial of the final score. Odd-columns present the estimates of equation (2.5.1). Even columns present the estimates of equation (2.5.1) and add controls interacted with wave fixed effects. Controls include gender, age, degree in education, years of experience, teaching at a single school, hours worked, receiving AEP, rural school, private-subsidized school, working conditions (top-50 school), student performance (top-50 school), SNED awarding school and education level of the main job: primary; all at time of application and excluding the outcome variable at time of application. Robust standard errors, adjusted for clustering in final score cells, in parenthesis.
* Indicates statistical significance at 10%
** Indicates statistical significance at 5%
*** Indicates statistical significance at 1%.

Appendix C

Chapter 3

C.1 Example

Consider the case of applicants 1 and 2, and their colleagues a, b and c. Applicant 1 and 2 teach in the same school, but applicant 1 is a primary school teacher, while applicant 2 is a secondary school teacher. As for their colleagues, teachers a and b teach at primary school, and teachers b and c teach at secondary school.

Perfectly Overlapping Reference Groups

Following the quasi-random assignment argument and defining peers according to reference group 1, the application decision of teachers a, b, and c at time $t+1$ are given by

$$\begin{aligned} y_a &= \alpha_a + \beta_1 D_1 + \gamma_1 s_1 + \beta_2 D_2 + \gamma_2 s_2 + \epsilon_a, \\ y_b &= \alpha_b + \beta_1 D_1 + \gamma_1 s_1 + \beta_2 D_2 + \gamma_2 s_2 + \epsilon_b, \quad \text{and} \\ y_c &= \alpha_c + \beta_1 D_1 + \gamma_1 s_1 + \beta_2 D_2 + \gamma_2 s_2 + \epsilon_c, \end{aligned}$$

where we drop the time index to ease the notation. In this case, the peers' application rate for 1 and 2 are exactly the same,

$$R_1 = R_2 = \frac{y_a + y_b + y_c}{3} \quad = \quad \overline{\alpha} + \beta_1 D_1 + \gamma_1 s_1 + \beta_2 D_2 + \gamma_2 s_2 + \overline{\epsilon}.$$

and would be estimated through

$$\begin{aligned} R_1 &= \overline{\alpha} + \beta D_1 + \gamma s_1 + \underbrace{\beta D_2 + \gamma s_2 + \overline{\epsilon}}_{\varepsilon_1}, \quad \text{and} \\ R_2 &= \overline{\alpha} + \beta D_2 + \gamma s_2 + \underbrace{\beta D_1 + \gamma s_1 + \overline{\epsilon}}_{\varepsilon_2}. \end{aligned}$$

Since the certification is quasi-randomly assigned around the cut-off, conditional on the score, D_1 and D_2 are uncorrelated and β is unbiased.

The advantage of having the left-hand-side variable defined at the level of the applicant i, instead that the level of the teacher j , is that it gives a straight forward criteria to select the relevant running variable. When the left-hand-side is teacher j specific, should we include the certification status and score of a or of b? In this case the election of the running variable would require an additional criteria. Moreira (2016) suggests that only the score of the applicant closest to the discontinuity threshold should be included. If the certification status is quasi-randomly this cut-off, the identification of the parameter of interest, β, would only come from the applicant marginally failing or marginally passing the exam.

Partially Overlapping Reference Groups

Now consider the same situation but, but define peers according to reference group 2. The application decision of teachers a, b, and c at time $t+1$ are given by

$$\begin{aligned} y_a &= \alpha_a + \beta_1 D_1 + \gamma_1 s_1 + \epsilon_a, \\ y_b &= \alpha_b + \beta_1 D_1 + \gamma_1 s_1 + \beta_2 D_2 + \gamma_2 s_2 + \epsilon_b, \quad \text{and} \\ y_c &= \alpha_c + \beta_2 D_2 + \gamma_2 s_2 + \epsilon_c. \end{aligned}$$

Now, the peers' application rate for 1 and 2 are given by

$$\begin{aligned} R_1 = \frac{y_a + y_b}{2} &= \overline{\alpha}_{ab} + \beta_1 D_1 + \gamma_1 s_1 + \frac{\beta_2}{2} D_2 + \frac{\gamma_2}{2} s_2 + \overline{\epsilon}_{ab} \quad \text{and} \\ R_2 = \frac{y_b + y_c}{2} &= \overline{\alpha}_{bc} + \frac{\beta_1}{2} D_1 + \frac{\gamma_1}{2} s_1 + \beta_2 D_2 + \gamma_2 s_2 + \overline{\epsilon}_{bc}, \end{aligned}$$

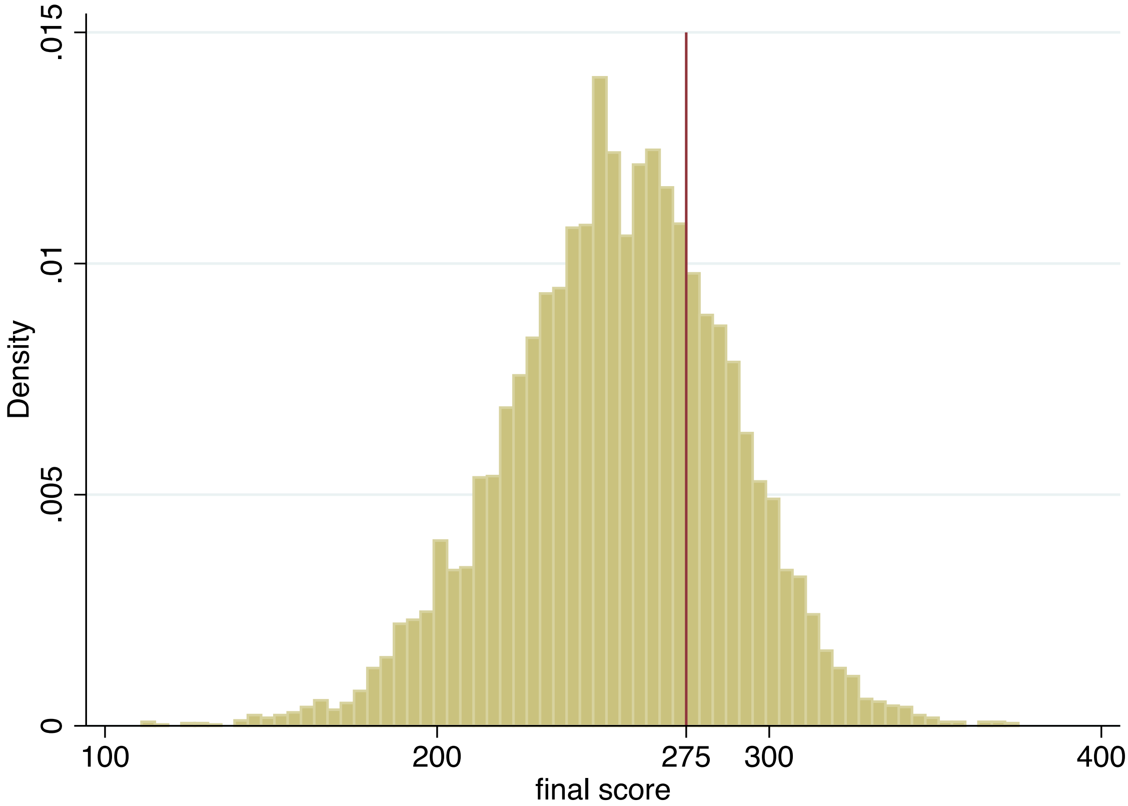
and would be estimated through

$$\begin{aligned} R_1 &= \alpha + \beta D_1 + \gamma s_1 + \underbrace{\frac{\beta_2}{2} D_2 + \frac{\gamma_2}{2} s_2 + \overline{\epsilon}_{ab}}_{\varepsilon_1}, \quad \text{and} \\ R_2 &= \alpha + \beta D_2 + \gamma s_2 + \underbrace{\frac{\beta_1}{2} D_1 + \frac{\gamma_1}{2} s_1 + \overline{\epsilon} + bc}_{\varepsilon_2}. \end{aligned}$$

As before, conditional on the score, the treatment status is quasi-randomly assigned around the discontinuity threshold so that β is unbiased. Moreover, even if the scores of applicant 1 and 2 are correlated, the magnitude of the bias decreases as the size of the reference group increases relative a few common peers (such as b). Following De Giorgi, Pellizzari, and Redaelli (2010), in a future stage of the project, we plan to use the partially overlapping reference group to disentangle the effect of applicant's certification status from the amplifier occurring through peers' simultaneous application decision.

C.2 Figures

Figure C.1: Distribution of the AEP score (Reference Group III)

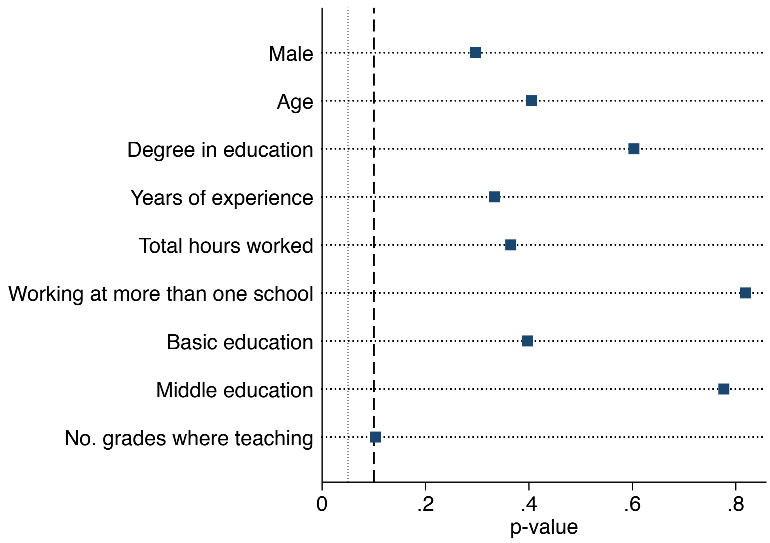


Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Sample of 8,608 applicants with at least two other teachers teaching at, at least, one same grade and school at the time of application.

Figure C.2: Balance of average characteristics of applicants' reference group I, at baseline

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Polynomial of order 1. The dashed line indicates statistical significance at 10%. The dotted line indicates statistical significance at 5%. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application. Average school characteristics excluded as, by construction, the school characteristics of the peers will be the same school characteristics of the applicant.

Figure C.3: Balance of average characteristics of applicants' reference group II, at baseline

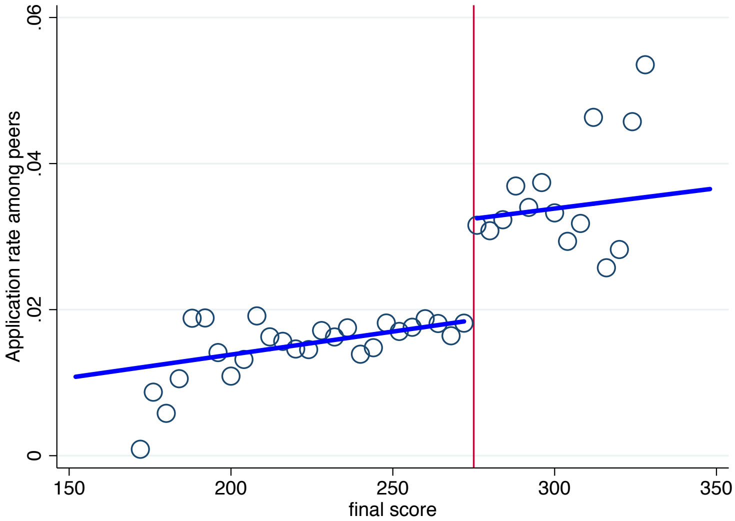


Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Polynomial of order 1. The dashed line indicates statistical significance at 10%. The dotted line indicates statistical significance at 5%. Definition of peers according to reference group II, i.e. teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application. Average school characteristics excluded as, by construction, the school characteristics of the peers will be the same school characteristics of the applicant.

Figure C.4: Balance of average characteristics of applicants' reference group III, at baseline

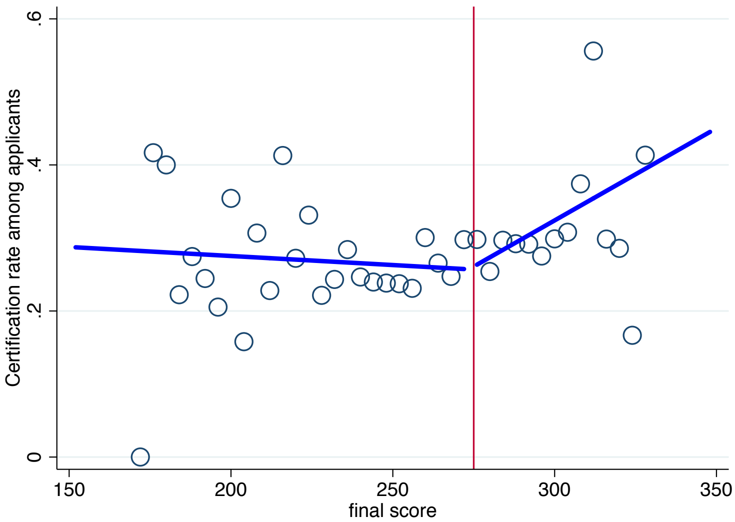
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Polynomial of order 1. The dashed line indicates statistical significance at 10%. The dotted line indicates statistical significance at 5%. Definition of peers according to reference group III, i.e. teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application. Average school characteristics excluded as, by construction, the school characteristics of the peers will be the same school characteristics of the applicant.

Figure C.5: Effect of AEP certification on peers' one period ahead application rate, (Reference group II)



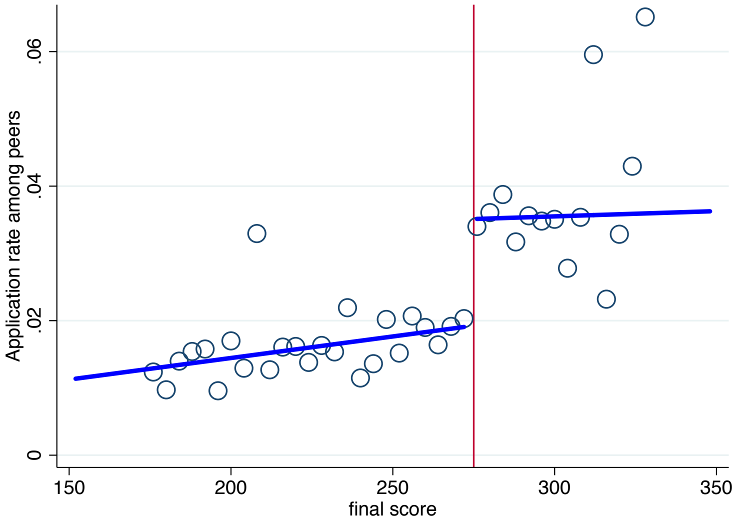
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: The circles represent mean of the outcome variable within bins of size 4 of the score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Errors clustered at the score cell level. Definition of peers according to reference group II, i.e. teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application.

Figure C.6: Effect of AEP certification on peers' one period ahead certification rate (Reference group II)



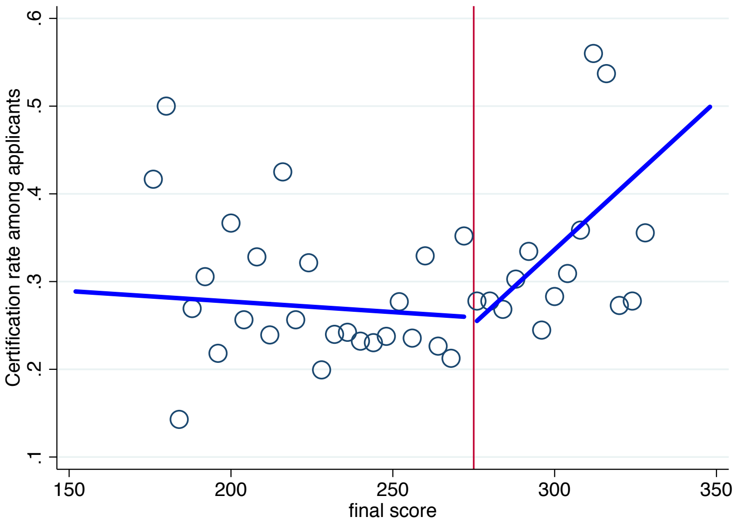
Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. The circles represent mean of the outcome variable within bins of size 4 of the score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Errors clustered at the score cell level. Definition of peers according to reference group II, i.e. teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application.

Figure C.7: Effect of AEP certification on peers' one period ahead application rate (Reference group III)



Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: The circles represent mean of the outcome variable within bins of size 4 of the score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Errors clustered at the score cell level. Definition of peers according to reference group III, i.e. teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application.

Figure C.8: Effect of AEP certification on peers' one period ahead certification rated (Reference group III)



Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate conditional on application. The circles represent mean of the outcome variable within bins of size 4 of the score. Bins with less than 20 observations are excluded. The solid lines show fitted values of a piecewise linear polynomial of the score in the 150-350 window. Errors clustered at the score cell level. Definition of peers according to reference group III, i.e. teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application.

C.3 Tables

Table C.1: Test for Continuity of the Final Score of AEP (Reference Group III)

	All	2003	2004	2005	2006	2007	2008	2009	2010
Frandsen Discrete test p-value	0.707	0.786	0.382	0.508	0.945	0.777	0.701	0.970	0.856
McCrary test p-value	0.250	0.316	0.533	0.930	0.622	0.443	0.425	0.738	0.441

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: McCrary (2008) test at the 275 cut-off, using the optimal bandwidth of 20 and bin size 1. Sample of 8,608 applicants with at least two other teachers teaching at, at least, one same grade and school at the time of application.

Table C.2: Balance of Applicants and Average Reference Group Characteristics at Baseline (Reference Group III)

	Applicants		Reference Group I (school)		Reference Group II (primary-secondary)		Reference Group III (grade)	
Dependent Variable	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Male	0.017	0.031	0.002	0.007	0.008	0.017*	0.013*	0.023**
	(0.017)	(0.022)	(0.006)	(0.008)	(0.007)	(0.010)	(0.008)	(0.012)
Age	-0.259	-0.498	-0.200	-0.394	-0.248	-0.465*	-0.270	-0.480*
	(0.342)	(0.535)	(0.190)	(0.252)	(0.191)	(0.248)	(0.213)	(0.290)
Degree in education	0.005	0.007	-0.000	-0.002	-0.003	-0.007	-0.003	-0.006
	(0.006)	(0.008)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
Years of experience	-0.396	-0.694	-0.228	-0.278	-0.282	-0.367	-0.288	-0.412
	(0.390)	(0.591)	(0.189)	(0.254)	(0.194)	(0.256)	(0.215)	(0.296)
Total hours worked	-0.048	-0.501	-0.145	-0.203	-0.104	-0.179	-0.024	-0.121
	(0.303)	(0.372)	(0.150)	(0.197)	(0.151)	(0.200)	(0.174)	(0.220)
Working at more than one school	0.015	0.013	-0.001	0.005	0.001	0.005	0.004	0.011
	(0.012)	(0.015)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
Basic education	-0.003	-0.017	0.016	0.017	0.007	0.000	0.006	0.000
	(0.015)	(0.018)	(0.012)	(0.015)	(0.012)	(0.015)	(0.012)	(0.014)
Middle education	0.010	0.012	-0.005	-0.007	0.008	0.015	0.007	0.013
	(0.017)	(0.021)	(0.013)	(0.018)	(0.014)	(0.018)	(0.014)	(0.018)
No. grades where teaching	0.016	0.006	0.046	0.025	0.059*	0.042	0.013	0.008
	(0.062)	(0.083)	(0.032)	(0.048)	(0.034)	(0.051)	(0.044)	(0.059)
Rural school	0.015	0.018	0.015	0.018	0.015	0.018	0.015	0.018
	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)
Private-subsidized school	0.024	0.025	0.024	0.025	0.024	0.025	0.024	0.025
	(0.023)	(0.034)	(0.023)	(0.034)	(0.023)	(0.034)	(0.023)	(0.034)
SNED awarded school	0.019	0.027	0.019	0.027	0.019	0.027	0.019	0.027
	(0.015)	(0.019)	(0.015)	(0.019)	(0.015)	(0.019)	(0.015)	(0.019)

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Columns one and two present the characteristics of the 8,608 applicants with at least two other teachers teaching at, at least, one same grade and school at the time of application. Columns three to eight, present the average characteristics of the applicants' reference group at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Robust standard errors, adjusted for clustering at score cells, in parenthesis. Column numbers in parenthesis indicate the order of the polynomial on the score centered around 275.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table C.3: Balance of applicants and average reference group characteristics at baseline

	Applicants		Reference Group I (school)		Reference Group II (primary-secondary)		Reference Group III (grade)	
Dependent Variable	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Male	0.018	0.028	0.023	0.018	0.019	0.008	0.025	0.022
	(0.016)	(0.021)	(0.020)	(0.027)	(0.021)	(0.027)	(0.023)	(0.032)
Age	-0.227	-0.431	-0.398	-0.080	-0.528	-0.069	-0.485	0.171
	(0.316)	(0.492)	(0.404)	(0.497)	(0.400)	(0.490)	(0.419)	(0.626)
Degree in education	0.006	0.008	-0.002	-0.006	-0.002	-0.007	0.002	-0.004
	(0.005)	(0.008)	(0.007)	(0.010)	(0.008)	(0.010)	(0.010)	(0.014)
Years of experience	-0.418	-0.654	-0.318	0.029	-0.422	0.021	-0.287	0.603
	(0.365)	(0.549)	(0.376)	(0.466)	(0.375)	(0.467)	(0.426)	(0.626)
Total hours worked	0.072	-0.294	0.488	0.684	0.415	0.602	0.649*	0.698
	(0.310)	(0.407)	(0.313)	(0.490)	(0.314)	(0.493)	(0.367)	(0.584)
Working at more than one school	0.019	0.018	-0.006	-0.021	-0.011	-0.029	0.004	-0.012
	(0.013)	(0.017)	(0.014)	(0.023)	(0.014)	(0.023)	(0.018)	(0.029)
Basic education	0.001	-0.009	0.047**	0.077***	0.048**	0.082***	0.044*	0.061*
	(0.015)	(0.018)	(0.020)	(0.026)	(0.020)	(0.028)	(0.025)	(0.036)
Middle education	0.005	0.003	-0.030	-0.067**	-0.031	-0.073**	-0.022	-0.042
	(0.016)	(0.020)	(0.022)	(0.030)	(0.023)	(0.031)	(0.024)	(0.033)
No. grades where teaching	0.004	-0.025	0.043	-0.124	0.037	-0.138	0.149**	0.006
	(0.060)	(0.083)	(0.070)	(0.101)	(0.074)	(0.106)	(0.073)	(0.112)
Rural school	0.015	0.024*	0.006	0.005	0.010	0.007	0.018	0.017
	(0.011)	(0.013)	(0.009)	(0.012)	(0.010)	(0.013)	(0.012)	(0.017)
Private-subsidized school	0.023	0.018	-0.001	0.077***	0.012	0.098***	-0.005	0.087**
	(0.024)	(0.034)	(0.020)	(0.029)	(0.020)	(0.029)	(0.025)	(0.035)
SNED awarded school	0.020	0.024	0.020	-0.012	0.009	-0.020	0.032	0.014
	(0.015)	(0.019)	(0.018)	(0.027)	(0.020)	(0.029)	(0.025)	(0.036)

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: OLS regression. Columns one and two present the characteristics of the 2003-2010 AEP applicants at the time of application. Columns three to eight, present the peers' characteristics at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the final score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Robust standard errors, adjusted for clustering at score cells, in parenthesis. Column numbers in parenthesis indicate the order of the polynomial on the score centered around 275.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table C.4: Effect of AEP certification on peers' future certification rate (Reference group I)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Certification rate among all peers	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0045***	0.0044***	0.0053***	0.0050***	0.0047***	0.0043***
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0017)	(0.0016)
N	8,934	8,934	6,384	6,384	4,645	4,645
Clusters	224	224	207	207	205	205
t+2	0.0015*	0.0014*	0.0017*	0.0015	0.0016	0.0014
	(0.0008)	(0.0008)	(0.0010)	(0.0009)	(0.0011)	(0.0011)
N	8,935	8,935	6,385	6,385	4,646	4,646
Clusters	224	224	207	207	205	205
t+3	-0.0006	-0.0006	0.0000	-0.0000	-0.0004	-0.0004
	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0007)	(0.0007)
N	8,933	8,933	6,383	6,383	4,644	4,644
Clusters	224	224	207	207	205	205
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate among all peers. Certification outcome coded as 0 for non-applicant peers. OLS regression. Polynomial of order 1. Definition of peers according to reference group I, i.e. teachers working at the same school as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table C.5: Effect of AEP Certification on peers' future certification rate (Reference Group II)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Certification rate among all peers	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0042***	0.0041***	0.0049***	0.0047***	0.0043***	0.0041**
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0017)	(0.0016)
N	8,934	8,934	6,513	6,513	4,831	4,831
Clusters	224	224	208	208	206	206
t+2	0.0009	0.0008	0.0013	0.0011	0.0016	0.0015
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0012)	(0.0012)
N	8,935	8,935	6,514	6,514	4,832	4,832
Clusters	224	224	208	208	206	206
t+3	-0.0004	-0.0005	0.0001	-0.0000	-0.0004	-0.0004
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0008)	(0.0008)
N	8,933	8,933	6,512	6,512	4,830	4,830
Clusters	224	224	208	208	206	206
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate among all peers. Certification outcome coded as 0 for non-applicant peers. OLS regression. Polynomial of order 1. Definition of peers according to reference group II, i.e. teachers working at the same school and teaching at the same level of instruction as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.

Table C.6: Effect of AEP Certification on peers' future certification rate (Reference Group III)

Dependent variable:	All		Closest to cut-off		Only one applicant	
Certification rate among all peers	(1)	(2)	(3)	(4)	(5)	(6)
t+1	0.0054***	0.0053***	0.0062***	0.0059***	0.0057***	0.0054***
	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0021)	(0.0021)
N	8,582	8,582	6,676	6,676	5,275	5,275
Clusters	222	222	209	209	207	207
t+2	0.0015	0.0014	0.0019	0.0018	0.0029*	0.0028*
	(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0017)	(0.0016)
N	8,574	8,574	6,668	6,668	5,267	5,267
Clusters	222	222	209	209	207	207
t+3	-0.0007	-0.0009	-0.0001	-0.0003	-0.0002	-0.0004
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0010)	(0.0010)
N	8,561	8,561	6,659	6,659	5,260	5,260
Clusters	222	222	209	209	207	207
Controls	0	1	0	1	0	1

Source: Own calculations based on data from the Ministry of Education (Chile).
Notes: Certification rate among all peers. Certification outcome coded as 0 for non-applicant peers. OLS regression. Polynomial of order 1. Definition of peers according to reference group III, i.e. teachers working at the same school and teaching at, at least, one same grade as the applicant at the time of application. Each cell reports the coefficient estimate of a dummy variable indicating if the score was at least 275 points. All specifications include wave fixed effects interacted with the piecewise polynomial of the final score. Odd-columns present the estimates of equation (3.4.1). Even columns present the estimates of equation (3.4.1) and controls by the reference group average characteristics. Robust standard errors, adjusted for clustering at score cells, in parenthesis.
* Indicates statistical significance at 10%.
** Indicates statistical significance at 5%.
*** Indicates statistical significance at 1%.