The Anatomy of Behavioral Responses to Social Assistance when Informal Employment is High*

Marcelo Bergolo
IECON-UDELAR and IZA

Guillermo Cruces
CEDLAS-FCE-UNLP, CONICET and University of Nottingham

September 2020

Abstract

The disincentive effects of social assistance programs on registered (or formal) employment are a first-order policy concern in developing and middle-income countries. We study the impact of a conditional cash transfer (CCT) program in Uruguay on the employment of adult members in beneficiary households in a context of high informality. Our research design relies on the sharp discontinuity introduced by program eligibility rules around a poverty score threshold combined with longitudinal administrative data. We find reductions of about 6 percentage points (a 13% drop) in formal labor force participation among all beneficiaries and of 8.7 percentage points (a 19% drop) for single mothers. The implied elasticity of participation in the formal sector with respect to the net-of-tax rate is about 0.78 for the full sample and about 1.3 for single mothers. The reduction in labor supply is stronger among individuals who have a medium propensity to be formally employed, with a smaller reduction in the case of infra-marginal individuals. We also present suggestive evidence that the reduction in formal employment increases inactivity and informal work in equal proportions. Finally, despite pervasive informality in the context of the Family Allowance assistance program (AFAM), the program's marginal value of public funds of 0.61 implies an efficiency cost within the range of cash transfer programs targeted to families in the United States.

JEL Classification: H31, I38, J22, O17.

Keywords: Cash transfer programs, labor supply, registered employment, efficiency costs

*M. Bergolo: Instituto de Economía (IECON), Departamento de Economía, Universidad de La República, G. Ramirez 1926, Montevideo, Uruguay. email: mbergolo@iecon.ccee.edu.uy. G. Cruces: CEDLAS - Universidad Nacional de La Plata, Facultad de Ciencias Económicas, Calle 6 n. 777, La Plata, Provincia de Buenos Aires, Argentina, CONICET and School of Economics, University of Nottingham, Nottingham NG7 2RD, United Kingdom. email: gcruces@cedlas.org. We would like to thank Leonardo Gasparini, Marco Manacorda and Andrea Vigorito for their valuable input at different stages of this project. We are also grateful to Mariano Bosch, David Card, Emily Conover, Paul Gertler, Corrado Giulietti, Hilary Hoynes, Clément Imbert, Hugo Jales, Carlos Lamarche, Stefan Staubli, Darío Tortarolo, Gonzalo Vázquez Baré, Andrea Weber and Josef Zweimüller for comments on earlier drafts and presentations, as well as seminar participants at the 9th IZA/World Bank Conference on Employment and Development, the European Association of Labour Economists 26th Annual Conference, ORT University seminar (Uruguay), IECON seminar, the IZA/CEPR 17th European Summer Symposium in Labour Economics, the 2015 Northeast Universities Development Consortium (NEUDC) Conference, and the SEU Annual Conference. Two anonymous referees made valuable contributions. G. Cruces would like to thank the generosity of CEGA at U.C. Berkeley, where part of this work was carried out during his visit in January 2015. Matías Giaccobasso provided excellent research assistance. A previous version of this paper circulated under the title "Work Incentives and Welfare Programs. Evidence on Real and Reporting Effects in Uruguay."

1 Introduction

The incentive effects of social assistance programs on labor supply have been at the center of economic policy debate in developed countries. In developing and middle-income countries with high levels of labor informality and poor enforcement of tax and labor regulations, the difficulties of targeting have potential consequences beyond overall labor supply.[1] Social assistance programs may introduce disincentives to registered (formal) employment, which has important implications. First, access to social insurance for workers and their families is typically tied to formal jobs (Levy, 2008, Levy and Schady, 2013). Second, unregistered employment means lower reporting of earnings and thus lower payroll and income-tax revenue. Third, a larger informal sector can lead to a variety of market distortions and efficiency losses, with potential consequences for productivity growth and economic development (La Porta and Shleifer, 2014, Meghir, Narita, and Robin, 2015). In Latin America, the recent expansion of conditional cash transfer (CCT) programs has prompted discussions of their effects on labor formality.[2] Empirical evidence on this specific issue, however, is limited and the impact of CCT programs on overall employment seems to be small or null (Alzua, Cruces, and Ripani, 2012; Banerjee et al., 2015). Moreover, with few exceptions (Gerard and Gonzaga, 2020), these discussions are limited to labor market outcomes in isolation and fail to quantify their welfare/efficiency costs.

Our analysis attempts to fill in some of these gaps in the existing literature by focusing on three main objectives.[3] We first establish the impact on formal employment of a CCT program – the Familly Allowance (henceforth, AFAM) – with a binding-income threshold and frequent income reassessment in Uruguay, a country with a high level of labor informality. Moreover, labor market attachment varies substantially across beneficiaries, some of whom may not be at the margin of (in-)formality. The CCT program in Uruguay, like many others in the developing world, has a "one-size-fits-all" approach that fails to accommodate these features, with potentially negative consequences. Our second objective is, thus, to study the heterogeneity of worker responses in terms of attachment to the formal sector and to decompose the formal employment response into informality and non-employment margins. Our third objective is to evaluate the implications of our results in terms of efficiency costs. We use the responses we estimate to recover the marginal value of public funds (MVPF)

[1] In this paper, "formal employees" refer to those workers who are compliant with social insurance regulations and do not evade payroll taxes. We use the terms "registered (unregistered)" and "formal (informal)" interchangeably, and both terms refer to the status of the worker with respect to the Social Security Administration.

[2] Gasparini and Tornarolli (2009) estimate that approximately 56% of wage earners in Latin America are informal workers—i.e., there are no payroll taxes nor social security contributions associated with their jobs.These workers are effectively excluded from social insurance benefits such as health and old-age pension coverage.

[3] We thank an anonymous referee for suggesting this organization of the discussion and some of the wording of the objectives.

following Hendren's (2016) methodology.

The empirical analysis relies on three matched sets of information linked through unique individual identifiers. The AFAM administrative records contain baseline socioeconomic and demographic information provided by applicants to the program from January 2008 to September 2010. We match adults in the applicant households to their registered employment records, constructed from data provided by Uruguay's Social Security Administration (henceforth, SSA), which is responsible for collecting and recording payroll taxes and social security contributions from registered employment. We thus obtain a rich longitudinal database that covers all spells of registered employment for individuals in the program from January 2005 to December 2012. Because the records only reveal periods of formal (or registered) employment, we complement them with a detailed follow-up survey of eligible and ineligible households that applied to AFAM. The information collected in that survey enables us to observe (self-reported) informal work and non-employment for each household member.

Our identification of AFAM's effects relies on the program's eligibility poverty score, which is based on a household's predicted level of poverty as a function of its characteristics. The authorities strongly enforced this eligibility rule, creating a sharp discontinuity in the likelihood of participation at the cutoff point. The available evidence rules out manipulation of the assignment rule by applicants. Despite that discontinuity, we find some degree of imbalance in registered employment at baseline. We thus identify the impacts of the program through a difference-in-discontinuity research design that compares labor market outcomes for adults in applicant households just above (i.e., the treatment group) and just below (i.e., the comparison group) the program's poverty score eligibility threshold, controlling for potentially confounding pre-application differences around the threshold (Grembi, Nannicini, and Troiano, 2016, Bertrand, Mogstad, and Mountjoy, 2019).

We find that AFAM reduces formal labor force participation by about 6 percentage points (a 13% drop) among all beneficiaries, and by 8.7 percentage points (a 19% drop) for single mothers, who make up about 43% of the total. This translates into an elasticity of formal sector participation with respect to the net-of-tax rate of about 0.78 for the full sample and about 1.3 for single mothers. This effect persists for the two-year period following application to the program. We also break down the behavioral responses to the program by subgroup and by type of employment. We study heterogeneous impacts for individuals with different predicted levels of attachment to the formal sector in the absence of the CCT program and find substantial variation in this dimension. The reduction is stronger (8.9 percentage points, or a 29% drop) among individuals who have a medium propensity to be formally employed, with an elasticity of about 1.64, and is even stronger for single mothers in this group—a drop of 12.4 percentage points (about 50%), yielding an elasticity of about 3.46. The impact on those who have a low propensity to be formally employed is -2.5 percentage points and is not statistically significant, probably because these individuals would not be employed formally irrespective of their participation in the assistance program. For those who have a high propensity to engage in formal work, on the other hand, the effect of about -4.7 percentage points is substantially lower than the average effect in proportional terms (about 6%), with a low elasticity of about 0.31. The latter group's higher propensity to be formally employed implies that its formal labor force participation probably also would not be substantially affected by participation in the assistance program. We then present suggestive evidence concerning AFAM's impacts on informal employment and non-employment using data from a follow-up survey that sampled households on both sides of the discontinuity. We find that the decline in formal employment is divided more or less equally between individuals moving into inactivity and those moving into informal employment. Finally, based on our estimated behavioral responses, we provide efficiency cost calculations by computing the marginal value of public funds (MVPF) for AFAM single mothers. Despite pervasive informality in the context of the program, its MVPF of 0.61 implies an efficiency cost similar to that of cash transfer programs targeted to families in the United States, for which Hendren and Sprung-Keyser (2020) estimate an average of 0.74 and a confidence interval of [0.36,1.47], though lower than the MVPF of similar programs that induce labor supply disincentives, such as AFDC (MVPF of 0.87).

The main contribution of our paper is to provide clear and credible evidence of the negative effect of a CCT program on formal employment in a context of high informality. The existing literature reports small or null effects on the overall labor supply (Alzua, Cruces, and Ripani, 2012; Banerjee et al., 2015), but the evidence on formal employment remains limited.[4] Garganta and Gasparini (2015) find that a large CCT program in Argentina reduced transitions to formal employment by 8.4 percentage points – a substantial drop of 40%. Their study reports intention-to-treat estimates, comparing poor households with and without children, and their empirical work relies on household survey data. Amarante et al. (2011), in turn, perform an exercise similar to ours for the PANES program in Uruguay (PANES was a more targeted, temporary, and smaller-scale predecessor to AFAM with eligibility also determined by a poverty score). The authors analyze the effect of the program on formal labor participation and earnings and find a negative impact on formal employment for men (but not for women) in eligible households of about 2.5 percentage points and of

[4]Besides the three papers cited in this paragraph, there are other studies and policy reports that provide suggestive evidence on CCTs and informality. In most cases, though, there are limitations in the underlying data or in the research design (see, for instance, Bosch and Manacorda, 2012, Araujo et al., 2017). In contrast, the empirical research on developed countries includes a sizable body of rigorous studies of the impact of social assistance on labor supply, as well as on the underlying mechanisms and magnitude of those impacts, the population groups most affected, and welfare implications (Chan and Moffitt, 2018). Moreover, these studies have generated a series of detailed analyses of social assistance programs' incentive effects on labor supply, which are often key to the design of welfare reforms aimed at improving the efficiency-equity trade-off (Blank, 2000, Scholz and Levin, 2001).

about 8 USD in monthly earnings. Bergolo and Galvan (2018), based on a sub-sample of the follow-up survey described below, study intra-household responses to AFAM and find a negative impact on the registered employment of married women. Bosch and Schady (2019) do not find effects of a CCT program on labor supply in Ecuador, although they report a small reduction in registered employment for women. These papers report some heterogeneous effects based on household and individual socioeconomic characteristics, but not based on an individual's propensity to be formally employed, which, as we discuss below, is a key aspect for the analysis and design of cash transfer programs. Moreover, the three papers report absolute and relative impacts of the programs on beneficiaries' labor supply, but fall short of providing efficiency cost estimates of the policies.

A second contribution is our analysis of the heterogeneity in responses. As stressed by Eissa, Kleven, and Kreiner (2008), efficiency implications vary substantially when the analysis of impacts is broken down by subgroup. The literature on labor supply and on informality fails to properly acknowledge the heterogeneity in labor market attachment across workers in developing countries – in particular, the fact that some workers may not be at the margin of (in)formality – and the attendant implications for the impact of social programs. Gerard and Gonzaga (2020) address this issue by exploiting heterogeneity across labor markets in Brazil and highlight the lower efficiency costs of unemployment insurance in labor markets with higher informality. Our analysis complements their findings with results from individuals within the same labor market.

A third contribution is the provision of efficiency cost impacts for the program. Existing papers on CCTs typically estimate impacts but fail to derive the implications of their results for welfare/efficiency. This point has also been recently emphasized by Gerard and Gonzaga (2020), who develop a full welfare analysis of unemployment insurance in the context of informality in Brazil. Unemployment insurance programs are still small in Latin America, especially in comparison to CCTs such as AFAM. We compute the marginal value of public funds in what is, to our knowledge, the first application of Hendren's (2016) methodology to a developing country. This allows us to compare the efficiency implications of the AFAM program to those of cash transfer programs in the United States as reported by Hendren and Sprung-Keyser (2020).

In a broader context, this paper also contributes to a growing body of literature on public finance and development, particularly to studies that use administrative microdata and quasi-experimental methods to evaluate the efficiency cost implications of social policies in developing countries.[5] Our results add to this literature by extending its analysis to CCT programs, and by presenting efficiency cost implications.

The paper proceeds as follows. Section 2 describes the context of Uruguay's social pro-

[5]See for instance Kleven and Waseem (2013), Best et al. (2015), Pomeranz (2015), Carrillo et al. (2017), Gerard and Gonzaga (2020), Naritomi (2019).

tection system, the AFAM program, its rules and characteristics, and the expected effects of AFAM on participants' labor market outcomes. Section 3 describes the data sources and the construction of the datasets and samples we use in our empirical analysis. Section 4 discusses the empirical approaches on which we base our identification strategy. Section 5 presents the main results on registered employment responses to the program, their heterogeneity, the implied elasticities of participation in registered employment and the decomposition of the effects between informality and non-employment. Finally, Section 6 discusses the efficiency implications of the program's financial incentives. Conclusions follow.

2 The AFAM Program: Financial Incentives and Expected Behavioral Responses

2.1 Context of the AFAM Program

Uruguay is a middle-income country in South America with a total population of approximately 3.3 million. The country has one of the oldest and most developed social protection systems in Latin America.[6] Access to most welfare and social insurance programs is linked to registered employment and financed through payroll taxes and contributions from both employers and employees, totaling about 32 percent of taxable wages. Registered (or formal) employees are those working in firms that reported them to the Social Security Administration (henceforth, SSA) and for which they paid the relevant taxes and contributions, which grants eligibility for social insurance benefits, mainly health insurance, family allowances, and old age pensions. A substantial fraction of employees are not registered with the SSA and thus are not covered by social insurance benefits. This phenomenon is referred to as labor informality (Gasparini and Tornarolli, 2009).

Uruguay suffered a severe economic crisis in 2002-2003. Unregistered workers, who lacked access to the counter-cyclical social assistance mechanisms provided by the SSA, were especially hard hit by this crisis. The government responded by launching a temporary conditional cash transfer program called *Plan de Atención Nacional a la Emergencia Social (PANES)* in 2005 targeted at the poorest 10 percent of households (see Manacorda, Miguel, and Vigorito, 2011 for details).

This emergency program was replaced in January 2008 by a new system of family allowances (Law 18.227), the AFAM program, as part of a broader progressive tax and transfer system reform. AFAM targeted poor households with children or pregnant women in the bottom 20 percent of the income distribution. It was implemented as a means-tested conditional cash transfer (CCT) program targeted to households in precarious socioeconomic

[6]See Appendix A.2.1 for more details on Uruguay's social insurance system and the context of the AFAM program.

conditions. The program's monetary transfers were conditional on health checks (both for pregnant women and children) and school attendance for children in beneficiary households. AFAM became the most important social assistance program in Uruguay in terms of both the extent of coverage (about 42 percent of all children under the age of 18 in Uruguay) and the magnitude of the cash benefits provided (about 0.35 percent of GDP).

2.2 Eligibility and Disqualification from AFAM

Eligibility for AFAM depended on two criteria: the household had to pass both an income test and a proxy means test. To participate in the program, households were first required to complete an application, in which they provided an array of socioeconomic information, including household characteristics (address, housing conditions, dwelling type, characteristics and quality, ownership, access to water and sanitation, etc.), and detailed information about household members, such as their national identification number, education levels, labor force participation, and income levels.

After completing the application, households were first subject to an income test: the per capita income of the household had to be below a predetermined threshold in order to qualify for the program. The household's income level for this test was computed by combining the information reported on the application form and SSA administrative records. The SSA computed total household income by matching its administrative earnings records with each household member's national identification number. These earnings included those from registered employment (as reported to the tax and social insurance authorities by employers), from pensions (paid by the SSA), and from other government transfers recorded in the SSA administrative records. According to the information provided by AFAM authorities, about 70% of the total verifiable household income corresponded to earnings from registered employment. We refer to this as "verifiable income" throughout the rest of the paper. The income test compared per capita household income with a given threshold. Crucially, unlike programs in other countries in the region (such as the Asignación Universal por Hijo in Argentina – Garganta and Gasparini, 2015), AFAM was compatible with formal employment (up to a certain level of earnings determined by the income test threshold). That is, formally employed individuals could still be eligible to receive AFAM cash transfers provided their income from formal employment was sufficiently low.

The second step for determining eligibility after the income test was a proxy means test. This test relied on a poverty score calculated by program officials and was based on a large set of socioeconomic characteristics provided by the household on the application form. The score's algorithm was devised in consultation with academics and social policy experts and details of its components and construction were never disclosed to the public.[7]

[7]The eligibility poverty score was devised by researchers at the Universidad de la República (UDELAR)

Households with income below the income test's threshold and with a poverty score above a predetermined level were deemed eligible to participate in the AFAM program. Households could still be rejected or disqualified for administrative reasons, such as, for example, if they failed to submit proof of children's school enrollment or of completion of health checks. Appendix Table A.4.1 summarizes reasons for rejection of AFAM applicants. About 50 percent of the applicant households were deemed ineligible because they failed the income test, the proxy means test, or both.

Most eligible households ended up participating in the program. The program rules specified three main situations that could lead to a household's disqualification after enrollment. First, a household's earnings computed from administrative records (its verifiable income) should not rise above the income eligibility threshold. The income test was applied to participant households every two months, and it was strictly enforced. Second, households had to adhere to the mandatory health check and school attendance requirements for children. Finally, households were disqualified from the AFAM program when all of the household's children reached 18 years of age. Our analysis of the 2009-2012 administrative records of AFAM's disqualifications, presented in Appendix Table A.4.2, indicates that, on average, about 15 percent of AFAM beneficiary households were disqualified each year during that period, and about 57 percent of disqualifications were attributable to households' verifiable incomes exceeding the income test threshold.

2.3 AFAM Monetary Benefits and Implicit Financial Incentives

Eligible households that participated in the program were entitled to a monthly cash transfer. A key feature of the AFAM program's incentives was that 100 percent of verifiable income up to the threshold amount for program eligibility was exempt from taxation, which yields an implicit marginal tax rate on the benefit of zero percent. Above this threshold, benefits abruptly drop to zero, creating a "notch" in households' budget constraints.[8]

The amount of the transfer varied according to the household's number of children under

in Montevideo, Uruguay (see Amarante and Vigorito, 2011). The algorithm is based on the coefficients of a probit model in which the dependent variable is equal to one if a household is within the first—i.e., the lowest —quintile of per capita income, and zero for those above the first quintile but below the median per capita income. The original model was estimated by means of a fully saturated function of household variables drawn from the ECH national household survey. The resulting coefficient estimates were used to predict the score for each applicant household based on data from the application form, and the eligibility thresholds for residents of Montevideo, the country's capital, were different from those for the rest of the country, to reflect differences in living costs. Uruguay's Social Security Administration, the *Banco de Previsión Social* (BPS), and the Ministry of Social Development (*Ministerio de Desarrollo Social*, henceforth MIDES), were responsible for computing the score and monitoring household income and conditionalities.

[8] This feature of the AFAM program is similar to most of the income support programs in Latin American countries, though not to traditional welfare programs in developed countries where income benefits are reduced at a specific rate. The benefit structure of the Job First program in Connecticut was similar to AFAM's structure in that program rules exempted from taxation 100 percent of earnings up to the monthly federal poverty line (Bitler et al., 2006, Kline and Tartari, 2016, Hartley and Lamarche, 2018).

the age of 18 and the number of children in the household attending secondary school. The transfer was larger for those in secondary education to encourage older children to attend and finish school.

The total benefit granted to a household was computed as follows:

$$G_{AFAM} = \begin{cases} 0 & \text{if } (1-\theta) \times Y^F > T \\ \beta \times (Kids0to17)^{0.6} + \delta \times (HighSchoolKids)^{0.6} & \text{if } \ (1-\theta) \times Y^F \leq T \quad where\, \theta = 0.15 \end{cases} \tag{1}$$

where Y^F is per capita household gross formal (or verifiable) income, T is the per capita income test threshold, θ is a fixed deduction equal to 15 percent of the household's gross verifiable income, $Kids0to17$ represents the number of children below 18 years old, $HighSchoolKids$ the number of children that attend secondary school, and β and δ are transfer levels.

The cash transfer and the income test's threshold levels were adjusted periodically according to the evolution of the official Consumer Price Index. The resulting cash transfer levels were relatively generous sufficiently so to affect beneficiary households' labor supply decisions: for 2008-2012, the transfer amounted on average to about 68 percent of the minimum wage for an average AFAM household and about 48 percent of the poverty line (Appendix A.2.3 details the levels of the parameters).

All households in our study sample passed the income test at the time of application and were deemed eligible or ineligible according to the proxy means test. The latter was only conducted at the time of the original application and was not subject to manipulation. On the other hand, beneficiary households could adjust their labor supply and were subject to the income test every two months. Our discussion focuses on this frequently enforced eligibility condition. The level of the benefit and the income test constitute the two key factors for understanding labor market responses to AFAM in the context of our paper.

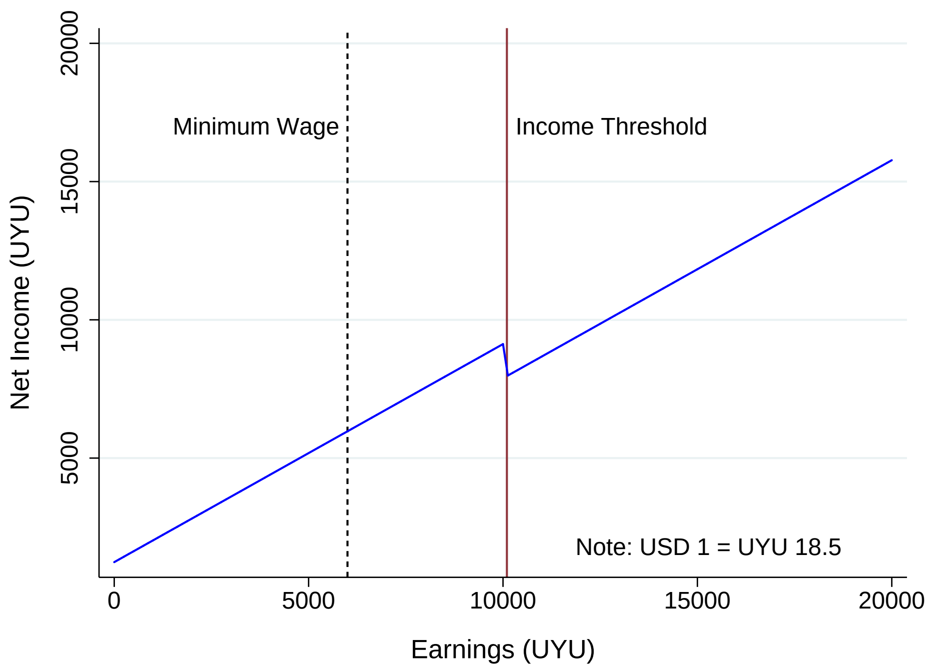
Figure 1 illustrates the program's implicit financial incentives to engage in formal employment. It depicts the simulated budget set for a representative individual in our setting – a single mother of one, beneficiary of AFAM, whose child attends secondary school (single mothers represent about 43% of AFAM adult beneficiaries). The tax schedule for most low income registered workers in Uruguay is essentially flat and corresponds to a payroll tax amounting to 21 percent of earnings – 21 percentage points out of the 32 percent total payroll tax is payed by the employee. The AFAM cash benefit is computed according to equation 1 with parameters corresponding to 2011. As a benchmark, the graph shows the level of the national minimum wage.

The AFAM income threshold for eligibility generates a large disincentive to work formally. There is a drop in net income as gross verifiable earnings exceed UYU 10,088 because at this earnings level the AFAM benefit is lost entirely.[9] If she engages in formal work, at

[9] While the income-eligibility threshold is computed in terms of household per-capita (formal) income, for

an earnings level of UYU 10,088, the single mother in the example depicted in Figure 1 loses UYU 2,118 (10,088×0.21) of her income in net tax payment and UYU 1,236 from the withdrawal of AFAM benefits. That is, the net income gain from working at a registered job declines from UYU 9,206 to UYU 7,970. The effective average tax rate relevant to the decision to engage in formal employment at this level of registered earnings is 33.2 percent – i.e., 21 percent from the tax and an additional implicit tax rate on registered employment of 12.2 percent (UYU 1,236/10,088) as a consequence of losing the cash transfer.

Figure 1: Budget Set for Single Mother with one Child



Notes: This figure presents a simulation of the stylized budget set for a single mother with one child. The dataset corresponds to the *AFAM administrative records* and Uruguay's ECH household survey for the year 2011. The figure is based on the AFAM cash benefit, AFAM income threshold, national minimum wage, individual earnings, and the nominal exchange rate for the year 2011. The AFAM cash benefit is computed using the following equation (1) : $UYU1,236 = 865 \times (1)^{0.6} + 371 \times (1)^{0.6}$. A household's income-eligibility threshold is transformed to an equivalent "per-person" amount to make it comparable with individual earnings – i.e., $T^* = T \times HHmembers/(1-\theta) = (4,287 \times 2/0.85) = UYU\,10,088$.

2.4 Expected Behavioral Responses to AFAM and Efficiency Implications of the Program

In a labor market with high informality, as is the case in most developing countries, the standard labor supply model predicts that an income-tested transfer program such as AFAM

the sake of simplicity we render it equivalent to a "per-person" rate to allow for comparison with individual earnings. Specifically, $T^* = T \times HHmembers/(1-\theta) = (4,287 \times 2/0.85) = UYU\,10,088$, where *HHmembers* corresponds to the ratio of a household's total number of members to the number of formally employed individuals. In this example, *HHmembers* is equal to 2 because we assume the mother is the only formally employed member in the household.

could induce a reduction in potential beneficiaries' registered employment at the extensive margin, resulting in either higher non-employment or higher informal employment.[10]

Our analysis of AFAM's effects on labor supply is based on a comparison of households that passed the income test at the time of application but were deemed either eligible or ineligible by the proxy means test (i.e., their eligibility poverty score). Enrolled households' post-application earnings from registered employment were monitored continuously and failure of the income test resulted in cessation of the benefit.[11]

We can illustrate the effects of the program with the cases of two comparable adults from similar households that applied to the program at the same time, passed the income test, but fell on different sides of the eligibility poverty score threshold. The adult in the household that passed the proxy means test and became a beneficiary of the program was more likely to limit her participation in registered employment because of an income effect from relatively higher unearned income. The adult in the beneficiary household also experienced a substitution effect that impelled her to limit her participation in registered employment because higher formal earnings could result in loss of the benefit due to the continuous formal earnings-based income test. The reduced formal labor participation in both cases could result in informal work or non-employment. The adult in the ineligible household, on the other hand, was more likely to enter or increase participation in formal employment, especially in the context of a growing economy as in the period under study.[12]

Behavioral changes induced by AFAM might be due either to substitution or income effects and might result in increased non-employment or, in a setting with a high level of informal employment, in more informal work. The total share of eligible adults in registered employment should decrease. Moreover, as long as there are two tiers of informality in the labor market (Fields, 2009), the negative effect on registered employment should be higher among individuals with a higher propensity to be formally employed than on potential beneficiaries with a lower propensity to be formally employed (see the discussion of these potential heterogeneous effects in section 5.2).

These effects can be used to evaluate the efficiency impact of the program. Standard public finance analysis indicates that substitution effects lead to distortions and thus have efficiency costs, whereas income effects do not generate distortions at the margin. Only

[10] In fact, AFAM's design implies that it might also affect the intensive margin of labor supply, because it might generate a notch at the income test threshold in households' budget constraints (Kleven, 2016). Unfortunately, we are unable to measure labor supply or misreporting at the intensive margin because our data do not cover working hours or earnings (see Section 3).

[11] Beyond the standard labor supply arguments, the program's conditionalities (requirements in terms of children's health checks and school attendance) also might have effects on the labor supply of adults in potential beneficiary households. We discuss these additional considerations in Appendix Section A.4.1.

[12] Uruguay's economy grew after the introduction of AFAM, with an increase in the overall rate of formal employment in the period under study (see Figure A.2.1 in Appendix A.2.2), which implied better labor market prospects for potential beneficiaries. The impact we estimate in our empirical work below thus corresponds to a relatively smaller increase (rather than an absolute decrease) in formal employment.

the former matters for understanding policies' efficiency implications. Our reduced-form estimates of the program's effects on employment do not allow us to separate between the income and substitution effects at play in the behavioral responses to AFAM. This imposes a clear limit on the interpretation of our results in terms of traditional efficiency cost analysis. In Section 6, we present an alternative to assess the efficiency implications of the program following Hendren's (2016) "policy elasticity" approach. This framework relies on a policy's effect on government revenue as a sufficient statistic for all behavioral responses, and posits that the welfare analysis of public policy changes can be based on individuals' willingness to pay out of their own income. The only behavioral response required to compute the efficiency cost impact of the program is the causal effect of the AFAM on the government's budget, which does not require computing the compensated behavioral responses of traditional welfare analysis.[13] We can thus bypass the impossibility of distinguishing between income and substitution effects in our setting.

3 Data Sources and Sample Construction

Our main analysis is based on a series of administrative and household survey datasets matched by means of a unique national individual identity number, the "Cédula de Identidad." Every citizen is assigned one of those numbers at birth, and it is required (and recorded) in all transactions related to registered employment, social insurance and welfare programs. This section provides a brief description of the original data sources, their characteristics and time frames. It also documents the matching process, the resulting datasets used in our empirical work, and the outcomes of interest. For more details see Appendix Section A.1.

3.1 Data Sources

AFAM Administrative Records. Our primary dataset is the *AFAM administrative records*, which includes individual- and household-level data on application to the program and during participation in it. This dataset was constructed from two sources. The first, the *AFAM baseline application records*, consists of responses to a detailed questionnaire on the socioeconomic and demographic characteristics of all individuals in households that applied to the program. This dataset also includes the date of application and, most importantly, the exact value of the household's eligibility poverty score computed by the authorities – we

[13] The envelope theorem implies that behavioral responses to marginal policy changes do not have a direct impact on individual utility, and thus the welfare analysis can rely solely on a policy's causal effect on government revenue (Hendren, 2016).

standardize that score for our analysis.[14] This rich baseline data contains information for both successful and unsuccessful applicant households, and it covers the period that runs from January 2008 to September 2010. The second dataset, the *AFAM participation records*, corresponds to the participation history of the individual/household once they were enrolled in AFAM – i.e., it covers a sub-sample of households in the *AFAM baseline application records*. This dataset covers the period January 2008 to March 2012. This dataset was compiled by the program's implementation institutions, MIDES and the SSA.

SSA Registered Employment Records. Our main outcome of interest is registered (or formal) employment. We obtain information on this outcome from a further source, the *SSA registered employment records*, which keeps monthly records of monetary contributions made by employers and employees to compute eligibility for and levels of social insurance services and transfers.[15] A formal employee is one who is "registered" with the SSA, and thus is covered by the social insurance benefits provided by the SSA. We have access to these records for program applicants for the period that spans from January 2005 to December 2012. The main advantage of this type of data is that it records all episodes of registered employment for both employees and self-employed workers, in both the private and public sectors. We construct a longitudinal database of monthly registered employment histories that covers the entire period under study and, most importantly, substantial periods of time before and after the application records. In terms of limitations, the *SSA registered employment records* do not include information on hours worked per day (or days per month) nor on earnings from registered work, which restricts the scope of our analysis. We cannot, for instance, determine the impact of AFAM on the intensive margin of labor supply or fully estimate all elasticities (see the discussion in Section 5.3).

AFAM Follow-Up Survey. An additional limitation of administrative databases for the study of labor market outcomes in developing countries is that, by definition, these sources do not have any information on unregistered (or informal) employment. Therefore, to complement the administrative data source, we use a follow-up household survey (the *AFAM follow-up survey*) specifically designed to study the effects of AFAM on household welfare and on individual labor market responses.[16] The survey's sampling frame was designed taking into account the discontinuity in the program's eligibility score threshold with the

[14] We standardize this score so that the eligibility threshold is zero. The standardized eligibility score for the population of all applicant households is in the range $(-0.257; +0.712)$. Eligible households have positive values and ineligible households have negative values.

[15] As detailed in Section 2, the SSA provides old age pensions, unemployment benefits, and maternity and child allowances, among other transfers, and determines health insurance eligibility.

[16] The survey was designed and implemented by researchers at the Instituto de Economía (IECON) of the Universidad de la República (UDELAR), in collaboration with MIDES and researchers at the Institute of Statistics and at the Department of Sociology at UDELAR (Amarante and Vigorito, 2011).

purpose of exploiting the quasi-random variation generated by the eligibility rule. It was based on a stratified random sample of eligible and ineligible households around the cutoff point. The interval of the (standardized) eligibility poverty score for the survey was set in the range $[-0.0426; 0.0727]$. The *AFAM follow-up survey* was conducted from September 2011 to February 2013. Its questionnaire consisted of a shortened version of Uruguay's National Household Survey (the *Encuesta Continua de Hogares,* henceforth ECH). It asked extensive questions about household living conditions, socioeconomic characteristics and, crucially, mutually exclusive labor market outcomes on the date of the interview – i.e., registered and unregistered employment, unemployment and non-participation – for each adult individual in the sample.

Encuesta Continua de Hogares-ECH. We also use the microdata from the ECH for the period 2008-2012. The ECH is a nationally representative household survey conducted according to international standards. It is carried out periodically by Uruguay's national statistical agency (*Instituto Nacional de Estadísticas*, INE). We use ECH data to describe labor market patterns in Uruguay (see Appendix Section A.2.2) and to impute income for estimating the participation tax rates (see Section 5.3).

3.2 Samples for the Empirical Analysis

From these multiple data sources, we construct two related but distinct samples and datasets for our empirical analysis.

Main Sample. We first matched the *AFAM administrative records* with *SSA registered employment records* to yield complete registered work trajectories for the period that runs from January 2005 to December 2012 for all adults in households that applied to AFAM between January 2008 and September 2010. The matched dataset also contains AFAM participation history for the same period and sample. We adjust this data to reflect the fact that application to AFAM occurred at different moments during the period under study (Appendix Section A.2.4 discusses patterns of AFAM application and participation over time). As in an event study, we re-center the work histories at the time of application to the program for individuals in both eligible and ineligible households.

We impose four restrictions to construct our study sample. First, we include only household heads and their spouses (when present) who were aged 18 to 57 at the time of application to AFAM during the period January 2008 to September 2010. These age limits restrict our sample to the economically active population over the whole period under study.[17] Second,

[17]The lower age limit means that individuals under age 18 would not have benefited from the program as children (e.g., they were not required to be enrolled in school). The upper age limit of 57 years reflects the fact that workers are eligible for retirement benefits in Uruguay once they reach 60.

we restrict the sample to the population of applicants in households that fall in the eligibility poverty score range $(-0.257; +0.257)$.[18] Third, we restrict our observation window for the data on registered employment to 36 months before and 24 months after the household's application date in order to generate and analyze a balanced panel. While we could study longer post-application periods for some individuals, our criterion balances the length of the pre- and post-application periods for all individuals in our sample, and thus avoids the selection issues that would arise if we mixed in our sample individuals from households with different lengths of enrollment. Fourth, to use a homogeneous sample across all of our empirical work, we discard a small number of observations (fewer than 5 percent) for which we cannot impute labor income due to missing covariates since this imputation is required to compute the elasticity of registered employment (see Section 5.3).[19]

From the population of 240,146 adults in all applicant households in the range $(-0.257; +0.712)$ of the eligibility poverty score, we use a sample (henceforth, the *Main Sample*) of 83,591 individuals (74,652 eligible and 8,939 ineligible) in households in the eligibility poverty score range $(-0.257; +0.257)$. For some specific analyses, we report results based on the $(-0.257; +0.712)$ bandwidth of the eligibility score (we refer to this as the *Population*). With these restrictions, the *Main Sample* consists of a five-year balanced panel of 5,099,051 monthly observations from 83,591 individuals observed for 61 months each (36 months pre-application, the application month, and 24 months post-application).

Follow-Up Sample. The second sample (henceforth, the *Follow-Up Sample*) consists of adults in households interviewed for the *AFAM follow-up survey* during the period September 2011 to February 2013. As described above, these households were also drawn from a range around the program's poverty eligibility score. There were 2,403 adults (18-57) interviewed in the follow-up survey (1,656 eligible individuals and 747 ineligibles). We matched the survey responses from these individuals with their and their households' information (supplied by the *AFAM administrative records* and the *SSA registered employment records*. Table A.1.1 in the Appendix presents summary statistics of selected socioeconomic characteristics for the samples we use in the empirical analysis.

3.3 Main Outcomes of Interest

The main outcome of interest in our analysis is registered employment as recorded in *the SSA registered employment records*. Specifically, registered employment in our data is an

[18] Restricting the sample to households symmetrically located around the zero value does not alter the choice of the optimal bandwidth (Calonico et al., 2014), but it does substantially reduce the computational time for obtaining our estimates.

[19] Our main results concerning the program's impact on registered employment remain qualitatively the same and quantitatively very similar if we include these discarded observations (comparison of results in Table 1 and Appendix Table A.6.8).

indicator variable equal to '1' if the individual is registered with the SSA in a given calendar month, and zero otherwise.

The analysis in Section 5.4 is based on the *AFAM follow-up survey*, and we exploit the information from the survey to explore responses to AFAM on additional margins of labor force participation. The first of those margins is registered employment, based on responses to a specific question in the follow-up survey and defined as an indicator coded as '1' for individuals who declared themselves to be a registered employee at the time of the interview and zero otherwise.[20] The second outcome of interest in our study is unregistered employment, again as reported by the respondent in the follow-up survey. This indicator is coded as '1' for individuals who work but state that they are not registered with the SSA, and zero otherwise. The third outcome is non-employment, which is an indicator variable equal to '1' when the individual declares that she is not working, and zero when she states that she is employed (either as a registered or as an unregistered worker).

The variable that defines assignment to the treatment in our empirical analysis is the eligibility condition for enrollment in AFAM. As discussed below, for identification reasons eligibility for AFAM is defined according to the household's eligibility poverty score (henceforth, "eligibility score") during the process of targeting household applicants. We standardize this score so that the eligibility threshold is zero, with positive values for eligible households and negative values for ineligible households.

4 Empirical Approach

Our empirical strategy exploits a specific feature of the AFAM program's eligibility rules. Households that apply to the program must first pass an income test based on household members' earnings from registered employment, as reported to tax and social security authorities. As described above, when household income is below a certain threshold (i.e., when the household passes the income-eligibility test), the household is subject to a proxy means test and assigned a eligibility score based on a detailed set of socioeconomic and demographic characteristics. The household is deemed eligible for the program only if its score is above a predetermined threshold. The authorities strongly enforced this eligibility rule, which created a sharp discontinuity in the likelihood of participation at the cutoff point as depicted in Figure 2. The figure plots the proportion of applicant households (vertical axis), both those deemed eligible and those deemed ineligible through the application process, that were enrolled in the program at any given point in time since its implementation in 2008 until the

[20]The specific question in the ECH and in the *AFAM follow-up survey* is: "Are you contributing to a retirement benefit through your current job?" (in Spanish, "¿Aporta a alguna caja de jubilaciones por su trabajo actual?"). This is a standard question to gauge registered or formal work in household surveys in Latin America (see, for instance, Gasparini and Tornarolli, 2009, and Galiani and Weinschelbaum, 2012).

end of our observation window in 2012, as a function of the eligibility score (horizontal axis). The figure clearly shows a sharp discontinuity in the probability of participation in AFAM at the threshold. Participation is about 96 percentage points higher for households with a score just above the eligibility threshold compared to those with a score just below it.[21] This finding confirms that the AFAM eligibility score assignment rule was strictly enforced.

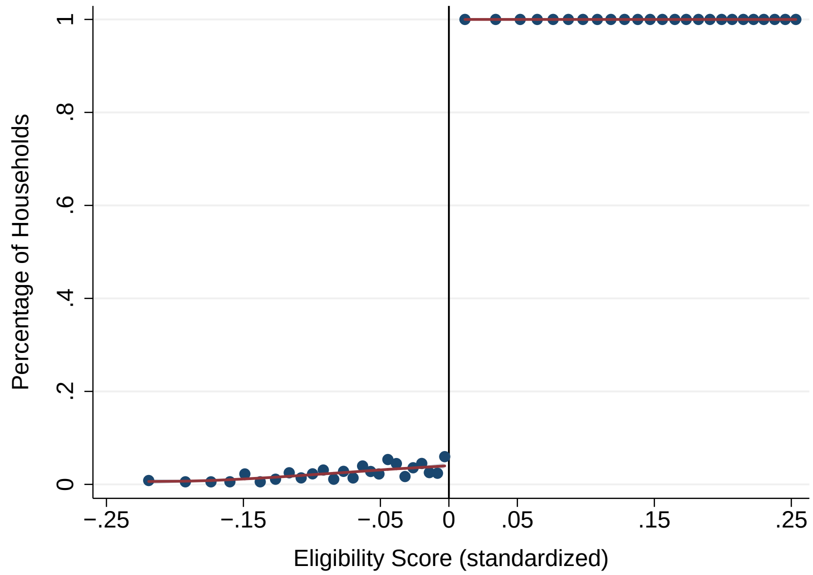
Figure 3 depicts the level of our main outcome of interest (registered employment) around the threshold before and after program application. In each panel in the figure, we plot for the *Main Sample* the mean unrestricted monthly registered employment (blue circle) in bins of width equal to 1 percentage point of the score and the estimated monthly means (solid red line) from a local linear regression applied to each side of the cutoff point along with the 95-percent confidence interval. Panel A in Figure 3 depicts the pre-application difference in registered employment around the AFAM eligibility threshold. There is a small but statistically significant jump at the threshold of about -3.8 percentage points (the RD estimate, discussed below, is reported in Table 1). Registered employment levels for the two groups start to diverge substantially after application to the AFAM. Panel B in Figure 3 shows a much larger jump at the eligibility threshold of about -9.4 percentage points (also reported in Table 1).[22]

Despite the sharp discontinuity in eligibility depicted in Figure 2, the imbalance in the pre-application period main outcome, and in some other baseline characteristics (as documented and discussed in Appendix Section A.4.3) may signal manipulation of the running variable, which would compromise identification of causal effects in the context of an RD research design based on Figure 3. If individuals were able to manipulate the program's eligibility process, the difference between the eligible and the ineligible at the threshold would reflect some systematic advantage favoring the eligible, instead of being determined by plausibly exogenous idiosyncratic factors. This would happen, for instance, if program officials favored households with adults engaged in informal jobs, or if applicants lied about their socioeconomic characteristics when filling in the application form as a strategy to gain eligibility. Either situation would introduce systematic differences between observable and unobservable characteristics of individuals just above and just below the eligibility threshold. Appendix Section A.4.3 discusses the results of a McCrary test – i.e., the density of the eligibility score and a smoothed density estimator based on a local linear regression on both sides of the threshold and other auxiliary evidence – which indicate that despite the

[21] Only a handful of households with a score below the cutoff participated in AFAM.

[22] It should be noted that registered employment fluctuates in the 30-40 percent range in the pre-application period, and in the 40-60 percent range in the post-application period. This is due to the fact that the program was launched when Uruguay's economy was recovering from the severe crisis of 2005, after which formal employment rose substantially. The pre-application period is closer in time to the crisis than the post-application period, and the difference in average registered employment for applicants in panels A and B in Figure 3 reflects the employment recovery during the period we study. Macroeconomic and labor market trends during our period of analysis in Uruguay are discussed in Appendix Section A.2.2.

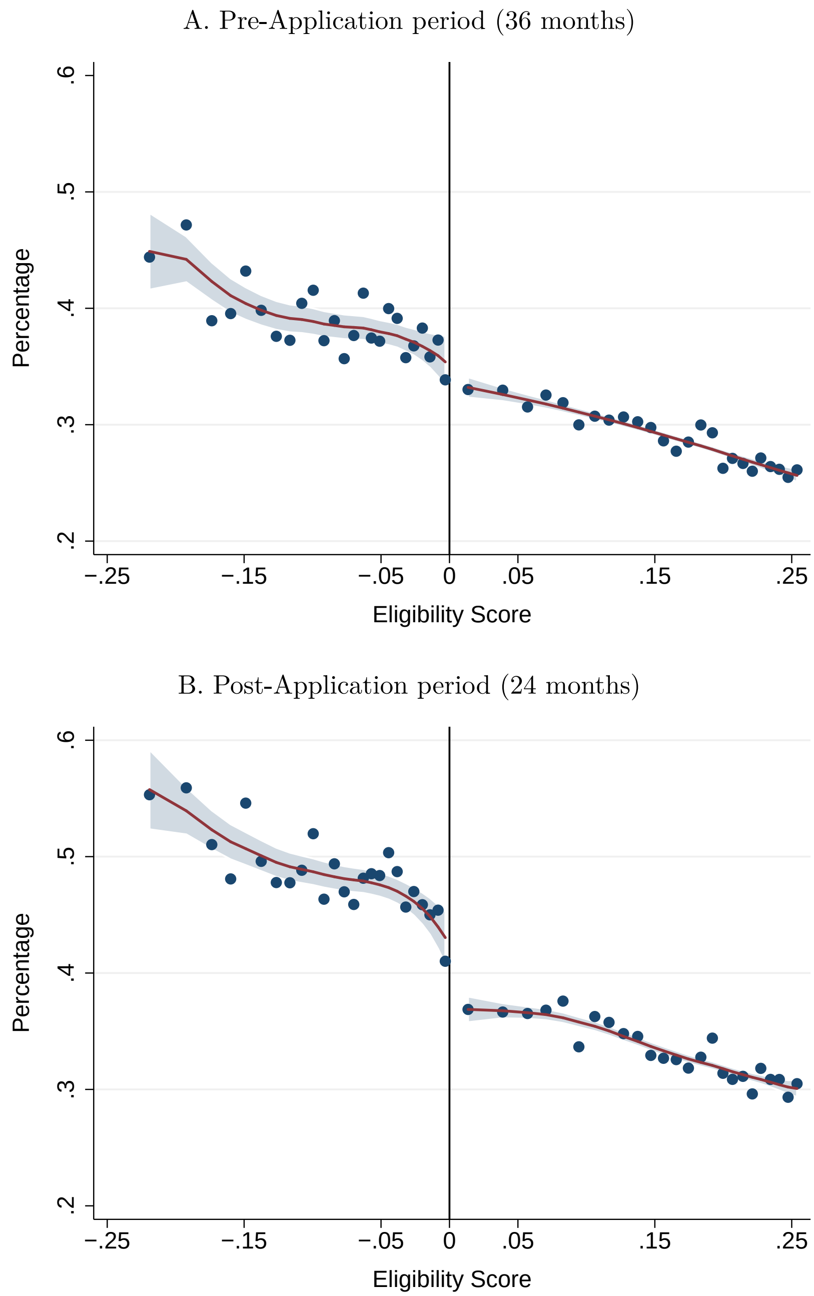
Figure 2: The AFAM Intake Process: Participation Rates and Eligibility Score



Notes: This figure plots participation in AFAM against the eligibility score for the sample of households with heads, and spouses of heads, aged 18 to 57 at the time they applied to the program. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating those in ineligible households. Each point (blue circle) in the plot represents the percentage of individuals in applicant households who have participated in AFAM in eligibility score bins of one percentage point. The solid red line plots estimated means from a local linear regression applied to each side of the eligibility threshold.

imbalance in some characteristics we cannot reject the null hypothesis of no discontinuity. Finally, besides this empirical evidence, the institutional context of AFAM also suggests the absence of selective sorting or of eligibility score manipulation by either beneficiaries or program officials. While individuals had incentives to complete the application form strategically to gain eligibility, they were limited by the fact that the government did not disclose the algorithm used to compute the score, the characteristics on which it was based, their underlying weights, or the level of the eligibility threshold. Moreover, the very large, sharp jump in participation at the threshold illustrated in Figure 2 indicates that the program's rules were strictly enforced.

Figure 3: Difference-in-Discontinuity: Registered Employment Rates by Eligibility Score, Pre- and Post-Application



Notes: This figure plots registered employment against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins of one percentage point. The solid red line plots estimated means from a local linear regression estimated at each side of the eligibility threshold, along with the 95-percent confidence interval.

While there are alternative ways to compute bounds in cases of a manipulated running variable in the context of RD (Gerard, Rokkanen, and Rothe, 2020), the empirical evidence does not point towards a case of manipulation. However, the imbalance in the main outcome (and in some of the covariates) may confound the impact of the program with these pre-application differences. We thus use an alternative to the standard RD model that exploits the longitudinal nature of our data in an RD framework to control for pre-application differences. Specifically, to identify the impacts of the program, we use a difference-in-discontinuity (DD-RD) research design that compares the discontinuity in post-application outcomes to the discontinuity in pre-application outcomes for individuals close to the threshold, thus controlling for potentially confounding pre-application differences around the threshold (see e.g., Grembi, Nannicini, and Troiano, 2016, Bertrand, Mogstad, and Mountjoy, 2019).[23] The DD-RD specification is similar to the standard RD estimation, but it also includes information from the pre-application period. Specifically, the DD-RD model that we estimate is given by the following equation:

$$Y_{it} = \alpha_0 + \beta_0 ELEG_i \times Post_t + \beta_1 ELEG_i + \beta_2 Post_t + \tag{2}$$

$$f(score_i)\left[\kappa_0 + \kappa_1 Post_t + \kappa_2 ELEG_i + \kappa_3 ELEG_i \times Post_t\right] + \lambda_t + \epsilon_{it}$$

where Y_{it} is the outcome of interest for individual i at time t; $ELEG_i$ is a dummy variable equal to 1 if the individual belongs to an applicant household eligible for the program (i.e., if $score_i > 0$), and zero otherwise; *score* is the value of the eligibility score, which in the RD literature is standardized relative to the eligibility threshold (c); $f(score_i)$ is an (unknown) functional form of the "assignment" or running variable *score*; and λ_t represents time fixed effects. Each variable's interactions with a post-application ($Post$) indicator represent the time dimension that the DD-RD model adds to the standard RD analysis.[24] The parameter β_0 captures the causal effect of AFAM on the outcome of interest. In the graphical interpretation, this coefficient is the difference between the jumps in registered employment at the discontinuities in panels A and B in Figure 3. [25]

The functional form of $f(.)$ and the window on each side of the cutoff threshold are key inputs for RD-based research designs.[26] Our baseline specification estimates this function

[23] Additional earlier studies based on RD designs that exploit longitudinal data include Lemieux and Milligan (2008), Cellini et al. (2010), Pettersson-Lidbom (2012), among others.

[24] Some subsets of our results are based on conventional RD estimates (for instance, those based on the follow-up survey, for which we only have one observation in time) which correspond to equation A.4.1 – i.e., the equation 2 excluding the $Post_t$ terms and the corresponding interactions.

[25] We implement the DD-RD estimator as follows: We first estimate equation A.4.1 using data from the pre-application period (36 months); we then subtract the RD regression coefficient from the outcome variable for those eligible, and we finally estimate the RD model using the post-application period (24 months). The RD coefficient of that regression corresponds to β_0 in Equation 2.

[26] See Lee and Lemieux (2010) for a detailed review of technical issues relevant to the RD research design.

using local linear regressions with a triangular kernel density for a given bandwidth (Cheng et al., 1997, Porter, 2003). When the analysis focuses on the *Main Sample*, we implement the bandwidth selection process following Calonico et al. (2014). Estimates based on the *Follow-Up Sample* use all of the survey's observations – the sample was drawn from an RD optimal bandwidth. We report MSE-optimal coefficients and standard errors (referred to as "conventional estimates" by Calonico et al., 2014), but rely on bias-corrected estimates for p-values, since these are optimal for inference in this context.[27] For estimates from the longitudinal *Main Sample,* we report robust standard errors clustered at the household level to account for potential serial correlation of outcomes for adults over time.[28] When we use the *Follow-Up Sample*, we report conventional robust standard errors.

The validity of our DD-RD design rests on the assumption that unobserved characteristics of eligible and ineligible individuals evolve similarly across time at the eligibility cutoff point. An indirect test of this assumption would be to check whether the discontinuities in observed characteristics, in the eligibility score, or in the outcome of interest change over time. Unfortunately, the covariates and the running variable are time-invariant in our setting – i.e., we observe this information only at the baseline.[29] However, given the longitudinal nature of our *Main Sample*, we can test whether there are differing discontinuity patterns around the eligibility threshold in our outcome of interest throughout the pre-application period. Reassuringly, inspection of Figure 4 (discussed in detail in the following section) indicates that the difference in registered employment at the threshold between eligible and ineligible individuals is roughly constant over the pre-application period.

5 Registered Employment Response to the Program

5.1 Registered Employment Response to AFAM

The essence of our main result is illustrated non-parametrically in panels A and B in Figure 3. AFAM generated a substantial decline in registered employment among its beneficiaries. In this section, we analyze this effect in greater depth. We first turn, in Table 1, to the detailed statistical results of the average effects of AFAM on registered employment over the whole

[27]The authors thank Gonzalo Vázquez Baré for this suggestion.

[28]We discuss program eligibility and income conditionalities at the individual level, although program participation and conditionalities were actually determined at the household level. We cluster standard errors at the household level in our empirical analysis to make up for this. However, the incentives to work or to not work, and to work formally or informally, may be a function of the combined responses and earnings potential of all household members (Galiani and Weinschelbaum, 2012). The elasticities we estimate in Section 5.3 might be biased if earnings potential is correlated across household members, although only about 10 percent of adults in our study sample belong to the same household.

[29]The results presented in the following sections are computed and presented with and without covariates to control for these baseline differences in characteristics. The results do not change significantly with the inclusion of these covariates.

period using the *Main Sample* of analysis. Column (1) reports the point estimates, while column (2) presents the optimal bandwidth used in the estimation following Calonico et al. (2014). Column (3), in turn, reports the average registered employment rate for ineligible individuals close to the threshold for the left-hand side of the bandwidth (which we label as our comparison mean). Column (4) presents the effects in column (1) as a proportion of the registered employment rates in column (3), while column (5) reports the number of observations (individuals by time). Rows (1) and (2) report estimates based on the RD specification given by equation A.4.1 using pre- and post-application data respectively. In turn, row (3) reports the main estimation results based on the DD-RD specification given in equation 2. All the estimates include time and date-of-application fixed effects to control for different potential underlying trends between eligible and ineligible groups. Panel A presents the estimates without socioeconomic controls, while panel B presents equivalent results, adding a standard set of covariates to the basic econometric models.[30]

The RD estimates in panel A show a large and statistically significant decline of 9.4 percentage points in the registered employment of adults eligible for AFAM relative to the rate for ineligible adults in the post-application period (column 1, panel A, first row). Meanwhile, the RD coefficient of the same model with registered employment in the pre-application period (second row) is also negative, smaller (-3.8 percentage points) and statistically significant at the $p<0.05$ level, which is consistent with the previous discussion regarding the balance in the pre-treatment outcomes between those eligible and those ineligible and with the evidence in Table A.4.3 and panel A in Figure 3. Accounting for this pre-application difference, our DD-RD estimate of the effect of eligibility for AFAM on registered employment is a negative and statistically significant effect of 5.6 percentage points (third row). In terms of the mean outcome of the comparison group, this effect represents a proportional decrease of about 12 % in rate of registered employment.

The regressions presented in panel B of Table 1 include controls for a series of predetermined covariates. The pre- and post-application RD estimates are statistically significant and similar in magnitude to those reported in panel A. The DD-RD coefficient for the impact on registered employment is slightly higher (6 relative to 5.6 – third row in panels B and A respectively), and they are both significant in statistical and economic terms. The pattern of results in Table 1 is consistent with the hypothesis discussed in Section 2.4: we find that eligibility for a social assistance program such as AFAM induced decline in registered or formal employment.[31]

Finally, we can also exploit the longitudinal nature of the administrative data to study

[30]The controls in the regressions are those described in Table A.1.1. They include individual characteristics such as gender, head of household status, age, marital status, educational level (in three categories), the number of children in the household aged 0-17, whether the household was enrolled in the PANES program, residence in Montevideo (Uruguay's capital), and whether the individual was employed. These covariates allow us to control for the imbalances documented in Table A.4.3, and they also allow us to gain precision

Table 1: Effect of AFAM Eligibility on Registered Employment

	Estimates	BW	Comparison Mean	%Δ w.r.t. (1)	Obs.
	(1)	(2)	(3)	(4)	(5)
Panel A. No Covariates					
Elegible(RD Post)	-0.094	0.049	0.462	-20.348	213,720
	(0.020)			(4.398)	
	[0.000]			[0.000]	
Elegible(RD Pre)	-0.038	0.062	0.373	-10.130	454,323
	(0.017)			(4.567)	
	[0.040]			[0.040]	
Elegible(DD-RD)	-0.056	0.050	0.462	-12.205	214,248
	(0.020)			(4.388)	
	[0.013]			[0.013]	
Panel B. With Covariates					
Elegible(RD Post)	-0.090	0.052	0.463	-19.412	225,768
	(0.019)			(4.042)	
	[0.000]			[0.000]	
Elegible(RD Pre)	-0.030	0.068	0.375	-8.048	516,779
	(0.016)			(4.156)	
	[0.061]			[0.061]	
Elegible(DD-RD)	-0.060	0.052	0.463	-12.948	226,056
	(0.019)			(4.039)	
	[0.003]			[0.003]	

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Panels A and B present the estimates from the RD – using either pre (row 1) or post (row 2) application data – and DD-RD (row 3) models in equations (A.4.1, 2) without and with socioeconomic covariates, respectively. Estimates from local linear regressions use a triangular kernel. The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. The covariates in the regressions in panel B include gender, head of household status, age, marital status, educational level (in 3 categories), the number of children aged 0-17 in the household, whether the household was enrolled in the PANES program, residence in Montevideo (Uruguay's capital), and whether the individual was employed. All regressions include time and date-of-application fixed-effects. Column (1) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "BW" in Column (2) reports optimal bandwidths following theCalonico et al. (2014) procedure. "Comparison Mean" in Column (3) reports the average of the dependent variable for ineligible individuals within the score bandwidth. Column (4) reports the program's effect from column (1) as a percentage of the mean of the dependent variable for ineligible individuals at the cutoff (columns 3). Column (5) reports the total number of observations.

the effect of AFAM eligibility on the evolution of registered employment over time before and after application to the program. We start with panel A in Figure 4, which gives a sense of the pattern of program participation after application. The figure presents the RD differences in the participation probability of eligible and ineligible groups since the household submitted its application, along with estimated confidence intervals.[32] The two vertical dashed lines indicate the semester when the household applied to the AFAM and the last semester of our data, respectively. As expected, the probability of program enrollment increases rapidly once the household applies and become eligible. In the application semester (t=0), eligible households are 31.8 percentage points more likely than ineligible ones to be enrolled in the program. This value increases to 90.6 percentage points in the third semester after application.

Figure 4 (panel B), in turn, plots the effect of the program on registered employment by semester, that is, the evolution of the RD estimates over time, with the baseline period given by each household's application date. The difference in registered employment between the two groups is not statistically significant in the semesters before application. In the application semester, when the difference in enrollment is still relatively low (31.8 percentage points), the difference in registered employment between the two groups is negative (3.2 percentage points) and significant at the $p<0.05$ statistical level. The difference in registered employment increases as the difference in enrollment rates increases. In the first semester after application, with an enrollment difference of 60.2 percentage points, the difference in registered employment is 7.9 percentage points, and it falls further in semesters 2 and 3 to a level of 9.2 percentage points when the enrollment difference is 72.3 and 90.5 percentage points respectively. There is a slight fall in the difference in registered employment to 8 percentage points in the fourth semester, when the difference in program participation rates also starts to fall (all differences after the application semester are highly statistically significant).

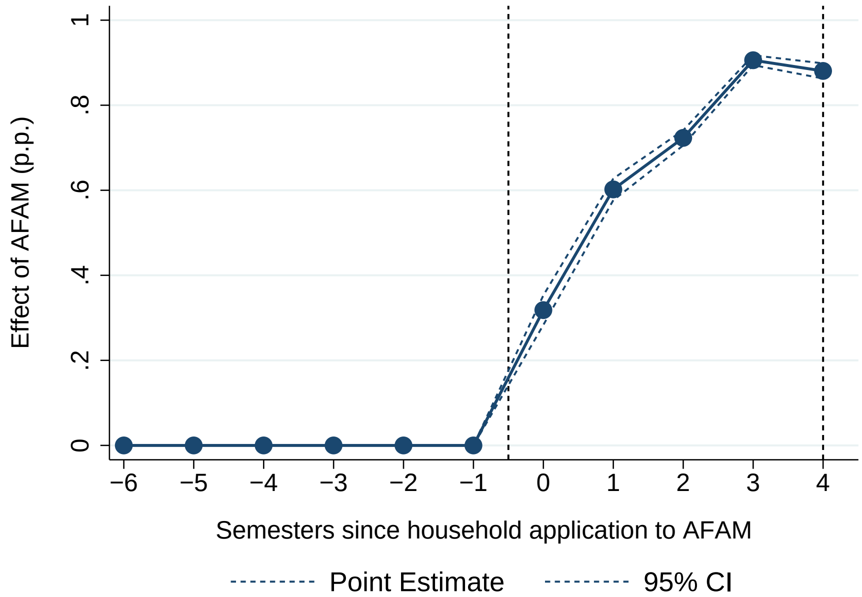
by reducing the residual variance.

[31]Except when noted, the regression results presented in the remainder of the paper are based on DD-RD estimates with covariates, because by controlling for pre-application differences they represent a more conservative estimate of the program's effects.

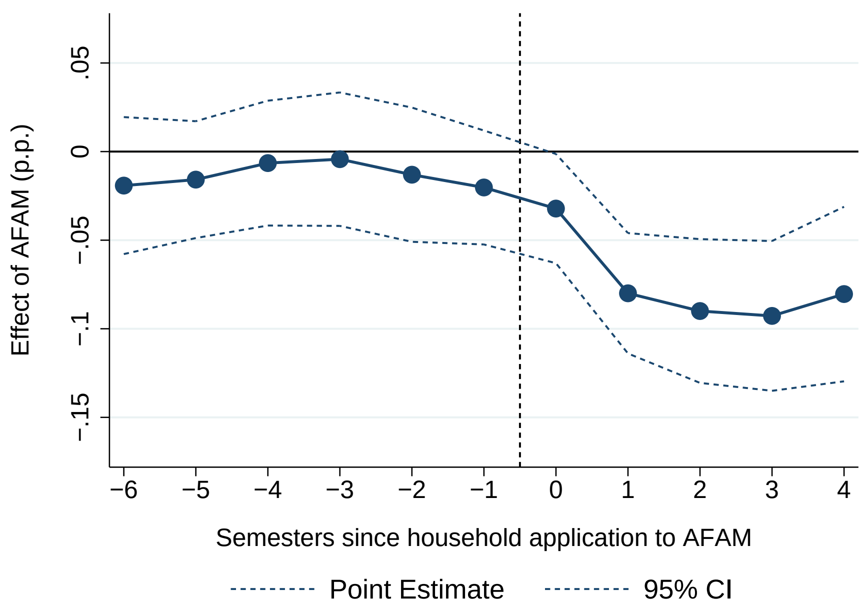
[32]While we report the RD difference for consistency with the bottom panel, inspection of Figure 2 indicates that enrollment for the non-eligible households was very low, and the RD differences reflect mostly enrollment levels for the eligible.

Figure 4: Effect of AFAM Eligibility on Registered Employment Over Time. Enrollment and Registered Employment Estimates by Semester

A. Program Participation Probability



B. Effect of AFAM Eligibility on Registered Employment Over Time



Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of AFAM application during the period January 2008-September 2010.The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Each point (blue circle) in the plot represents the estimate of the RD effect of eligibility on the probability of participation in AFAM (panel A) or registered employment (panel B) for a specific semester from one to six semesters before application to AFAM, and up to four semesters after applying to the program. The estimates correspond to a local linear regression, estimated separately on either side of the eligibility cutoff. Regressions in panel B. include covariates as controls (see notes to Table 1). The dotted blue lines plots a 95% confidence interval based on standard errors clustered at the household level.

5.2 Distributional Effects

The previous analysis considered average effects among all adult applicants. To elucidate the range of responses to the AFAM program, we now examine the degree of heterogeneity of responses by subgroups of individuals. Specifically, we analyze the heterogeneity of AFAM's impact by formal labor market attachment.[33] As Perry et al. (2007) argued, there are two ways of explaining the presence of informal work in an economy. According to the "exit" view, workers voluntarily choose to engage in informal employment because they place less value on the social insurance benefits tied to registered jobs relative to their costs. That is, they seek to avoid the associated payroll taxes and contributions because, in the context of weak enforcement of tax and labor regulations, informal employment results in better net-pay from a private cost-benefit calculation. According to the "exclusion" view, by contrast, workers are informal because formal jobs are harder to find (e.g. because search costs are higher for formal jobs). These two views can be considered complementary.[34] In any case, with or without segmented labor markets, theory predicts differential effects among individuals with different propensities to work as formal employees – i.e., those closer (marginal) or further away (infra-marginal) from the registered/not registered employment margin of choice (Galiani and Weinschelbaum, 2012). Even if all AFAM-eligible individuals face the same incentives from the program, their outside options might differ (for instance, given their levels of education), which is what our proxy for the probability of being formally employed (described below) captures. The negative effect on registered employment should be higher among individuals with higher propensities to be employed as formal workers relative to potential beneficiaries with lower propensities to be formally employed.

We combine a broad set of background characteristics into a single predicted registered employment index as a proxy for the propensity to work formally in the absence of AFAM. We estimate a regression model with registered employment as a function of our standard set of covariates only for the group of ineligible individuals (Appendix Table A.6.1 reports the results from this auxiliary regression). We then use the estimated coefficients to generate predicted values of the probability to work as a registered employee for all individuals in the sample (i.e., also including those eligible for AFAM). We then divide our sample into three roughly equally-sized groups based on this score. Finally, we estimate the effects of the program for each of the three subgroups separately. This approach follows what Abadie

[33]Responses to AFAM could also have varied as a function of other individuals' and households' observable characteristics. However, this does not seem to be the case. Appendix Table A.6.2 (Appendix A.6.2) presents the DD-RD estimates for different subgroups of the *Main Sample*. These estimates do not reveal any clear pattern of heterogeneous effects by type of individuals, with the exception of single female households, for whom the effect of AFAM is substantially larger than for other groups.

[34]Early versions of the exclusion view considered formal and informal sectors as segmented – i.e., the presence of dual labor markets (Fields, 2009). However, several studies using longitudinal data challenged this view by showing that workers often transit between formal and informal jobs in Latin America (Maloney, 2004).

et al. (2018) call the "endogenously stratified" group method.[35]

We perform the distributive analysis considering both the full sample of individuals and the group of single mothers. Single mothers represent a large share of our full sample (about 43%) of AFAM beneficiaries. Moreover, the literature on welfare programs' effects has documented stronger behavioral responses for this group compared to the rest of the population (Moffitt, 2002, Chan and Moffitt, 2018). Indeed, we find that single mothers is the group that responds most strongly to AFAM in our sample (Table A.6.2).

Table 2 reports the estimates of the effect of AFAM eligibility on registered employment for these three sub-groups, both for all beneficiaries (panel A) and for single mothers (panel B). The table follows the column layout of Table 1, but presents DD-RD estimates only.[36] Column (3) presents the average of the dependent variable, registered employment, for ineligible individuals in each subgroup at the cutoff. The average for the full sample is 46.3%, with a slightly lower level (45.2%) for single mothers. As expected, individuals in the group with the lowest propensity to work formally exhibit a low level of registered employment—13.4%—in the post-application period (panel A), with an even lower level for single mothers (12.3%, panel B). The levels are substantially higher for individuals in the middle and high propensity groups – 30.7% and 78.7%–, with lower levels for single mothers (25.0% and 71.5%).

The effects of the program on registered employment are presented in Column (1) of Table 2. The effect for the full sample is 6 points percentage negative (panel A, as in the panel with covariates in Table 1), and substantially higher in absolute value for single mothers – 8.7 percentage points (panel B). Both effects are highly statistically significant. The DD-RD estimate for the group with a low propensity to work formally is still negative but lower in magnitude and not statistically significant at conventional levels for the full sample (2.5 percentage points), and positive but not statistically significant for single women with a low propensity. For the group with a middle propensity, the effect is negative and substantially higher for the full sample (-8.9 percentage points) and for single mothers (-12.4 percentage points). Finally, it is also negative for the group with a high propensity to work formally at -4.7 percentage points for the full sample and -9.8 percentage points for single mothers, although only the latter is statistically significant at conventional levels. The effects as a proportion of the comparison mean differ substantially: they represent about -13% for the full sample and -19.3% for single mothers, but they are much larger for the middle-propensity group (-29% and -49.5%, respectively) than for the high propensity group (-5.9%

[35] Alternatively, we can use the individual history of registered employment in the pre-application period as a proxy for the propensity to be a formal worker. We define it as the proportion of months that an individual was a registered worker in SSA records during the 36-month pre-application period. The estimates remain qualitatively similar when we implement this alternative for group stratification. These results are available upon request.

[36] Appendix Tables A.6.3 and A.6.4 report the RD estimates underlying our DD-RD results using pre- and post-application data, respectively.

and -13.6%). The strongest behavioral response to the program, both in absolute and in relative terms, is observed for the group with a middle propensity to work formally, whereas the statistically insignificant effects for the low propensity groups (for the full sample and for single mothers) indicate that this group is much less responsive to the incentives introduced by the program.[37]

These results are consistent with the fact that individuals in these groups probably face different incentives and have different opportunities in the labor market. On the one hand, those with a low propensity to work formally may not respond much to the financial incentives of the program because they have limited opportunities to work as registered employees to begin with. On the other hand, individuals with a high propensity to work formally are less affected by the program's disincentives for registered employment, because most of them are likely to work formally regardless. The group with the middle propensity to be formally employed seems to be the one closest to the margin of choice between formal and informal employment, and thus individuals in this group are those who react most strongly to the new incentives. The presence of heterogeneous treatment effects in the dimension of propensity to work formally accords with the evidence documented in Gerard and Gonzaga (2020), who find heterogeneity across labor markets with different formal employment rates in Brazil. Our evidence complements these results by analyzing heterogeneity between groups within the same labor market.

[37]The difference between the coefficient for the middle propensity group and that for the other two groups is statistically significant for single mothers, although only the difference between the middle- and low-propensity groups is significant for the full sample.

Table 2: Effect of AFAM Eligibility on Registered Employment, by Propensity to be a Registered Employee, Full Sample and Single Mothers

	Estimates (DD-RD)	BW	Comparison Mean	%Δ w.r.t. (1)	Obs.
	(1)	(2)	(3)	(4)	(5)
Panel A. Full Sample					
Average	-0.060	0.052	0.463	-12.948	226,056
	(0.019)			(4.039)	
	[0.003]			[0.003]	
Low	-0.025	0.075	0.134	-18.572	113,568
	(0.017)			(12.478)	
	[0.182]			[0.182]	
Medium	-0.089	0.073	0.307	-29.062	118,056
	(0.025)			(8.239)	
	[0.000]			[0.000]	
High	-0.047	0.051	0.787	-5.926	86,496
	(0.026)			(3.342)	
	[0.104]			[0.104]	
Panel B. Single Mothers					
Average	-0.087	0.064	0.452	-19.265	147,672
	(0.023)			(5.119)	
	[0.000]			[0.000]	
Low	0.039	0.038	0.123	31.771	18,408
	(0.030)			(24.696)	
	[0.118]			[0.118]	
Medium	-0.124	0.056	0.250	-49.518	38,160
	(0.036)			(14.236)	
	[0.001]			[0.001]	
High	-0.098	0.052	0.715	-13.649	47,088
	(0.039)			(5.407)	
	[0.019]			[0.019]	

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Each row presents the estimates from the DD-RD model in Equation (2) with covariates at time of application to the program for the corresponding subgroup, as in the notes to Table 1. Estimates from local linear regressions use a triangular kernel. The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. Panels A and B present the regression estimates for the full sample and the sample of single mothers, respectively. The first row in each panel reports the average estimate for each sample, while rows 2 to 4 report estimates for the subgroups of individuals with Low, Medium, and High propensity to work as a registered employee, respectively. See Section 5.2 for a detailed explanation of the procedure used to determine these probabilities. Column (1) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "BW" in Column (2) reports optimal bandwidths following the Calonico et al. (2014) procedure. "Comparison Mean" in Column (3) reports the average of the dependent variable for ineligible individuals within the score bandwidth. Column (4) reports the program's effect from column (1) as a percentage of the mean of the dependent variable for ineligible individuals at the cutoff (column 3). Column (5) reports the total number of observations.

5.3 The Implied Elasticity of Participation in Registered Employment

The results presented thus far indicate that individuals eligible for AFAM respond to the program's financial incentives by reducing their participation in registered employment. In this section, we establish the magnitude of these responses in relation to the size of the program's financial incentives. We use these results to establish a more general parameter, the aggregate elasticity of participation in registered employment to net-of-(registered employment) participation tax rate – that is, the percentage change in formal employment for each percentage point change in the net average tax rate.[38] This parameter captures movements into and out of formal employment as a consequence of the tax and transfer schedule.

We compute the aggregate elasticity of participation in registered employment ϵ_R in the context of our DD-RD analysis by closely following the procedure developed by Kostol and Mogstad (2014). Specifically we define ϵ_R as:

$$\epsilon_R = \frac{1 - ptr_{ineligible}}{R_{ineligible}} \cdot \frac{\triangle R}{\triangle(1 - ptr)} \tag{3}$$

where $(1 - ptr_{ineligible})$ denotes the mean of the net-of-participation tax rate in registered employment for the ineligible group, $\triangle(1 - ptr)$ is the difference in the net-of-participation tax rate between eligible and ineligible individuals, $R_{ineligible}$ represents the registered employment rate for AFAM ineligible individuals, and $\triangle R$ denotes the difference in registered employment between eligible and ineligible individuals. It should be noted that the ptr corresponds to the effective average tax rate, which includes both payroll taxes on formal earnings and the loss of the program's benefit following entry into registered employment (when the resulting earnings exceed the program's income test threshold).

We obtain estimates of $R_{ineligible}$ as the average registered employment rate for ineligible individuals at the cutoff point, while $\triangle R$ corresponds to the DD-RD estimates with controls discussed in the previous section. The main challenge to compute this elasticity is to find a measure of the $ptr_{ineligible}$ – i.e., the implied ptr for the difference between disposable income when formally employed (subtracting payroll taxes) and when not formally employed (including the AFAM cash transfer when relevant). This requires measures of individual earnings and of net-of-tax transfers in both employment states. Since we do not have information on earnings in the SSA records, we impute them for formal and informal employment based on Uruguay's ECH household survey for the period 2008-2012 (i.e., the period covered by our SSA data). As expected, ineligible individuals have higher average earnings than eligible

[38] As discussed in Section 2.4, AFAM's income test eligibility rule introduces a discrete change in potential beneficiaries' budget sets. In the context of such a tax and transfer schedule, the participation tax rate is generally considered more relevant to behavioral responses than the marginal tax rate (Eissa et al., 2006, 2008).

individuals. We also compute individual net taxes and transfers based on Uruguay's tax code and on the AFAM transfer schedule, which we apply to each individual and her household characteristics. Importantly, the simulated schedule accounts for AFAM benefit loss when earnings from registered earnings are above the program's eligibility income threshold. Based on these inputs, we estimate the disposable income and compute the *ptr* for every individual in our database. Finally, we compute $\triangle(1-ptr)$ as the weighted difference in the net participation tax rate between eligible and ineligible individuals. The weights are given by the conditional density of the registered earnings of ineligible individuals who work as registered employees.[39]

Table 3 presents the estimated elasticities for the three subgroups of propensity to be a registered employee for individuals in the full sample (panel A) and for single mothers (panel B). Each row reports our estimates of the components of equation (3), and the final row in each panel presents the resulting elasticity and its confidence interval.

The change in registered employment, our proxy for $\triangle R$, is presented for each subgroup in the first row of each panel, and corresponds to the estimates reported in Table 2. The change for the full sample is -6 percentage points, and -8.7 for single mothers, with higher effects for the group with medium propensity to be a formal worker. The average registered employment level for ineligible individuals at the cutoff, our proxy for R, is 46.3% for the full sample and 45.7% for ineligible single mothers. These levels vary accordingly by propensity to be formally employed: the registered employment rates are 13.4%, 30.7% and 78.7% for those with low, medium and high propensities, respectively, to work as registered employees in the full sample, and 12.3%, 25.0%, and 71.5% percent for the corresponding subgroups of single mothers.

The third row in each panel presents the average change in the net-of *ptr*, given by the loss of the entire AFAM cash transfer at the income-eligibility threshold. While the implicit tax rate on the benefit is 100% at the threshold, we average out over the whole distribution to obtain the change in the average tax rate. This rate is -6.2% for the full sample and -5.5% for single mothers. The rate ranges from -5.5% to -7.1% for the three subgroups of the full sample, whereas it is more homogeneous for the three subgroups of interest among single mothers – from -4.7% to -5.3% . Finally, the forth row in each panel shows that the net-of-participation tax rate in formal employment for ineligible individuals is 37.5% for the full sample and 37.2% for single mothers, with little variation between the subgroups (36.6%-39.4% for subgroups of the full sample, and 35.8%-37.0% for single mothers).

The bottom row in each panel of Table 3 presents the resulting estimates for the registered employment elasticity. For the full sample, it stands at 0.784, and it is substantially higher

[39] Appendix A.3 details the earnings imputation procedure, the computation of the weights, and other details on the derivation of this elasticity in our context. Appendix Figure A.3.1 illustrates the net-of-tax *ptr,* i.e. 1-*ptr*, and the density of simulated registered employment earnings by eligibility status for the year 2011.

Table 3: Elasticity of Participation in Registered Employment

		Propensity to be formal		
	Average	Low	Medium	High
	(1)	(2)	(3)	(4)
Panel A. Full Sample				
Δ Registered Employment	-0.060	-0.025	-0.089	-0.047
	(0.019)	(0.017)	(0.025)	(0.026)
Registered Employment (Ineligible)	0.463	0.134	0.307	0.787
	(0.499)	(0.340)	(0.461)	(0.409)
Δ Net-of-Participation Tax Rate	-0.062	-0.055	-0.065	-0.071
	(0.001)	(0.001)	(0.001)	(0.001)
Net-of-Participation Tax Rate (Ineligible)	0.375	0.394	0.368	0.366
	(0.026)	(0.023)	(0.033)	(0.026)
Elasticity (ϵ_R)	0.784	1.336	1.641	0.308
	[0.612-0.912]	[0.987;1.812]	[1.235;1.911]	[0.224-0.424]
Panel B. Single Mothers				
Δ Registered Employment	-0.087	0.039	-0.124	-0.098
	(0.023)	(0.030)	(0.035)	(0.039)
Registered Employment (Ineligible)	0.457	0.123	0.25	0.715
	(0.498)	(0.328)	(0.433)	(0.451)
Δ Net-of-Participation Tax Rate	-0.055	-0.047	-0.053	-0.052
	(0.001)	(0.002)	(0.001)	(0.001)
Net-of-Participation Tax Rate (Ineligible)	0.372	0.358	0.370	0.370
	(0.060)	(0.057)	(0.060)	(0.061)
Elasticity (ϵ_R)	1.288	-2.415	3.463	0.975
	[1.080;1.428]	[-3.684;-1.362]	[2.688;3.989]	[0.776;-1.230]

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within an eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records, and to the ECH for the period 2008-2012 (see Section 3.1 for a description of the data). The elasticity formula is given in Equation (3) and discussed in Section 5.3 and A.3. Panels A and B report the estimates for the full sample and the sample of single mothers, respectively. In the rows, "△Registered Employment" represents the differences in registered employment between eligible and ineligible individuals as given by the DD-RD estimates in Table 2. "Registered Employment (Ineligible)" denotes the registered employment rate for ineligible individuals at the cutoff (the comparison group). The "Net-of-Participation Tax Rate (Ineligible)" is given by Equation A.3.7, and computed as the mean value for ineligible individuals (see Section A.3 for a detailed explanation). The differences in this tax between eligible and ineligible individuals are given by Equation A.3.6. Column (1), "Average", refers to the estimates for the whole sample. Columns (2) to (4) report the estimates for the subgroups of individuals with Low, Medium, and High propensity to work as a registered employee, respectively. Confidence intervals on the estimated elasticity (square parentheses) are computed using a bootstrap method.

(1.288) for single mothers. These figures mean that a reduction of 1% in the net-of-tax share of income that individuals can keep reduces registered employment by about 0.8% for the full sample, and by about 1.3% for single mothers.

Consistent with the previous results regarding the heterogeneity of AFAM's impact, we find a substantially larger elasticity for the group with a medium propensity to be a registered employee: 1.641 for the full sample, and 3.463 for single mothers, reflecting the fact that these are probably the individuals closest to the formal/informal margin. The elasticities are lower for the group with a higher propensity to be formal (0.308 for the full sample and 0.975 for single mothers), reflecting the fact that individuals in this group are likely to work as registered employees in any case. Finally, the elasticities for the group with a low propensity to be formally employed are large and positive for the full sample (1.336), due to the low rate of registered employment, and large and negative for single mothers (-2.415), due to the positive estimated impact on registered employment for this subsample. It should be noted that even if the confidence intervals for the last two elasticities do not include zero, the reduced-form effects (-0.025 and 0.039) are not statistically significant at standard levels (i.e., p-values are 18.2% and 11.8%, respectively).

Our results for Uruguay indicate that individuals in developing countries make decisions about entering formal employment in response to the incentives implied by the tax and transfer schedule they face. Moreover, these estimates accord with those found in developed countries. For instance, the corresponding elasticity for single mothers in the USA derived from variation in the generosity of the AFDC program is about 0.7 (based on estimates from Meyer and Rosenbaum, 2001), compared to our estimate of about 1.3 for this group. However, in a context like ours with high informality, our estimates reflect the response in terms of non-participation and also in terms of movements between formal and informal employment. We find substantial heterogeneity along this dimension: the elasticity is much higher (about 3.5) for single mothers on the margin of being formal employees, with lower levels for those with either very high or very low propensities to work formally. The following section attempts to shed some light on how our results concerning lower registered employment map onto these two alternative outcomes.

5.4 Decomposing the Effects of the Program: Informality and Non-Employment

The main analysis in the previous section was based on the *Main Sample*, a longitudinal dataset of formal work histories from *SSA registered employment records*. As discussed above, this high-quality data source has a limitation in terms of the outcomes of interest: while we can establish the effects of the program on changes in registered employment, we cannot separate the responses in terms of inactivity or non-registered employment –

i.e., we cannot know whether the null registered employment status in the administrative data represents a period in which the individual unemployed or, alternatively, a period of informal employment. This is an important distinction for understanding how programs such as AFAM interact with a labor market featuring high informality.

Our *Follow-Up Survey* provides information to complement the benchmark specifications for registered employment, allowing us to complete the anatomy of responses to the AFAM's financial incentives by establishing the comparative magnitude of responses along the non-employment and informal work margins. The analysis, however, has some clear limitations. Unlike the *Main Sample*, the *Follow-Up Sample* is a cross-sectional dataset, and a relatively small sample at that. We thus cannot compute DD-RD estimates and must rely on standard RD, which may be biased because we cannot correct for factors such as the baseline differences in the main outcome identified in Section 4.[40]

Moreover, we do not have enough statistical power to distinguish between the full sample and single mothers, and thus we only present results for the former. For these reasons, we interpret the following results as merely suggestive.

Table 4 presents the baseline estimated RD effects (i.e., equation 2 without the interaction terms) for this sample on three mutually exclusive labor market outcomes: registered employment, non-employment and unregistered or informal employment. For comparison purposes, we present two sets of results. First, in column (1) we present the RD estimates by using the self-reported outcome from the *Follow-Up Survey*. Second, in column (2), we present our preferred set of RD results based on the sub-sample of individuals whose self-reported registered employment coincided with their registered employment status in the SSA records at the time of the survey's interview.[41] We denote this group the "consistent sample".

The results in Table 4 indicate that the negative effect on registered employment is similar in size for the different estimations, and all estimates are statistically significant at the $p<0.01$ level (row 1). Our preferred estimates (those based on the consistent sample) indicate a statistically significant effect on registered employment of about 10 percentage points, which might be biased upwards because we are not controlling for baseline differences in the outcome. The following two rows present the main result from this exercise. Focusing on the consistent sample (column 2) as a benchmark, the 10 percentage points reduction in registered employment induced by AFAM corresponded to an increase in non-employment of about 5 percentage points, and a similar increase in informal employment. However, only

[40] Appendix Table A.4.3 presents the covariate balance tests for adults in the *Follow-Up Sample* and Figures A.5.4 to A.5.6 the depict the corresponding RD plots.

[41] Appendix Tables A.6.5 through A.6.7 present cross-tabulations of self-reported and SSA-recorded registered employment. The classification errors in the registered employment variable are about 10 percent, and are similar in magnitude when conditioning by eligibility status (10% for eligible individuals vs. 8% for the ineligible). Figure A.5.7, in turn, indicates that the RD difference in these reporting errors is balanced by eligibility condition.

Table 4: Effect of AFAM Eligibility on Different Margins of Participation

	Self-Reported Follow-Up Survey (FUS)	SSA Records=FUS	Comparison Mean
	(1)	(2)	(3)
1. Registered Employment	-0.109	-0.100	0.548
	(0.041)	(0.042)	
	[0.003]	[0.002]	
2. Non-Employment	0.050	0.051	0.189
	(0.036)	(0.037)	
	[0.056]	[0.043]	
3. Informal Employment	0.059	0.049	0.263
	(0.043)	(0.044)	
	[0.218]	[0.209]	
Joint p-value (2 & 3 = 0)	0.065	0.104	
Joint p-value (2 = 3)	0.911	0.975	
Observations	2,403	2,157	

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010. This sample corresponds to the subset of individuals who were interviewed for the program's follow-up survey during the period September 2011-February 2013. The survey's sample was drawn from households within the eligibility score range $[-0.0426; +0.0727]$ (see Section 3.1 for more details). Each row presents the estimates from the DD-RD model in Equation (2) excluding the $Post_t$ terms and the corresponding interactions and with covariates at the time of application to the program, as the notes to Table 1 explain. Estimates from local linear regressions use a triangular kernel. In the first row, the dependent variable is the registered employment status (one if she is working as a registered employee, zero otherwise). In the second row, the dependent variable is non-employment, which is equal to on1 if the individual is not working, and equal to zero if she is working (as a registered or as a non-registered employee). In the third row, the dependent variable is informal employment at the time of the interview, which is equal to 1 if the individual is working as an informal (or non-registered) employee, and zero otherwise (i.e., if she is working as a registered employee or not working). The fourth (fifth) row presents the p-value of the statistical test that the eligible coefficients in the "Non-Employment" and "Informal Employment" regressions are jointly equal to zero (equal). Column (1), "Self-Reported Follow-Up Survey (FUS)", presents results for individuals in the *Follow-Up Sample* using the self-reported information they provided at the interview to measure the outcome variables. Column (2), "SSA Records=FUS," reports results based on the subsample of individuals who reported being employed as registered employees in a manner consistent with the SSA records for the same period when they were interviewed. We report "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "Comparison Mean" in Column (4) reports the average of the dependent variable for ineligible individuals close to the cutoff, i.e. within the $[-0.0426; 0]$ score bandwidth.

the coefficient for non-employment is statistically significant at the usual statistical levels. The p-value for the test that both coefficients are equal to zero is 0.104 (row 3), and we cannot reject the null hypothesis that they are both equal (p-value of 0.975, in row 4).

In the discussion of the expected effects of the program, we proposed that the cash transfer could cause an increase in non-employment or a shift towards informal employment. Even with the study's limitations, the suggestive evidence in this section indicates that both effects are probably present in the context of AFAM.

6 Efficiency Cost Implications of the AFAM Program

The empirical analysis performed thus far has shown that the AFAM program might generate behavioral distortions in the labor market by reducing participation in registered employment. While this might be an important policy question, the ultimate relevant outcome is the welfare impact of the program and its incentives. We follow Hendren's (2016) approach to translate the estimated causal effects of a public policy change to a measure that captures its efficiency cost implications.[42] The key statistic from this approach is the marginal value of public funds ($MVPF$).

This measure is given by the ratio of marginal benefits to marginal costs of the extra funds spent on the policy. Specifically, the marginal benefit is defined by the beneficiaries' marginal willingness to pay (WTP) for the increase in expenditure out of their own income to the cost to the government of the additional funds spent per beneficiary. This net cost includes both the increase in government expenditure due to the policy in the absence of any behavioral change (denoted as the "*mechanical cost*") and all other impacts on the government budget of behavioral responses to the policy (denoted "*fiscal externalities*"). Thus, the $MVPF$ of a policy j is defined as*:*

$$MVPF^j = \frac{WTP^j}{Net\,Cost^j} \tag{4}$$

In our setting, the object of interest is the $MVPF^{AFAM}$ of the extra local currency unit (UYU) spent on the AFAM program. If we make the plausible assumption that AFAM's beneficiaries value each UYU in monetary transfers as cash earnings, the envelope theorem suggests that they are willing to pay for the mechanical cost of the AFAM benefit – i.e., UYU 1 transfer is valued at UYU 1.[43] Normalizing the $MVPF$ in equation 4 by the mechanical

[42]Hendren (2016) shows that the government is the only distortion between private prices and social costs. Thereby, the causal (behavioral) response to a policy change regarding the government's budget is the only statistic required for welfare calculations.

[43]For AFAM beneficiaries who do not change their registered employment behavior in response to the program, the 1 UYU transfer is valued at 1 UYU. For those who modified their behavior because of the

cost of AFAM, the *MVPF* is valued at UYU 1 in mechanical cost per program beneficiary. Thus, the MVPF of an extra UYU 1 spent in the AFAM cash transfer is:

$$MVPF^{AFAM} = \frac{1}{1 + FE^{AFAM}} \tag{5}$$

where FE^{AFAM} describes the effect of the behavioral response to AFAM on the government's budget (i.e., the fiscal externality) per UYU increase in the mechanical cost per AFAM beneficiary. Importantly, $MVPF^{AFAM}$ captures the relative gain for beneficiaries of receiving the AFAM transfer vs. receiving a lump-sum from the government, holding the government budget constant. Thus, in our setting, $MVPF^{AFAM}$ measures the efficiency cost of the AFAM program. In what follows, we describe how we estimate $MVPF^{AFAM}$.

WTP: The willingness to pay for the AFAM transfer is given by the marginal benefit of the program per UYU 1 of mechanical cost, and is equal to UYU 1 (numerator in equation 5). Intuitively, this means that increasing the cash benefit by UYU 1 is valued at UYU 1 for the average AFAM recipient. We do not need the estimated causal effects for this estimate.

Net Cost: This component represents the marginal cost of increasing the AFAM cash transfer by UYU 1, and it is given by $1 + FE^{AFAM}$ (the denominator in equation 5), which adds the fiscal externality per UYU 1 induced by AFAM to the mechanical cost (UYU 1). In our calculations, we only add the fiscal externality resulting from the reduction in adults' registered employment as a response to AFAM during the period under study – i.e, two years after the beneficiaries enrolled in the program. Thus, the estimated *MVPF* does not account for long-run responses in the labor market nor for spillover effects on children (for instance, in terms of increased education given the program's conditionalities). We discuss below the implications of including these effects in our *MVPF* estimate.

The impact of the decline in registered employment in response to AFAM has two fiscal externality components that we incorporate in FE^{AFAM}. First, the behavioral response reduces tax revenue – i.e., payroll taxes – collected by the authorities because of the fall in registered employment earnings. Second, AFAM's effects on registered employment also imply higher government expenditures, because in the absence of the program some of the individuals who would be registered workers would also not draw the AFAM benefit. Note that some of the individuals who become informal would have been AFAM beneficiaries even when formally employed, because their earnings would be below the eligibility threshold. Among those who changed their behavior, we should include in our accounting only the fraction who received AFAM payments because of their behavioral responses.

program (e.g., they exited registered employment to remain eligible), the envelope theorem implies that, since they already are optimizing, they should be indifferent to the policy change. Their willingness to pay is thus not counted for these computations – we only take into account the willingness to pay of infra-marginal individuals.

We need to make a series of assumptions to estimate these components. First, we assume that the decline in registered employment due to AFAM does not create fiscal externalities for other tax bases. Although capital income taxation is not an issue for AFAM beneficiaries, this assumption rules out fiscal externalities with respect to other important tax bases in developing countries, such as consumption. Second, we also rule out fiscal externalities of the AFAM program with respect to take-up of other social assistance programs. Third, as in our previous calculations, we assume that the AFAM population is only subject to the payroll tax, which amounts to about 32% of gross earnings (the sum of the employee and employer taxes) and is paid by the employer. Fourth, we use the distribution of *imputed* formal earnings (used for calculations in Section 5.3) for eligible individuals to compute the average counterfactual (gross) formal earnings of AFAM beneficiaries who reduced their level of registered employment because of the program and the proportion of individuals that would be above or below the eligibility threshold in this counterfactual.[44] The earnings distribution is based on imputation procedure developed to compute elasticities in Section 5.3 (see also Appendix A.3). Finally, we assume that the only mechanical cost to the government for providing the program to the beneficiary household is the cash transfer – that is, we do not factor in the program's administration and monitoring costs.

For the purposes of simplification and of comparability with previous studies, the following calculations illustrate the case of single mothers. Table 5 reports main inputs for the calculations and our results. Column 1 shows results for the average single mother, while columns 2 to 4 report the estimates based on breaking down the sample of single mothers into the three groups according to their propensity to be registered employees, as described in the previous sections.

Row A in Table 5 reports the mechanical cost calculations, which correspond to the average monthly benefit of AFAM per single mother household. We apply the benefit determination rule given by equation 1 (Section 2.3) to compute the AFAM benefit for the average single mother household and for the three subgroups. The average number of children per household is similar between groups, which explains the very similar values of the benefit between groups.

Row B, in turn, presents the change in payroll tax revenue as a consequence of AFAM-induced behavioral responses. We use the DD-RD estimates of the program's effect on registered employment for the average eligible single mother and for the three subgroups (rows 5 to 8 in Table 2). Second, we apply a 32% tax rate on earnings. Third, we need an estimate of the average counterfactual earnings from registered employment – i.e., the

[44] Alternatively, we could have used the imputed distribution for ineligible individuals as a counterfactual to compute the statistics needed for MVPF calculations. The results are qualitatively the same and quantitatively very similar under this alternative assumption. For instance, the MVPF for all single mothers in this alternative is 0.55, similar to the value of 0.61 reported in Table 5. The full set of alternative results analogous to those in Table 5 is available upon request.

gross registered employment earnings that single mothers would have accrued were they ineligible for AFAM. We proceed as follows: The DD-RD estimates indicate an average 8.7 percentage point decline in registered employment, a 19.3% decrease with respect to the average employment levels of the ineligible single mothers around the cutoff point (column 1). We compute gross counterfactual earnings from registered employment of UYU 9,597 for an average single mother beneficiary in 2011. A 32% tax rate on earnings implies that the net change in tax collected from an individual due to the AFAM-induced reduction in registered employment is equal to UYU 592.7 ($0.193 \times 9,597 \times 0.32$). Similarly, we compute the change in payroll tax revenue for single mothers according to their propensity to be formally employed (columns 2 to 4). It should be noted that the DD-RD estimate for single mothers with a low propensity to be formally employed is positive, in contrast to the negative estimate for the average of the group and for the full sample. While the coefficient is not statistically significant at standard levels, we report the results for this group for completeness.

Row C in Table 5 reports the calculations for the loss of potential government budget savings. The average monthly AFAM benefit for eligible single mothers is UYU 1,086, and this represents the unitary cost of the program for this group of beneficiaries (column 1). Based on the distribution of imputed registered employment earnings, we estimate that 55.3% of single mothers would have had earnings from registered employment above the income-eligibility threshold. Recall also that we estimate a 19.3% decline in registered employment for single mothers. Therefore, the percentage of single mothers who receive AFAM benefits because of their program-induced decrease in registered employment is 10.7 = ($0.553 \times 0.193 \times 100$). The government thus foregoes average savings of about UYU 115.9 = ($1,086 \times 0.107$) per single mother due to the beneficiaries who would have been engaged in registered employment in the absence of a behavioral response to the program and, thus, would not have drawn the benefit because they would not have passed the income test. We proceed analogously for the three subgroups, noting that for those with a low propensity to be a registered employee AFAM represents a revenue increase, given the positive estimate of the program's effect on registered employment for this group.

The average total cost of the AFAM program per single mother is thus UYU 1,794.6 – i.e. the sum of the mechanical cost (UYU 1086, row A) and behavioral responses (UYU 592.7 plus 115.9 , rows B and C, respectively). This implies a total fiscal externality of UYU 0.65 ($[592.7 + 115.9]/1,086$) as reported in row D. About 84% (0.55/0.65) of the total fiscal externality corresponds to the loss of tax revenue due to the decline in registered employment, while roughly 16% (0.11/0.65) can be attributed to the loss of potential government budget savings. Columns 2 to 4 in row D show the fiscal externality estimates for single mothers according to their propensity to be formally employed.

Finally, the marginal value of public funds for the AFAM program for single mothers, $MVPF^{AFAM}$, is UYU 0.61 (1/[1+0.65]). An MVPF below 1 indicates that individuals would prefer a UYU 1 lump-sum transfer as opposed to UYU 1 in AFAM benefits. Our result implies that for every UYU 1 spent by the government on AFAM benefits, AFAM beneficiaries would be willing to pay UYU 0.56. This figure is higher – 0.67 – for the subgroup of single mothers with a high propensity to be a registered employee and lower – 0.37 – for the group with a medium propensity. Following Hendren and Sprung-Keyser (2020), we assign an infinite MVPF value to the group with a low propensity to be a registered employee, because this group exhibits a positive willingness to pay for the program but a negative net cost for the government. While this finding implies that the policy "pays for itself" for this group, the effect on registered employment for this group is not statistically significant.

Table 5: Estimates of AFAM's Marginal Value of Public Funds (MVPF) for Single Mothers

Component	Propensity to be formal			
	Average	Low	Medium	High
	(1)	(2)	(3)	(4)
A - Mechanical Cost (UYU)	*1,086*	*1,080*	*1,080*	*1,126*
Average monthly AFAM benefit (UYU)	1,086	1,080	1,080	1,126
B - Change in Tax Revenue (UYU)	*592.7*	*- 956.7*	*1,534.4*	*448.3*
Registered Employment Change (%)	19.3	-31.7	49.5	13.6
Payroll Tax Rate on Gross Earnings (%)	32	32	32	32
Gross Earnings of Single Mothers (UYU)	9,597	9,431	9,687	10,302
C - Loss of Budget Saving (UYU)	*115.9*	*-171.2*	*304.2*	*105.9*
Average monthly AFAM benefit (UYU)	1,086	1,080	1,080	1,125
Registered Employment Change (%)	19.3	-31.7	49.5	13.6
% with Formal Earnings > Threshold	55.3	50.0	56.9	69.2
D - Fiscal Externality (B+C)/A	*0.65*	*-1.04*	*1.70*	*0.49*
MVPF (1/1+D)	0.61	∞	0.37	0.67

Notes: Estimates of "Average Monthly AFAM Benefit" are obtained from the AFAM administrative data. "Registered Employment Change" is taken from Table 2Estimates of "Gross Earnings of Single Mothers" and "% with Formal Income > Threshold" are obtained from Section A.3.3. The "MVPF" estimates follow equation 5.

.

This simple MVPF calculation provides an estimate of the order of magnitude of the efficiency cost implications of the program's effects on registered employment. However, our simple estimates abstract from some potentially important additional effects of the policy. First, our MVPF calculation ignores potential effects of AFAM on beneficiaries' children. For instance, the program's conditionalities could induce an increase in health check-ups and in school attendance, in turn generating an improvement in their educational, health and labor market outcomes in the long run. An increase in the earnings of beneficiaries' children represents a positive externality that would increase the MVPF with respect to our baseline calculation. While this scenario is plausible, the available preliminary evidence indicates that AFAM did not, in fact, induce a significant increase in health checkups or in school attendance (Bergolo et al., 2016). Moreover, while several studies of other developing countries find positive long-term impacts of CCTs on children's human capital accumulation, they fail to find higher lifetime earnings (Millan et al., 2019). Second, our MVPF calculation relies on a short-term impact of AFAM on the registered employment levels of adult beneficiaries. If the cash transfer distorts labor market behavior in the short run but has a positive impact on (registered) labor earnings in the long run, we should impute a higher MVPF to the program. However, this type of longer-run effect on employment is unlikely, because AFAM did not include any job training or job intermediation component.

An additional value of computing the MVPF is that we can compare our results to those obtained for other similar programs. Hendren and Sprung-Keyser (2020) report the MVPFs of 133 historical policy changes in the USA. Their evidence suggests lower MVPFs for policies targeting adults/households, as in the case of AFAM, than for policies that directly target children. These lower MVPFs reflect the fact that most of the programs targeting adults reduce (reported) labor earnings through labor market distortions. Thus, as in the case of the AFAM program, the average cost per $1 of mechanical government spending for these policies is greater than $1 (Hendren and Sprung-Keyser, 2020, Figure V). Our MVPF estimate for single-mother AFAM beneficiaries of 0.61 implies an efficiency cost within the range of cash transfer programs targeted to families in the United States, for which Hendren and Sprung-Keyser (2020) estimate an average of 0.74 and a confidence interval of [0.36,1.47] (their Table II). However, our estimate for AFAM is lower than the MVPF of similar programs that induce labor supply disincentives, such as AFDC (MVPF of 0.87). Despite pervasive informality in the AFAM context, the efficiency cost of this policy is still within the range of that of cash transfers in the United States, where informality is substantially lower.

7 Conclusions

The AFAM program's eligibility rules, which are based on verification of reported earnings, create a strong disincentive for registered employment – a notch in beneficiaries' budget constraints. Individuals in households that applied to the program reacted as predicted by economic theory. We built an anatomy of labor market responses induced by the program's financial incentives along four dimensions. First, we established that beneficiaries responded to the program's incentives by reducing their levels of registered employment. Second, we found substantial heterogeneity in these effects: the program had virtually no effect on individuals with a low propensity to be formally employed and a greater-than-average negative effect on individuals with a medium propensity to be a registered employee. Those with a low propensity to work formally did not respond much to the program's financial incentives, probably because they had limited opportunities to work as registered employees to begin with. This suggests that different groups face different incentives and opportunities, and policy design should try to accommodate these differences. Consistent with the findings for developed countries, we also found evidence of strong behavioral responses to social assistance for single mothers. Third, by matching administrative data with a follow-up survey, we provide suggestive evidence that the decline in registered employment translated into both an increase in unregistered employment and a shift towards non-employment. Fourth, in addition to the strong financial incentives of the AFAM program, the behavioral responses we observed translate into an elasticity of participation in registered employment for the whole population of about 0.78, and of 1.3 for single mothers. Finally, the transfers and registered employment responses we estimate yield a MVPF of AFAM on the order of 0.61.

Our evidence shows that neither the extensive margin of (formal) employment elasticity nor the efficiency cost of a welfare program disincentivizing (formal) employment are particularly large in a context of high informality, compared to the context of developed countries such as the United States where informality is much lower. We also highlight important heterogeneity underlying this result. As in developed countries, single mothers are more responsive than is the average beneficiary. Moreover, the elasticity is much greater for workers at the margin of formal employment and much less for (infra-marginal) workers with either low or high attachment to the formal sector, although the efficiency cost does not vary as much between these groups.[45]

Our results allow us to draw valuable conclusions that can be used to minimize programs' disincentives and to improve the equity-efficiency trade-off of policies of this type. In designing cash transfer programs in contexts of high labor informality, policymakers in developing countries must consider reactions along the registered/unregistered employment margin and find ways to mitigate these unintended adverse results, although our findings in

[45] We thank an anonymous referee for this interpretation of the results.

this regard are only suggestive and we need more evidence regarding these margins. Mitigating movements from formal to informal employment requires reducing the implicit tax rates or perceived tax wedges on formal earnings, and, wherever possible, modifying the program to incorporate incentives to work formally. Hendren and Sprung-Keyser (2020) show that programs that incorporate these incentives have lower efficiency costs – for instance, the AFDC as compared to the EITC.

A first implication of our analysis is that AFAM authorities (and those designing similar initiatives in countries with widespread informal employment) should consider ways to smooth out the cash notch implied by the program's eligibility rule. For instance, the income test could be modified by introducing a more continuous schedule with phase-in and phase-out regions that withdraw the benefit gradually and with lower implicit tax rates rather than generating a total loss of the benefit at the income threshold. Alternatively, authorities could test the effects of allowing beneficiaries to continue in the program while generating earnings from registered employment above the threshold for a transitory period. This temporary disregard of earnings could ease the transition to formal employment and mitigate the possibility that beneficiaries prefer lower but more stable income from the program to higher but riskier earnings from registered employment. Moreover, the employment of low income workers is often seasonal – e.g., tourism is an important industry in Uruguay with a marked seasonal pattern of high peaks during the three summer months. Allowing short spells of registered employment without total loss of AFAM benefit would probably reduce the negative effects of the program on registered employment.

Our results also highlight the importance of computing the efficiency effects of welfare policies in developing countries. A full welfare analysis might yield unexpected conclusions. For example, in studying the unemployment insurance program in Brazil, Gerard and Gonzaga (2020) found relatively low efficiency costs, in part because high informality levels preclude larger impacts to begin with.

The significant heterogeneity of responses to the program as a function of individuals' propensity to engage in formal employment also has direct consequences for the design of programs of this type, and particularly for their incentive structures. For instance, specific incentives or conditionalities that are costly to monitor (such as minimum working hours requirements) may only be worthwhile for those reasonably expected to react to the program's incentives. The poorest beneficiaries, who are also those with the lowest propensity to be a formal employee—a characteristic that might be proxied by low education levels—seem to be infra-marginal in terms of the formal/informal work decision. This means that concerns about disincentives to pursue registered employment are less relevant for this group. They could benefit more from other policy initiatives such as job training, job intermediation, or public employment programs. As highlighted by Hendren and Sprung-Keyser (2020), the marginal return on policies that transfer resources to adults are relatively low if they are not accompanied by measures to promote beneficiaries' income-generating prospects. At the same time, the high responsiveness of less poor beneficiaries suggests that the program could incorporate differential incentives for this group. Just as work incentives guided welfare reforms in developed countries, the new generation of cash transfer and social assistance programs in developing countries could incorporate measures to reduce disincentives to engage in formal employment. More broadly, concerns about the efficiency of welfare programs may become more salient as middle-income countries develop and levels of formality increase.

References

Abadie, A., Chingos, M. M., West, M. R., 2018. Endogenous Stratification in Randomized Experiments. Review of Economics and Statistics 100 (4), 567–580.

Alzua, M. L., Cruces, G., Ripani, L., 2012. Welfare programs and labor supply in developing countries: experimental evidence from Latin America. Journal of Population Economics 26 (4), 1255–1284.

Amarante, V., Ferrando, M., Vigorito, A., 2013. Teenage School Attendance and Cash Transfers: An Impact Evaluation of PANES. Economia 14 (1), 61–102.

Amarante, V., Manacorda, M., Vigorito, A., Zerpa, M., 2011. Social Assistance and Labor Market Outcomes: Evidence from the Uruguayan PANES. Mimeo.

Amarante, V., Vigorito, A., 2011. Una propuesta para la evaluacion de impacto del programa Asignaciones Familiares-Plan de Equidad y la Tarjeta Alimentaria del MIDES. Tech. rep., FCEA-UDELAR, Montevideo.

Araujo, M. C., Bosch, M., Maldonado, R., Schady, N., 2017. The Effect of Welfare Payments on Work in a Middle-Income Country. IDB Working Paper Series 830.

Banerjee, A., Rema, H., Gabriel, K., Benjamin, O., 2015. Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs Worldwide. The World Bank Research Observer 32 (2), 155–184.

Bastagli, F., 2009. From social safety net to social policy?: the role of conditional cash transfers in welfare state development in Latin America. IPC-IG Working paper 60, The International Policy Centre for Inclusive Growth (IPC-IG), Brasil.

Bergolo, M., Dean, A., Perazzo, I., Vigorito, A., 2016. Evaluacion impacto del programa Asignaciones Familiares-Plan de Equidad y la Tarjeta Alimentaria del MIDES. Tech. rep., IECON-MIDES, Montevideo.

Bergolo, M., Galvan, E., Mar. 2018. Intra-household Behavioral Responses to Cash Transfer Programs. Evidence from a Regression Discontinuity Design. World Development 103, 100–118.

Bertrand, M., Mogstad, M., Mountjoy, J., Mar. 2019. Improving Educational Pathways to Social Mobility: Evidence from Norway s Reform 94. Working Paper 25679, National Bureau of Economic Research.

Best, M. C., Brockmeyer, A., Kleven, H. J., Spinnewijn, J., Waseem, M., Dec. 2015. Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan. Journal of Political Economy 123 (6), 1311–1355.

Bitler, M. P., Gelbach, J. B., Hoynes, H. W., 2006. What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. American Economic Review 96 (4), 988–1012.

Blank, R. M., 2000. Distinguished Lecture on Economics in Government: Fighting Poverty: Lessons from Recent U.S. History. Journal of Economic Perspectives 14 (2), 3–19.

Bosch, M., Manacorda, M., 2012. Social policies and labor market outcomes in Latin America and the Caribbean: a review of the existing evidence. CEP Occasional Papers 32, Center for Economic Performance.

Bosch, M., Schady, N., 2019. The effect of welfare payments on work: Regression discontinuity evidence from ecuador. Journal of Development Economics 139, 17 – 27.

Brewer, M., Saez, E., Shephard, A., 2008. Means-testing and tax rates on earnings. Prepared for the Report of a Commission on Reforming the Tax System for the 21st Century, Chared by Sir James Mirrlees, Institute for Fiscal Studies.

Calonico, S., Cattaneo, M. D., Titiunik, R., Nov. 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. Econometrica 82 (6), 2295–2326.

Carrillo, P., Pomeranz, D., Singhal, M., Apr. 2017. Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement. American Economic Journal: Applied Economics 9 (2), 144–164.

Cellini, S. R., Ferreira, F., Rothstein, J., Feb. 2010. The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. The Quarterly Journal of Economics 125 (1), 215–261.

Chan, M. K., Moffitt, R., 2018. Welfare Reform and the Labor Market. Annual Review of Economics 10 (1), 347–381.

Cheng, M.-Y., Fan, J., Marron, J. S., Aug. 1997. On automatic boundary corrections. The Annals of Statistics 25 (4), 1691–1708.

Eissa, N., Kleven, H., Claus Kreiner, 2006. Welfare Effects of Tax Reform and Labor Supply at the Intensive and Extensive Margins. In: Agell, J., Sorensen, P. B. (Eds.), Tax Policy and Labor Market Performance. MIT Press, Cambridge, pp. 147–185.

Eissa, N., Kleven, H. J., Kreiner, C. T., Apr. 2008. Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers. Journal of Public Economics 92 (3-4), 795–816.

Fields, G., 2009. Segmented Labor Market Models in Developing Countries. In: Kincaid, H., Ross, D. (Eds.), The Oxford handbook of philosophy of economics. Oxford University Press, Oxford, pp. 476–510.

Fiszbein, A., Schady, N. R., 2009. Conditional cash transfers: reducing present and future poverty. World Bank, Washington D.C.

Galiani, S., Weinschelbaum, F., Jul. 2012. Modeling Informality Formally: Households and Firms. Economic Inquiry 50 (3), 821–838.

Garganta, S., Gasparini, L., 2015. The impact of a social program on labor informality: The case of AUH in Argentina. Journal of Development Economics 115, 99–110.

Gasparini, L., Tornarolli, L., 2009. Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata. Desarrollo y Sociedad 63, 13–80.

Gerard, F., Gonzaga, G., 2020. Informal Labor and the Efficiency Cost of Social Programs: Evidence from the Brazilian Unemployment Insurance Program. American Economic Journal: Economic Policy (forthcoming).

Gerard, F., Rokkanen, M., Rothe, C., 2020. Bounds on treatment effects in regression discontinuity designs with a manipulated running variable. Quantitative Economics 11 (3), 839–870.

Grembi, V., Nannicini, T., Troiano, U., Jul. 2016. Do Fiscal Rules Matter? American Economic Journal: Applied Economics 8 (3), 1–30.

Hartley, R. P., Lamarche, C., 2018. Behavioral Responses and Welfare Reform: Evidence from a Randomized Experiment. Labour Economics 54, 135–151.

Hendren, N., 2016. The Policy Elasticity. Tax Policy and the Economy 30 (1), 51–89.

Hendren, N., Sprung-Keyser, B., 2020. A Unified Welfare Analysis of Government Policies. Working Paper.

Kleven, H. J., 2016. Bunching. Annual Review of Economics 8 (1), 435–464.

Kleven, H. J., Waseem, M., 2013. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. The Quarterly Journal of Economics 128 (2), 669–723.

Kline, P., Tartari, M., 2016. Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach. American Economic Review 106 (4), 972–1014.

Kostol, A. R., Mogstad, M., Feb. 2014. How Financial Incentives Induce Disability Insurance Recipients to Return to Work. American Economic Review 104 (2), 624–655.

La Porta, R., Shleifer, A., 2014. Informality and Development. Journal of Economic Perspectives 28 (3), 109–126.

Lee, D. S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. Journal of Economic Literature 48 (2), 281–355.

Lemieux, T., Milligan, K., 2008. Incentive effects of social assistance: A regression discontinuity approach. Journal of Econometrics 142 (2), 807–828.

Levy, S., 2008. Good intentions, bad outcomes. Social policy, informality and economic growth in Mexico. Brookings Institution Press, Washington.

Levy, S., Schady, N., 2013. Latin America's Social Policy Challenge: Education, Social Insurance, Redistribution. Journal of Economic Perspectives 27 (2), 193–218.

Maloney, W. F., 2004. Informality Revisited. World Development 32 (7), 1159–1178.

Manacorda, M., Miguel, E., Vigorito, A., Jul. 2011. Government Transfers and Political Support. American Economic Journal: Applied Economics 3 (3), 1–28.

Meghir, C., Narita, R., Robin, J.-M., 2015. Wages and Informality in Developing Countries. American Economic Review 105 (4), 1509–1546.

Meyer, B. D., Rosenbaum, D. T., 2001. Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers. The Quarterly Journal of Economics 116 (3), 1063–1114.

Millan, T. M., Barham, T., Macours, K., Maluccio, J. A., Stampini, M., Feb. 2019. Long-Term Impacts of Conditional Cash Transfers: Review of the Evidence. The World Bank Research Observer 34 (1), 119–159.

Moffitt, R. A., 2002. Welfare programs and labor supply, chapter 34 , Handbook of Public Economics. Vol. 4. Elsevier, pp. 2393–2430.

Naritomi, J., Sep. 2019. Consumers as Tax Auditors. American Economic Review 109 (9), 3031–3072.

Perry, G., William, M., Arias, O., Fajnzylber, P., Mason, A., Saavedra-Chanduvi, J., 2007. Informality: Exit and Exclusion. The World Bank, Washington D.C.

Pettersson-Lidbom, P., Apr. 2012. Does the size of the legislature affect the size of government? Evidence from two natural experiments. Journal of Public Economics 96 (3), 269–278.

Pomeranz, D., Aug. 2015. No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. American Economic Review 105 (8), 2539–2569.

Porter, M., Aug. 2003. The Economic Performance of Regions. Regional Studies 37 (6-7), 549–578.

Scholz, J., Levin, K., 2001. The Evolution of Income Support Policy in Recent Decades. In: Danziger, S. H., Haveman, R. H. (Eds.), Understanding Poverty. Harvard University Press and Russell Sage.

SEDLAC, 2017. Socio-Economic Database for Latin America and the Caribbean (CEDLAS and World Bank).
URL http://www.cedlas.econo.unlp.edu.ar/wp/estadisticas/sedlac/

Waseem, M., 2015. Taxes, Informality and Income Shifting: Evidence from a Recent Pakistani Tax Reform. Working Paper, University of Manchester.

APPENDIX: FOR ONLINE PUBLICATION ONLY

The Anatomy of Behavioral Responses to Social Assistance when Informal Employment is High

Marcelo Bergolo and Guillermo Cruces

September 2020

A.1 Description of Administrative and Survey Data

In this section, we provide additional information about the four datasets used in the paper: AFAM program records, Social Security Administration (SSA) records, a follow-up survey of program applicants, and the National Household Survey (ECH).

A.1.1 AFAM Program Records

The *AFAM administrative records* includes individual/household-level data from the AFAM program. These records include two datasets. The first, the baseline application records, correspond to a detailed questionnaire about the socioeconomic and demographic characteristics of all individuals in households that applied to the program. This rich baseline data contains information for both successful and unsuccessful applicant households, and it covers the period January 2008 to September 2010.

The detailed application form was designed to produce the proxy means score, and thus included a host of information on the households' living conditions. It was completed by SSA staff and a member of the applicant household. Its design was based on the typical questions found in household and labor force surveys. The individual and household characteristics elicited by this process include demographics, schooling, labor force participation and income, housing conditions and durable asset ownership, and region of residence, among others. The records also include the date of application and, most importantly, the exact value of the household's proxy means score computed by the authorities.

The second dataset, the AFAM participation records, corresponds to the participation history of the individual/household once they were enrolled in AFAM. This data contains information on households that successfully applied to the program between January 2008 and September 2010 – i.e. on a subsample of households for whom we have baseline data from the application records—and covers the period January 2008 to March 2012. We were granted access to the *AFAM administrative records* with unmasked unique national individual identifiers by the Social Development Ministry (SDM). An individual's national identification number corresponds to the unique number on his or her identification card ("*Cedula de Identidad*") which is issued at birth and renewed periodically for all citizens of Uruguay. It is uniquely linked to tax and social security records, and it is widely used as proof of identity in public offices and for private commercial services.

The variables used in the paper are the following:

- **document**: the individual national identifier number in Uruguay.
- **idnucleo**: a unique household identifier used by the MIDES and SSA institutions.
- **fecha**: the current month and year – "calendar time" (MM_YYYY). We use this variable to construct two additional variables, **mes** and **year**, which record the prevailing month and year, respectively. We also use this variable to construct a indicator variable for the calendar month, and we include a dummy for each of the twelve calendar months when we control for time-fixed effects in our regressions.
- **solicit_fecha_corr2**: the date of application to the AFAM program –"calendar time" (DD_MM_YYYY). We use this variable jointly with **fecha** to construct a "relative time" variable, **event_t**, which records

the number of months relative to the month and year the household applied to AFAM ("the event"). This variable ranges from -36 months to 24 months, while a value of 0 corresponds to the month in which the household applied to the program. In addition, we use the **event_t** variable to construct three variables: **post,** which is an indicator variable for the post-application period; **pre,** which is an indicator variable for the pre-application period; **and semester,** which records the number of semesters pre/post the application date.

- **fecha_nac_corr**: the individual's birthdate (DD_MM_YYYY). We use this variable jointly with **solicit_fecha_corr2** to construct a variable for the invidual's age at the time of the household application: **edad_ap**.
- **sexo**: an individual's gender (1 female, 0 otherwise)
- **jefe**: this variable records whether the individual is the head of the household (1 household head, 0 otherwise). We use this variable jointly with **sexo** to construct the variable **head_w** (1 female household head, 0 otherwise).
- **parentesco**: the individual's family relationship. We use this information to construct additional variables: **couples** (1 married/partnership, 0 otherwise), **single_w** (1 single female, 0 female in couple). We include an indicator variable for missing information when we control for marital status in our regressions.
- **region**: the department in Uruguay where the household is located. There are 19 departments in Uruguay. We include an indicator variable for each of the 19 departments when we control for region dummies in our regressions.
- **nivel_iecon**: the highest education level the individual has achieved . It is recorded as a categorical variable as follows: no education, pre-primary school, completed primary school, completed high school, completed technical school (UTU), completed tertiary education, no college, completed college. We transform this variable into a categorical "**level**" of education variable which ranges as follows: 1 "Completed Primary or less," 2 "Secondary or more". We include 3 indicator variables, one for each educational category plus one that captures missing information when we control for level of education dummies in our regressions.
- **child**: this variable indicates whether the individual is the son of the head of household. We use this variable jointly with **edad_ap** to construct the variable **children_17,** which records the number of children in the household below 18 years of age.
- **integrantes:** the number of household members.
- **panespart**: this variable records whether the individual has participated in the PANES program before applying to the AFAM program (1 PANES participation, 0 otherwise).
- **postulante**: the individual within the household who filled out the AFAM application form (1 Applicant, 0 otherwise).

- **employed_b**: this variable records whether the individual was employed at the moment of applying to the AFAM program (1 employed, 0 non-employed). We include an indicator variable for missing information when we control for employment at the baseline in our regressions.

- **t_afam**: the length in months between the household´s application date and the date the AFAM program was launched (01/2018).

- **indice_estd**: the household eligibility score, which was a function of the applicant households' characteristics. This variable is standardized relative to the eligibility threshold. Thus, it ranges from (-1;1) where positive values correspond to eligible households and negative values correspond to ineligible ones.

- **eligible**: whether the individual belongs to an applicant household eligible for the program (1 indice_estd > 0, 0 otherwise).

- **afam_corr**: whether the individual participated in the AFAM program during a given "calendar time" (1 if the individual participated in MM_YYYY , 0 otherwise). The period corresponds to January 2008 to March 2012.

- **acumula_trat_foto**: whether the individual was enrolled in the program at any given point in time since its implementation in January 2008 to March 2012 (1 enrolled at any time, 0 otherwise).

- **cant_men**: the number of children enrolled in primary school in a given AFAM beneficiary household.

- **cant_liceal**: the number of children enrolled in secondary school in a given AFAM beneficiary household.

- **benef_afam:** the value of the AFAM payment granted to the household. To construct this variable we use equation (1) in the main text, which describes the AFAM payment schedule, and plug in the values recorded in variables "cant_men" and "cant_liceal." The transfer baseline levels used in equation (1) are as follows:

Year	β (UYU)	δ (UYU)
2008	700	300
2009	764	327
2010	809	346
2011	865	370
2012	939	432

- **threshold**: the value of the income threshold for eligibility to participate in AFAM. While the income-eligibility threshold is computed as per capita household income, for some exercises in the paper we transform it to "per-person" so that it we can compare it with the individual earnings. This variable is called **threshold_imp**. In Section 2.3 of the main text, we explain in detail how we compute this variable.

- **flag**: the reason an applicant household was denied admission to the AFAM program. **flag** is recorded as an indicator variable with seven different possible values: "Does not pass the proxy means test"; "Does not pass the income test"; "Age of child above threshold"; "Did not present certificate of study"; "Did not present certificate of health checks"; "No children in the household"; "Other reasons".

A.1.2 SSA Records

The SSA records the monetary contributions made by employers and employees to social insurance services every month. A formal employee is one who is "registered" with the SSA as either an employee or self-employed worker in either the private or public sector, and thus is covered by the social insurance benefits provided by this institution. We have access to SSA records for program applicants for the period from January 2005 to December 2012. We construct a longitudinal database of registered employment histories by month that covers the three-year period before (January 2005 to December 2007) and after the period covered by the program application records (January 2008 to December 2012). Unfortunately, this dataset does not include information about hours worked per day (or days per month) or about earnings from registered work. We were granted access to the SSA employment history records of AFAM applicants.

The variables used in the paper are the following:

- **document**: the individual national identification number in Uruguay;
- **fingreso_bps**: a worker's date of registration with the SSA;
- **fegreso_bps**: the date of separation from the SSA;
- **fecha**: the current month and year – "calendar time" (MM_YYYY).

We use these variables to construct the following key variables:

- **cotiza:** this variable records whether the individual is registered with the SSA in a given month and year. Thus, this variable is the main outcome used in our paper and captures the employment status of the worker as a registered (formal) employee over time (1 if registered in MM_YYYY= t, 0 if not registered in MM_YYYY= t).
- **avg_cotiza_pre36**: this variables records the proportion of month that a worker was registered in SSA during the 36 months before application to AFAM.
- **groups**: the "propensity to be a registered employee" in terciles. As indicated in Section 6.5, this variable is constructed by using two different approaches. The first uses baseline registered employment as a proxy, defined as the number of months an individual was registered with SSA during the 36 months before the household applied to AFAM (i.e "**avg_cotiza_pre36**"). The second approach combines a broader set of background characteristics into a single predicted registered employment index as proxy for propensity to work formally (see Abadie et al. (2018) for a review of this method).

A.1.3 Follow-up Survey Data

An important limitation of administrative databases for the study of labor market outcomes in developing countries is that, by definition, these sources do not have any information about unregistered or informal employment. Individuals typically appear in these databases as working as registered employees for whom social insurance contributions and payroll taxes are paid, since the main purpose of these databases is to determine eligibility for social insurance benefits. In some cases, as in our data for Uruguay, individuals may also appear as beneficiaries of social assistance programs – typically child-related cash transfers or unemployment insurance. Individuals who do not appear in the database may thus be inactive, unemployed (and not receiving unemployment insurance payments), or working as unregistered or informal employees. While these are good data sources for determining registered employment status, they do not provide a complete picture of labor market outcomes because we cannot distinguish between inactivity, unemployment and unregistered work.

To complement the administrative data source, therefore, MIDES commissioned a group of researchers based at IECON-De La Republica University (UDELAR) to develop and implement a follow-up household survey specifically designed to study the effects of AFAM on household welfare and on individual labor market responses. In order to limit strategic responding, the surveyed households were informed that they were part of a research study being conducted by the UDELAR. They were further informed that, as research participants, their answers were not going to be used by the SSA and that the confidentiality of their individual information is legally protected (Statistical Secrecy Law No. 16.616). The survey was designed with the evaluation's RD identification strategy in mind. Thus, the survey's sampling frame was based on a stratified random sample of eligible and ineligible households that were close to the cutoff point according to AFAM baseline application records, in order to exploit the quasi-random variation generated by the eligibility rule. The optimal interval of the (standardized) predicted income score was $[-0.0426; 0.0727]$.

The follow-up survey was conducted from September 2011 to February 2013. Overall, 40% (1,441) of the stratified random sample of 3,565 households were interviewed, with a slightly higher proportion of non-respondents among ineligible households (44%) than among eligible households (39%). Despite the high-level of non-response, there is no significant correlation between non-response and eligibility status. For an analysis of this issue, see Bergolo and Galvan (2018).

The survey questionnaire was basically a shortened version of the Encuesta Continua de Hogares (the household survey carried out periodically by Uruguay's national statistical agency). It collected detailed information about household living conditions and individual labor market outcomes. More specifically, the survey data allows us to determine the mutually exclusive labor market outcomes at the time of the interview – i.e., registered and unregistered employment, unemployment and non-participation – for each individual in the sample.

The variables used in the paper are the following:

- **pobpcoac**: the specific employment status of an individual. It was generated by the IECON´s researchers using answers to a series of questions in the survey. Specifically, the variable identifies: individuals below 14 (legal age for working), employed, unemployed, not in the labor force, and working in a public-jobs program.

- **p68a**: this variable records whether an individual declared him- or herself to be a registered employee at the time of the interview. Specifically, the question in the follow-up survey is: "are you contributing to a retirement benefit through this job?" ("¿Aporta a alguna caja de jubilaciones por su trabajo actual?"). This is a standard criterion in the analysis of household surveys in Latin America — see, for instance, Gasparini and Tornarolli (2009) and Galiani and Weinschelbaum (2012).

We use the above variable to construct the following outcome variables:

- **nocupado**: this variable records whether the individual is non-employed at the time of the interview (1 non-employed, 0 otherwise).
- **informal**: this variable records whether the individual is employed as informal at the time of the interview (1 informal, 0 otherwise).
- **formal**: this variable records whether the individual is formally employed at the time of the interview (1 formal, 0 otherwise). Note that the three variables above are constructed as mutually exclusive employment statuses.
- **cotiza_enc**: this variable records whether the individual is employed as a registered employee at the time of the interview, but uses SSA records instead of the follow-up questionnaire.

In addition, we construct the variable:

- **campo**: the two time periods during which the survey was conducted, i.e. September 2011 to March 2012 and November 2012 to February 2013. We use this indicator as a dummy variable to control for our regression in Table 6.

A.1.4 National Household Survey (ECH)

We complement our empirical analysis with microdata from Uruguay's 2008-2012 National Household Survey (*Encuesta Continua de Hogares*, henceforth ECH). The ECH is a nationally representative household survey conducted according to international standards. It combines elements of living standard and labor force surveys, and is the main source for socioeconomic, labor, and demographic indicators in Uruguay. The microdata is made available to the public by the Instituto Nacional de Estadísticas (http://www.ine.gub.uy). We use ECH data to describe labor market patterns in Uruguay in Section 2.4 and for the imputation procedure for calculating participation tax rates in Section 8.

The variables used to describe labor market patterns are the following:

- **age:** the age of an individual at the time of the interview.
- **pobpcoac**: this variable records information as described in Section A.1.3
- **nocupado**: this variable records information as described in Section A.1.3
- **informal**: this variable records information as described in Section A.1.3

- **formal**: this variable records information as described in Section A.1.3. Note that the above three variables are constructed as mutually exclusive employment statuses.

- **tratado**: this variable records wheter the individual belongs to an AFAM beneficiary household. To construct this indicator, we use variables at the individual (household) level that indicate whether the individual (household) receives AFAM payments.

The variables used for the imputation procedure in Section 8 are the following:

- **sexo**: the gender of an individual (1 female, 0 otherwise)

- **jefe**: this variable records whether the individual is the head of the household (1 household head, 0 otherwise).

- **age**: the age of the individual at the moment of being interviewed.

- **couples**: this variable records whether the individual is married or partnered (1 married/partnership, 0 otherwise).

- **region**: the department (i.e., regional jurisdiction) in Uruguay where the household is located.

- **level**: the individual's level of education. This variable can take the following values: 1 "Less than Primary," 2 "Complete Primary," 3 "Secondary or more."

- **children_17**: the number of children under 18 years of age.

- **formal**: this variable records whether the individual is formally employed at the time of the interview (1 formal, 0 otherwise).

- **pareja form**: this variable records whether the individual's partner is formally employed (1 couple is a formal employee, 0 otherwise).

- **ing_form**: the monthly labor income of the individual from formal sources of earnings . Those sources of income include salary, overtime, commissions, incentives, tips, bonus, public transport tickets, payments for food and drinks out of the workplace, housing, and other types of payments. In cases of self-employment, it also includes withdrawals for household expenses, utility distributions, and withdrawals for one's own consumption. We define a new variable, **ing_form_nom,** which measures the individual's formal labor income in nominal terms.

- **ing_no_lab_corr**: the individual's monthly non-labor income . It includes all the monetary transfers received by the household, excluding the AFAM payment (i.e., social security benefits, pensions and retirement income, and transfers from other households).

- **benef_afam**: the value of the AFAM payment granted to the household. We construct this variable using the information from *AFAM administrative records* as described in Section A.1.1.

- **pesoano**: the weight attached to a given individual in the ECH sampling framework. This variable makes sure that the statistics that are based on ECH surveys are representative at the national level.

A.1.5 Summary Statistics

Table A.1.1presents summary statistics for selected socioeconomic characteristics of adult (18-57 years of age) individuals in all applicant households (column 1), for those from the *Main Sample* (column 2) and for individuals in the *Follow-Up Sample* (column 3). There seem to be no substantial systematic differences between the characteristics in columns 2 and 3. However, households and individuals in our two selected groups differ from those in the full population on some key characteristics. For instance, the individuals in our two samples had a higher level of education; 37.56% of the *Main Sample* and 27.09% of the *Follow-Up Sample* completed only primary school or less compared to 51.98% of the full population who did so. Also, the *Main Sample* and *Follow-Up Sample* households have fewer children (an average of 1.56 and 1.41, respectively, compared to 2.09 for the population), and fewer former beneficiaries of the PANES emergency cash transfer program, which targeted the extreme poor (19.41% and 9.45%, respectively compared to 42.46% for the population).

Table A.1.1: Comparison of Samples: Applicants to AFAM, Main and Follow-Up Samples

	Population (All applicant HHs)	Main Sample	Follow-up Sample
	(1)	(2)	(3)
Female applicants (%)	68.30	70.76	72.99
	(46.53)	(45.49)	(44.41)
Household head (%)	76.32	78.45	79.28
	(42.51)	(41.12)	(40.54)
Female head (within heads) (%)	72.90	76.37	80.73
	(44.45)	(42.48)	(39.45)
Age at application to AFAM	35.00	35.99	37.61
	(9.17)	(9.27)	(8.95)
Complete Primary or less (%)	51.98	37.56	27.09
	(49.96)	(48.43)	(44.45)
Secondary or more (%)	40.81	52.24	59.18
	(49.15)	(49.95)	(49.16)
Married/in couple (%)	49.20	44.59	44.04
	(49.99)	(49.71)	(49.65)
Single mother (within singles) (%)	85.73	87.25	90.36
	(34.97)	(33.35)	(29.52)
Number of children	2.09	1.56	1.41
	(1.35)	(0.94)	(0.83)
Enrolled in PANES (%)	42.46	19.41	9.45
	(49.43)	(39.55)	(29.25)
Montevideo (capital city) (%)	31.13	27.57	31.42
	(46.30)	(44.69)	(46.43)
Employed (%)	53.73	54.59	55.35
	(49.86)	(49.79)	(49.72)
Application Date (Months since 01/2008)	6.70	8.93	11.31
	(6.57)	(6.92)	(7.28)
Observations	240,146	87,834	2,403

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application and during the period, January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the SSA's administrative records for the period January 2005-December 2012 (see Section 3.2 for a detailed description of the data). "Population" in column (1) refers to individuals in these age categories in all households applying to AFAM during the period under study – i.e., the the entire population of AFAM applicants. The "*Main Sample*" in column (2) refers to the subset of individuals from households with eligibility scores in the range $[-0.257; +0.257]$ – i.e., individuals close to the eligibility score threshold, which we use for our main results. The "*Follow-Up Sample*" in column (3) corresponds to the subset of individuals interviewed for the program's follow-up survey during the period September 2011-February 2013, drawn from households with eligibility scores in the range $[-0.0426; +0.0727]$. All individual/household characteristics presented in this table are measured at the dtime of application, i.e. before an administrative decision on their enrollment in the program. See Appendix A.1 for a detailed definition of these variables.

A.2 Further Details about the AFAM Program and its Context

A.2.1 Social Insurance and Social Assistance Programs in Uruguay

Uruguay is a middle-income country in South America with a total population of approximately 3.3 million. In 2015, the annual GDP per capita reached around USD 15,000, making Uruguay the country with the second-highest per capita GDP in the region, surpassed only by Chile. The country has one of the oldest and most developed social protection systems in Latin America. This system follows a contributive, European Bismark-type model, where access to most welfare and social insurance programs is linked to registered employment and financed through payroll taxes and contributions from both employers and employees. Registered (or formal) employees are those working in firms that reported them to the Social Security Administration and for which they paid the relevant taxes and contributions. The overall payroll tax amounts to roughly 32% of taxable wages. Formal status makes these workers eligible for social insurance benefits such as health and unemployment insurance, sickness and disability benefits, maternity leave, family allowances, and old age pensions.

As in many middle-income countries, enforcement of labor market regulations is far from universal. There is widespread non-compliance with social insurance regulations and evasion of payroll taxes is quite pervasive. A substantial fraction of employees are not registered with the SSA and are thus not covered by social insurance benefits. This phenomenon is referred to as labor informality (Gasparini and Tornarolli, 2009).

The social assistance system in Uruguay during the period of our study (2008-2012) dates back to the mid-2000s. A severe economic crisis hit Uruguay in 2002-2003 and unregistered workers, who lacked access to the risk-mitigation mechanisms provided by the SSA, were hit especially hard by this crisis. The government responded by launching a series of reforms to the social protection system to expand the coverage of social assistance programs.[46] In particular, the government launched a temporary social assistance program called *Plan de Atención Nacional a la Emergencia Social (PANES)* in 2005. This program, which targeted the poorest 10% of households in Uruguay, provided a cash transfer conditional on a series of health and education requirements for children in beneficiary households.[47] This emergency program was replaced in January 2008 by a new system of family allowances (Law 18.227), the AFAM program, as part of a broader progressive reform of the tax and transfer system. AFAM targeted poor households—the bottom 20% of the income distribution—containing children or pregnant women, . It became the most important social assistance program in Uruguay both in terms of coverage and in the magnitude of the cash benefits provided.

AFAM was implemented as a means-tested conditional cash transfer (CCT) program targeted to households in precarious socioeconomic conditions. The program's monetary transfers were contingent on health checks (both for pregnant women and children) and school attendance for children in beneficiary households. At the beginning of 2008, AFAM covered 275,000 children. In 2014, the program reached nearly

[46] Those reforms were in line with a number of policies that many countries in Latin America implemented during the decade of 2000s to expand social protection and non-contributive programs by (Fiszbein and Schady, 2009).

[47] See Manacorda, Miguel, and Vigorito (2011) for more details on the goals, components and implementation of PANES.

370,000 children, which constituted about 42% of all children under the age of 18 in Uruguay. The budget for the cash transfer component of the program in 2013 was just over 0.35% of GDP. AFAM was among the largest programs of its type in Latin America in terms of its relative coverage and of its budget as a proportion of GDP. For instance, Brazil's Bolsa Familia reached almost 24% of the country's population, and had a budget of 0.4% of GDP in 2006, whereas Mexico's Progresa/Oportunidades covered 20% of the population with a budget of 0.4% of GDP in the same year (Bastagli, 2009).

A.2.2 The Uruguayan Labor Market and Some Key Patterns

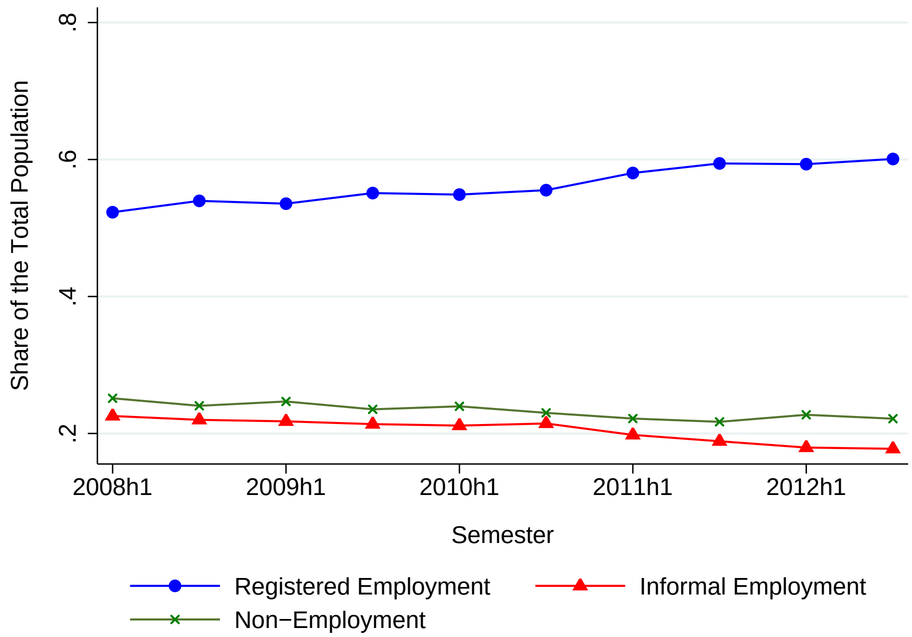
After the severe economic crisis that hit Uruguay in 2002-2003, ten years of extraordinary economic performance and growth of main labor market indicators followed. In 2015, 90.8% of men aged 18-60 and 74.6% of women worked or were looking for employment. The employment rate was about 76.4% while the unemployment rate was roughly 7.5%. Disaggregating by sex, the employment and unemployment rates were about 85.2% and 6.3% for men, respectively, and 67.9% and 9.0% for women, respectively. As in other developing economies, including all Latin American countries (LAC), the labor market in Uruguay featured both registered workers and a substantial share of informal employees. The unregistered employees represented about a quarter of the total number of salaried employees during 2000-2009 and declined to about 12% in 2015 (SEDLAC, 2017).

Next, we document key empirical patterns of the main labor employment outcome used in this study, i.e. registered employment, and the two complementary employment statuses, i.e., informal employment and non-employment (defined as being outside the labor force or unemployed) for the period of analysis (2008-2012). These three employment statuses are mutually exclusive and defined for the entire population of individuals aged 18-60. The statistics are computed using microdata from Uruguay's national household survey (*Encuesta Continua de Hogares*) described in the Appendix A.1. Figure A.2.1 depicts the evolution of those outcomes for the entire population.

The Figure shows that the share of the population who were registered employees increased steadily from 52.3% in 2007 to 60% in 2012, consistent with the extraordinary period of economic growth in Uruguay. This process of formalization seems to be associated with a continuous decrease in both non-employment and informal employment, whose shares declined from 25.1% and 22.6%, respectively, in 2008 to 17.8% and 22.2% in 2012, respectively.

Figure A.2.1: Labor Market Trends in Uruguay (2008-2012)

a) Share of population by labor market outcomes. Individuals aged 18-60



Notes: The data corresponds to Uruguay´s ECH for the years 2008-2012. See notes to Table (6) for a detailed definition of variables used in the figures. See also Appendix A.1 for a detailed definition of these variables.

A.2.3 Determination of the AFAM Benefit

As discussed in the main text, the level of the cash transfer varied according to the number of children in the household under the age of 18, and the number of children attending secondary school in the household.[48] The transfer was larger for those in secondary education to encourage older children to attend and finish school.

The total benefit granted to a household was computed as follows:

$$G_{AFAM} = \begin{cases} 0 & \text{if } (1-\theta)Y^F > T \\ \beta \times (Kids0to17)^{0.6} + \delta \times (HighSchoolKids)^{0.6} & \text{if } (1-\theta) \times Y^F \leq T \quad where\, \theta = 0.15 \end{cases} \tag{A.2.1}$$

where $Kids0to17$ represents the number of children below 18 years old, $HighSchooKids$ the number of children that attend secondary school, β and δ the transfer levels, Y^F the per-capita household gross formal income, and T the per-capita income test threshold. θ is a fixed deduction equal to 15% of the household's gross verifiable income that the authorities used to conduct the income test.

The cash transfer and the income test's threshold levels were adjusted periodically according to the evolution of the official Consumer Price Index. For instance, in 2011, the parameters in equation A.2.1 were set at β = UYU 865 (USD 47) and δ = UYU 371 (USD 20).[49] The average income transfer for a

[48] There was also an extra component for disabled children. These cases represent a very low fraction of our study sample.
[49] All figures in USD correspond to the June 2011 exchange rate.

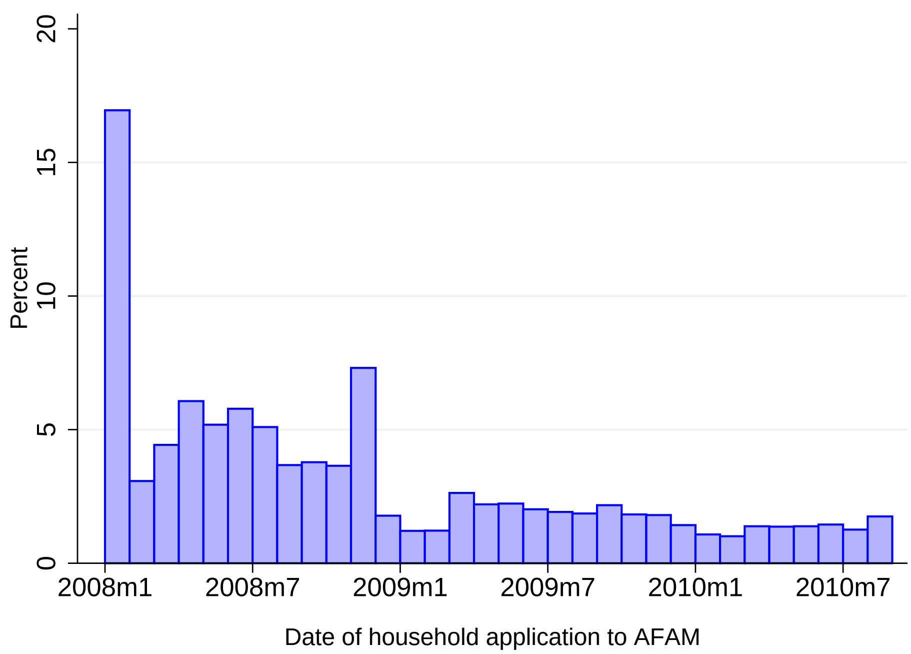
beneficiary household in 2011 was UYU 1,571 (USD 85), about 26% of the monthly national minimum wage. That same year, the income threshold was set at a monthly level of UYU 4,287 (around USD 231). These cash transfer levels were relatively generous and influenced beneficiary households' labor supply decisions: for 2008-2012, the transfer amounted on average to about 68% of the minimum wage for a household with two children, and about 48% of the poverty line.

All households in our study sample passed the income test at the time of application and were deemed eligible or ineligible according to the proxy means test. The latter was only conducted at the time of the original application and was not subject to manipulation. On the other hand, beneficiary households could adjust their labor supply and were subject to the income test every two months.

A.2.4 Timing of application to AFAM

As illustrated in Figure A.2.2, application to AFAM occurred at different points in time during our period of analysis, although most households applied in the first year following the program's launch (2008).

Figure A.2.2: Distribution of Households by AFAM Application Date



Notes: The sample corresponds to the population of applicant households with adults aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The outcome variable is date of application to AFAM defined as the month and year of the household application date.

The potentially non-random timing of applications might be related to potential outcomes. Table A.2.1, shows which household and individual characteristics from the baseline application records are related to the date of application. The dependent variable in this regression analysis is the month and year of a household's application date expressed as an index equal to 1 in January 2008. The evidence from this table indicates that applicants who were female, whio lived in households with a higher eligibility score , who were older, more educated, married, employed before the application process and had a higher number of children applied earlier to AFAM than did those who did not have these characteristics. Moreover, consistent with the implementation of the program, individuals who were beneficiaries of the previous PANES transfer program also applied to AFAM earlier in the process. While we find some statistically significant correlations between individual characteristics and the timing of application to AFAM, the effects are quantitatively unimportant and most of the variation remains unexplained. This result is consistent with the evidence that most of the applicants' observable characteristics are continuous around the eligibility cutoff, although we cannot rule out discontinuous changes in unobservable characteristics given by the timing of application.

Table A.2.1: Determinants of Household Application Date to AFAM

	Dep. Var.: Date of application
Income Score	-1.912***
	(0.286)
Female applicants	-0.204***
	(0.075)
Age at application to AFAM	-0.078***
	(0.003)
Household head	0.035
	(0.085)
Montevideo	0.254***
	(0.066)
Complete Primary or less	-0.395***
	(0.100)
Secondary or more	0.103
	(0.098)
Married	1.394***
	(0.078)
Married missing	3.180***
	(0.159)
Enrolled PANES	-5.237***
	(0.068)
Number of children	-0.771***
	(0.031)
Employed	-1.358***
	(0.064)
Employed missing	-6.080***
	(0.267)
Registered 36 months pre	2.953***
	(0.086)
Constant	15.380***
	(0.190)
R-squared	0.11
Observations	87,834

Notes: The sample corresponds to the entire population of applicant households with adults aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The dependent variable is date of application to AFAM, defined as the month and year of the household application date and expressed as an index equal to 1 in January 2008. All individual/household characteristics presented in this table are measured on the date of application, i.e. before the administrative decision regarding enrollment in the program. Huber-White robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. See Appendix A.1 for a detailed definition of these variables.

.

A.3 The Implied Elasticity of Participation in Registered Employment

A.3.1 Computation of the Elasticity of Participation in Registered Employment

Eissa et al. (2006) derive the participation elasticity in a context of labor force participation responses to taxes and benefits.[50] Waseem (2015) extends this setup to a context where formal employment is the relevant extensive margin of response. As in Waseem (2015), we incorporate the decision to be a registered employee in the basic model of labor supply by assuming that formal employment provides a discrete utility cost q^i to individuals. This utility cost includes direct costs associated with formal employment (e.g. regulations, transportation costs to large urban areas where formal firms tend to agglomerate, etc.), as well as indirect costs, such as other social benefits that the worker would have to forego as a registered employee (Galiani and Weinschelbaum, 2012). The utility maximization process is described in two stages. First, an agent chooses her optimal hours (or earnings) conditional on participation in registered employment. Second, the agent must decide whether to participate or not. Let w represent earnings from registered employment, let the function $T(y)$ be the tax and transfer schedule, and let -$T(0)$ represent the program's benefit (e.g., the AFAM cash transfer) which, for simplicity's sake, we assume is received only by those who are not formally employed. An individual participates in formal employment if, and only if, the utility of doing so, $u^i\left(y - T(y)\right) - q^i$, exceeds the utility from non-participation, which we assume to be $u_0^i\left(-T(0)\right)$.[51] This leads to the following condition for participation in registered employment:

$$q^i \leq u^i\left(y - T(y)\right) - u_0^i\left(-T(0)\right) \equiv \bar{q}_i \tag{A.3.1}$$

This expression defines an upper bound $\bar{q}_i$ on the discrete utility gain from participation in registered employment. The size of $\bar{q}_i$ reflects the utility gain from participating in registered employment accounting for taxes and transfers. That is, individuals with fixed cost q^i below $\bar{q}_i$ decide to participate in the labor market as formal employees. If q^i is distributed among the population according to a smooth function $R(q)$, the fraction of individuals in registered employment is given by $R(\bar{q})$. The elasticity of participation in registered employment captures movements into and out of formal employment as a consequence of the tax and transfer schedule:

$$\epsilon_R = \frac{1 - ptr}{R(q)} \cdot \frac{\triangle R(q)}{\triangle(1 - ptr)} \tag{A.3.2}$$

This elasticity is defined as the percentage change in participation in registered employment, $R(q)$, induced by a one percentage point change in the average net-of-participation in registered employment tax rate, $1 - ptr$.[52] It should be noted that the ptr corresponds to the effective average tax rate, which includes

[50] In the context of AFAM, the potential beneficiaries may decide to work either as formal or informal employees, where we assume full compliance and non-compliance, respectively. In that case, for a given level of enforcement, the decision to work or not to work can be replaced without loss of validity by the decision to work and report full earnings, or to work and not report earnings (Brewer et al., 2008).

[51] This fixed cost is only incurred if the individual participates in registered employment. To simplify, we assume that it is additively separable in utility.

[52] For small changes in taxes and transfers, $\triangle(1 - ptr) = d(1 - ptr)$ and $\triangle R(q) = dR(q)$, so that the elasticity is given by

the loss of the program's benefit following entry into registered employment (when the resulting earnings exceed the program's income testing threshold).

We compute the aggregate elasticity of participation in registered employment, ϵ_R, in the context of our empirical analysis by closely following the procedure developed by Kostol and Mogstad (2014) in their analysis of the effects of disability insurance on labor supply in Norway.

By using a change of notation, we can express ϵ_R in equation A.3.2 as:

$$\epsilon_R = \frac{1 - ptr_{ineligible}}{R_{ineligible}} \cdot \frac{\triangle R}{\triangle(1 - ptr)} \tag{A.3.3}$$

where (1-$ptr_{ineligible}$) denotes the mean net-of-participation tax rate on registered employment for the group ineligible for AFAM, $\triangle(1 - ptr)$ is the difference in the net-of-participation tax rate between eligible and ineligible applicants, $R_{ineligible}$ expresses the registered employment rate for AFAM-ineligible individuals, and $\triangle R$ denotes the difference in registered employment between eligible and ineligible individuals.

A.3.2 Estimating the Change in Participation Tax Rate in Registered Employment

The main challenge to compute ϵ_R is deciding how to measure the $ptr_{ineligible}$. The ptr on registered employment is defined as 1 minus the financial gain from working as a formal employee as a proportion of gross registered earnings. It measures how the tax and benefit system affects the financial reward of working as a formal employee. Specifically, we define a measure of the ptr at a given earnings level as:

$$ptr_k = 1 - \frac{I_k - I_0}{y_k} \tag{A.3.4}$$

where I_k denotes disposable income with a level of registered gross earnings y_k, and I_0 corresponds to the amount of disposable income if the individual is not working formally – i.e., the earnings derived from the two employment states that we can observe in the SSA data – (adding the AFAM cash transfer when relevant). Importantly, non-formal employment includes both non-employed and informally employed status, because in the SSA records we do not know which status corresponds to an individual coded as unregistered. To compute disposable income, we need three inputs: the individual registered earnings y_k, the net-of-tax transfers $T(y)$ – such as $I_k = y_k - T(y)$ –and the level of earnings y_0 associated with informal employment – note that $I_0 = y_0 + T(0)$. We make this computation by following two steps.

First, since we do not have information on earnings in our data, we obtain earnings from an imputation based on Uruguay's ECH for the period 2008-2012 (i.e. the same period for which we have available SSA records on registered employment). Specifically, based on a sample of registered/unregistered employed individuals we regress registered earnings on a set of standard covariates.[53] We use the estimated coefficients to predict registered/unregistered employment earnings for all the individuals in our (AFAM) *Main Sample*, $\hat{y}$. In Section A.3.3, we describe in detail the procedure for imputing earnings.

$\epsilon_R = \frac{1-ptr}{R(q)} \cdot \frac{dR(q)}{d(1-ptr)}$

[53]We use the standard covariates listed in the notes to Table 1, with the exception of the variables for enrollment in the PANES program and employment at baseline, which are not available in the household survey.

The second step is to obtain the individual net taxes and transfers $T(\widehat{y})$ based on Uruguay's tax levels and the AFAM transfer schedule according to the characteristics of each individual and household. The tax schedule for low-income registered workers in Uruguay is fairly simple – their earnings are only subject to a payroll tax (in fact, social security contributions) amounting to 18.09% for individuals with very low incomes and 21.13% for the next range of gross earnings, paid by the employer.[54] The income tax minimum threshold is well above the earnings of low-income workers, so we assume that AFAM's target population is not liable for personal income taxes. We compute the exact AFAM benefit for each individual according to the formula given by equation (1) (see main text of the paper). We obtained the information required to compute the exact benefit from the program's administrative database. The effective *ptr* also accounts for the loss of the AFAM benefit when earnings from registered earnings exceed the program's eligibility threshold.

Based on these inputs – imputed earnings $\widehat{y}$, taxes and transfers $T(\widehat{y})$ – we estimate the disposable income and compute the *ptr* for every individual in our database. Finally, the effective *ptr* at a given level of imputed earnings from registered employment $\widehat{y}_k$ is given by 1 minus the difference in disposable income at that level of income $\widehat{I}_k$ and at zero earnings from registered employment, $\widehat{I}(0)$:

$$ptr_k = 1 - \frac{\widehat{I}_k - \widehat{I}(0)}{\widehat{y}_k} \tag{A.3.5}$$

We then define the difference in the *ptr* as the weighted difference between the *ptr* rates for eligible and ineligible individuals:

$$\triangle ptr = \sum \left[E\left(ptr_k|eligible\right) - E\left(ptr_k|ineligible\right)\right] p_k \tag{A.3.6}$$

where we define $k = 20$ bins of earnings from registered earnings $\widehat{y}_k$ by dk increments, and we compute the ptr_k at the average income for each bin. $E\left(PTR_k|eligible\right)$ is the average *ptr* for eligible individuals and $E\left(PTR_k|ineligible\right)$ is the average *ptr* for ineligible individuals. The weights p_k reflect the density of the income distribution function for the ineligible individuals, and they are computed as:

$$p_k = \frac{Pr(k \leq \widehat{y}_k < k + dk \mid ineligible)}{\sum_{k>0} Pr(k \leq \widehat{y}_k < k + dk \mid ineligible)} \tag{A.3.7}$$

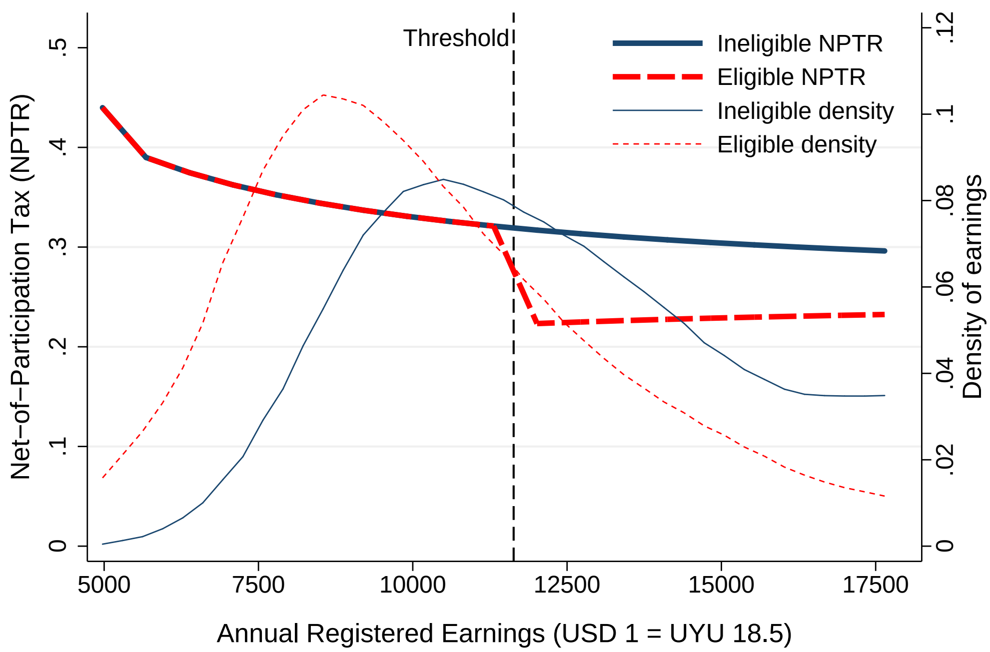
This implies that $\triangle ptr$ is the weighted sum of the differences between registered employment participation tax rates for eligible and ineligible individuals, with weights given by the (conditional) density of the registered earnings of ineligible individuals who work as registered employees. In the context of our analysis, the payroll tax does not vary substantially between individuals, so the main source of variation in incentives to participate in registered employment is the variation given by the gain/loss of the AFAM transfer above/below the program's registered earnings eligibility threshold.

For illustration purposes, Figure A.3.1 depicts the net-of-tax *ptr,* i.e. 1-*ptr*, and the density of earnings from registered employment by eligibility status for the year 2011. Some important points emerge from this figure. First, average earnings from registered employment for individuals eligible for AFAM are lower

[54] Employers are liable for payroll taxes on earnings in the form of social security contributions of up to 13% of gross earnings. We implicitly assume that workers and employers each pay the statutory rate assigned by the tax authorities. If registered employees bear the full cost of the sum of employee and employer payroll taxes, the effective tax rate would be higher than what we use in our computation.

than the same type of earnings for ineligible individuals. Second, most eligible individuals would remain eligible even when working in registered employment – i.e., their earnings from registered employment would be below the income-eligibility threshold. Third, the tax and transfer system and, more specifically, the AFAM eligibility threshold, creates a discontinuity and a wedge which declines as earnings grow, but remains substantial for relevant levels of earnings for the population under study. This AFAM-induced notch lowers the financial rewards associated with working as a registered employee by about 10%, and it is larger for formal earnings close to the income-eligibility threshold.

Figure A.3.1: Distribution of Earnings and Participation Tax Rates by Eligibility Status – Estimation Based on 2011 Data



Notes: This figure is based on Figure 5 in Kostol and Mogstad (2014). It plots the density of monthly earnings (right axis) and the net-of-participation tax rate (left axis) for earnings above the income-eligibility threshold for ("average") eligible and ineligible individuals. The horizontal axis corresponds to monthly earnings in UYU. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application and during the period, January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the SSA's administrative records for the period January 2005-December 2012, and information from Uruguay's ECH. The participation tax rate (ptr) is computed following the equation (A.3.5) and the simulated tax and transfer schedule that a low-income individual would face if she was, or was not,eligible for the AFAM, respectively. Earnings are constructed using the ECH data and the imputation procedure described in Section A.3.3. The dashed vertical line represents the income threshold for an "average" household in our sample. To transform the income-eligibility threshold measured as household per-capita income to an equivalent "per-person income-threshold," we follow the same procedure described in Section 2.3.

Our simple calculations of the elasticity of participation in registered employment suffer from a number of issues, which warrant cautious interpretation of the results. First, the elasticity we estimate is not a structural parameter depending solely on individual preferences. It depends on the specific features of the Uruguayan tax and benefit system and its enforcement rules. Moreover, we abstract from specific characteristics of AFAM eligibility design for our calculations. For instance, the income-eligibility condition is not conducted at the individual level but also depends on the formal employment status of other individuals in the household. Second, for the above reason, and because the AFAM-induced response that we are estimating probably is capturing income and substitution effects, the calculated elasticity should not be interpreted as a sufficient statistic to quantify the efficiency costs of the program, at least if one uses traditional measures as the marginal excess burden. Finally, to construct the elasticity, we consider the AFAM in isolation, i.e., we abstract from other policy instruments that interact with the AFAM program and that probably also affect the financial incentives the eligible population face.

A.3.3 Imputation of Individual Earnings from Registered Employment

A key input to compute the participation tax rate is the individual level of earnings when formally employed and when informally employed. Information on earnings is not available in our data. Instead, we rely on a standard earnings' imputation procedure based on data from Uruguay's ECH during the period 2008-2012 . This procedure consists of a series of steps.

First, we estimate a probit model for the likelihood that an individual participates in (and has earnings from) registered/unregistered employment in a baseline year:

$$P((U)RE_i/=1) = \Phi(Z_i\gamma) \tag{A.3.8}$$

where $(U)PRE$ is participation in (un)registered employment – unregistered status indicates the individual is either non-employed or informally employed – for individual i, and variable Z denotes a set of controls X, plus an additional variable indicating whether at least one adult other than i in the household is a registered employee.[55]

Second, we estimate a simple earnings equation, specified as follows:

$$log(YRE_i) = X_i\alpha + \widehat{P_i^{RE}}\beta + \varepsilon_i \tag{A.3.9}$$

$$log(YURE_i) = X_i\alpha + \widehat{P_i^{URE}}\beta + \varepsilon_i \tag{A.3.10}$$

where YRE_i and $YURE_i$ are registered income and "unregistered" income (i.e., income observed in the survey when an individual is not formally employed) for individual i, respectively, and X denotes the same controls as in equation A.3.8. To account for the self-selection of individuals in registered employment, we include propensity scores, $P^{RE}$and P^{URE}, which denote an individual's estimated probability of being formally and informally employed, respectively, derived from equation A.3.8, as an additional control in the earnings equation. Note that we identify the selection effect based on the registered/unregistered status of other adults belonging to an individual's household.

The final step implies using the estimated coefficients of the earnings equations for each year to predict the earnings from registered employment among individuals in our administrative AFAM database. We impute potential earnings for individuals who are working as registered employees and for those who are unregistered according to the SSA records. Finally, we use the predicted earnings to simulate the participation tax rates (with and without the AFAM program) as described in the previous section.

[55] X corresponds to a set of demographic and household characteristics, including age, head of the household status, marital status, education level (in three categories), the number children in the household under 18, residency in the capital city (Montevideo), and an indicator for participation in AFAM.

A.4 Additional Discussions and Results

A.4.1 The Role of Program Conditionalities

Beyond these standard economic theory arguments based on the program's rules and benefits, the program's conditionalities might also induce changes in the labor supply of adults. On the one hand, the requirement that children attend school might free up time that adults in the household previously spent on childcare. On the other hand, if conditionalities are effective in curbing child labor, the net effect of transfers on households' incomes is reduced, which might mitigate the program's potential disincentive for adults to seek paid employment (Alzua et al., 2012).

The combination of these channels in addition to the "financial channel" implies that the overall effect of AFAM on adults' labor force participation is ambiguous from a theoretical point of view. However, we expect the reduction of employment that results from the financial disincentive to be of first-order importance relative to the more ambiguous incentives introduced by the conditionalities. First, as discussed above, it is not clear to what extent the government really enforced these conditionalities, at least during the period that we study. In fact, evidence from our follow-up survey suggests that about 40% of the beneficiaries were unaware of conditionalities attached to the program (Bergolo et al., 2016). Manacorda, Miguel, and Vigorito (2011) also note that the conditionalities were *de facto* not enforced in the case of PANES, the program that preceded AFAM, because of the lack of coordination between public institutions. Second, school attendance is nearly universal for primary school children in Uruguay, and thus the child labor argument would only apply to teenagers and not to all children in the household. In fact, Amarante, Ferrando, and Vigorito (2013) did not find any evidence that PANES affected school attendance or child labor for children aged 14 to 17.

A.4.2 Applicant Rejection and Disqualification from AFAM Participants

A further consideration concerning the expected effects of the program is related to the process of disqualification from AFAM. On average, about 15% of households were removed from the program each year in the period under study, and 57% of those disqualifications were due to households failing the income test. About three-quarters of those disqualified for failing the income test regained eligibility for the AFAM benefit later on (row 4, Appendix Table A.4.2).

Adults in households disqualified from the program for failing the income test (i.e., because their earnings from formal employment surpassed the threshold, probably because adults in the household entered registered employment in the context of the growing economy) had an incentive to move to informal employment (with non-verifiable earnings) or to exit the labor force altogether and, hence, reduce their formal labor earnings to regain eligibility. The patterns of exit and re-entry to AFAM suggests that this effect was relevant in our context.

Table A.4.1: Reasons for Rejection of AFAM Applicants (2008-2010)

	Frequency (%)
Do not pass the proxy means test	37.20
Do not pass the income test	13.83
Age of child above threshold	1.85
Do not present certificate of study	3.31
Do no present certificate of health controls	3.30
No children in the household	9.86
Other reasons	30.64
Num. of Households	24,684

Notes: The sample corresponds to the entire population of applicant households with adults aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The table shows the distribution of the major reasons for rejection of households who applied to the AFAM program.

Table A.4.2: AFAM Disqualification Statistics (2009-2012)

	2009	2010	2011	2012	Average
AFAM households removed (%)	10.81	12.54	17.05	17.81	14.55
Income test removal w.r.t total removed (%)	43.72	60.16	62.40	60.13	56.60
Num. of times removed (Avg.)	1.169	1.265	1.226	1.148	1.202
New enrollment after an income test removal (%)	72.11	70.36	78.44	73.83	73.69

Notes: The dataset corresponds to the *AFAM administrative records* (January 2008-September 2012). The estimates shown in this table were provided by AFAM program authorities; we do not have the microdata necessary to replicate it.

A.4.3 Balance and Robustness Checks for RD Design-Main Sample

As discussed in the body of the text, we implement a variation of the basic regression discontinuity (RD) design that leverages the time dimension of our data (DD-RD). In the standard RD setting, we would exploit the sharp discontinuity in the AFAM assignment rule of the program to identify its effects on applicants' behavioral responses. We can estimate a regression model within a narrow window around the AFAM eligibility threshold specified as follows:

$$Y_{it} = \alpha + \beta ELEG_i + \delta ELEG_i \times f(score_i) + f(score_i) + \lambda_t + \epsilon_{it} \tag{A.4.1}$$

where Y_i is the outcome of interest for individual i at time t; $ELEG_i$ is a dummy equal to 1 if the individual belongs to an applicant household eligible for the program (i.e., if $score_i > 0$), and zero otherwise; *score* is the value of the eligibility score, which in the RD literature is standardized relative to the eligibility threshold (c); $f(score_i)$ is an (unknown) functional form of the "assignment" variable *score*; and λ_t represents time fixed-effects. The parameter β captures the causal effect of AFAM on the outcome of interest in the RD design.

Identification in the RD research design requires that $E[Y_i(1) \mid score]$ and $E[Y_i(0) \mid score]$ be continuous functions at the eligibility threshold c.[56] Since *ELEG* is a discontinuous function of the eligibility score, and the control function $f(.)$ in equation A.4.1 is, by assumption, continuous at c, the RD coefficient β is identified if the continuity condition is met. This continuity assumption would be violated if individuals were able to manipulate the program's eligibility process.

Table A.4.3 presents the average of selected socioeconomic characteristics at baseline for ineligible individuals close to the cutoff in the *Main Sample* (column 1) and in the *Follow-Up Sample* (column 4). Column (2) reports the RD estimates for the difference in mean value between eligible and ineligible individuals (the β coefficient in equation A.4.1) for each characteristic in the *Main Sample* (with the optimal bandwidth reported in column 3), while column (5) reports the estimates for the *Follow-Up Sample*. Most of the RD estimates of the differences in socioeconomic characteristics at baseline are not statistically significant at the standard levels. Moreover, most of the significant discontinuities—in age, fraction of individuals with secondary education or more, number of children, and date of application—are economically small (). However, there are substantial and significant differences in enrollment in PANES in both datasets.[57] Most importantly, we find some level of imbalance at baseline in our main outcome of interest, registered employment, which is 37.33% for the ineligible individuals and lower by about 3.8 percentage points for eligible individuals (p-value of the difference = 0.040).

Despite the sharp discontinuity in eligibility depicted in Figure 2, the imbalance in the pre-application period main outcome, and in some other baseline characteristics in Table A.4.3 (depicted in RD form in Figures A.5.1, A.5.2 and A.5.3) may signal manipulation of the running variable, which would compromise identification of causal effects in the context of an RD research design based on Figure 3.

A standard prediction consistent with a non-manipulated regression discontinuity setting is that the distribution of the assignment variable itself should be continuous at the eligibility threshold when potential beneficiaries are unable to manipulate the underlying score. Panel A in Appendix Figure A.4.1 presents the distribution of the assignment variable, the standardized eligibility score for the main samples. There

[56] In terms of the Rubin causal model, $Y_i(1)$ and $Y_i(0)$ denote the potential outcomes for eligible and ineligible individuals respectively.

[57] The large difference for being located in Montevideo (Uruguay's capital) for the *Follow-Up Sample* is a result of the survey's sampling design. There were quotas of eligible and ineligible individuals, and to boost the sample to meet the eligible interviewees quota the fieldwork was carried out in Montevideo for cost reasons. These observations were balanced in other relevant dimensions.

Table A.4.3: Regression Discontinuity Differences in Baseline Characteristics, Eligible and Ineligible Individuals, Main and Follow-Up Samples

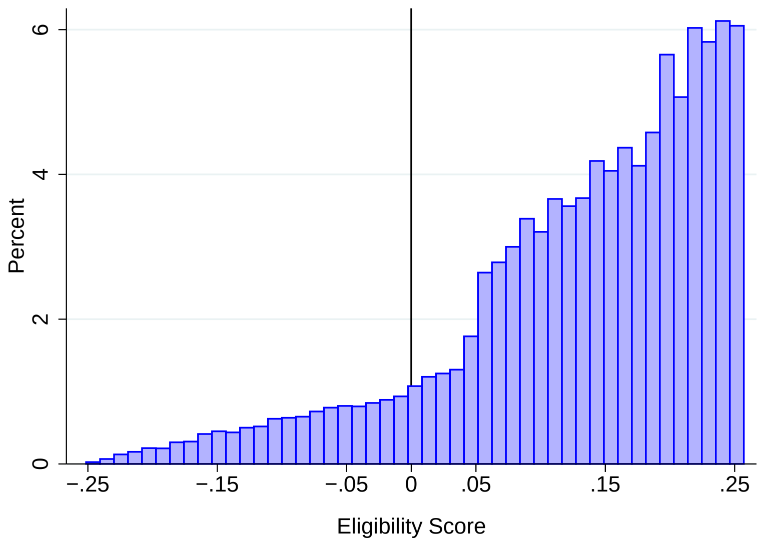
	Main Sample			Follow-up Sample	
	Comparison Mean	Estimates	BW	Comparison Mean	Estimates
	(1)	(2)	(3)	(4)	(5)
Female applicants (%)	70.45	-0.737 (1.566) [0.463]	0.056	72.96	-1.101 (3.146) [0.381]
Household head (%)	76.06	-0.146 (1.318) [0.899]	0.068	79.25	0.337 (2.999) [0.315]
Female head (within heads) (%)	76.95	-0.237 (2.310) [0.703]	0.048	81.42	2.093 (4.244) [0.098]
Age at application to AFAM	38.50	-1.328 (0.441) [0.005]	0.058	38.85	-1.494 (0.868) [0.564]
Complete Primary or less (%)	27.41	1.343 (2.169) [0.397]	0.058	26.10	1.008 (4.186) [0.761]
Secondary or more (%)	60.02	-4.836 (2.546) [0.025]	0.050	60.37	-5.366 (4.738) [0.685]
Married/in couple (%)	47.72	1.175 (2.671) [0.547]	0.063	44.20	4.373 (5.735) [0.997]
Single mother (within singles) (%)	87.73	0.263 (1.876) [0.986]	0.062	90.35	3.642 (4.108) [0.038]
Number of children	1.26	0.161 (0.041) [0.000]	0.070	1.27	0.145 (0.094) [0.333]
Enrolled in PANES (%)	4.38	3.395 (1.098) [0.001]	0.080	4.42	8.371 (2.724) [0.000]
Montevideo (capital city) (%)	34.00	-3.636 (2.326) [0.099]	0.068	36.68	-33.096 (5.274) [0.001]
Employed (%)	57.43	-1.037 (2.041) [0.825]	0.063	57.30	-3.149 (4.336) [0.249]
Application Date (Months since 01/2008)	10.00	2.460 (0.368) [0.000]	0.064	10.16	3.623 (0.821) [0.001]
Registered 36 months pre (p.p.)	37.33	-3.781 (1.705) [0.040]	0.062		
Observations		17,404			2,403

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application and during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). The *Main Sample* refers to the subset of individuals from households within the eligibility score range $[-0.257; +0.257]$. The *Follow-Up Sample* corresponds to the subset of individuals interviewed for the program's follow-up survey during the period September 201-February 2013, drawn from households within the eligibility score range $[-0.0426; +0.0727]$. All individual/household characteristics included in this table refer to the household's application date, with the exception of the last two rows in the main panel. "Registered 36 months pre" refers to each individual's average registered employment rate for the 36 months before application to the AFAM program. See Appendix A.1 for a detailed definition of the variables shown in this table. Column (1) presents the average characteristics for individuals from ineligible households close to the cutoff in the *Main Sample*. Column (2) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) according to Calonico et al. (2014) from a RD specification described in Equation (A.4.1), with the respective characteristic as the dependent variable. Estimates from local linear regressions use a triangular kernel. "BW" in Column (3) reports optimal bandwidths following the Calonico et al. (2014) procedure. Columns (4) and (5), in turn, report the same estimates as columns (1) and (2), but for those individuals in the *Follow-Up Sample*.

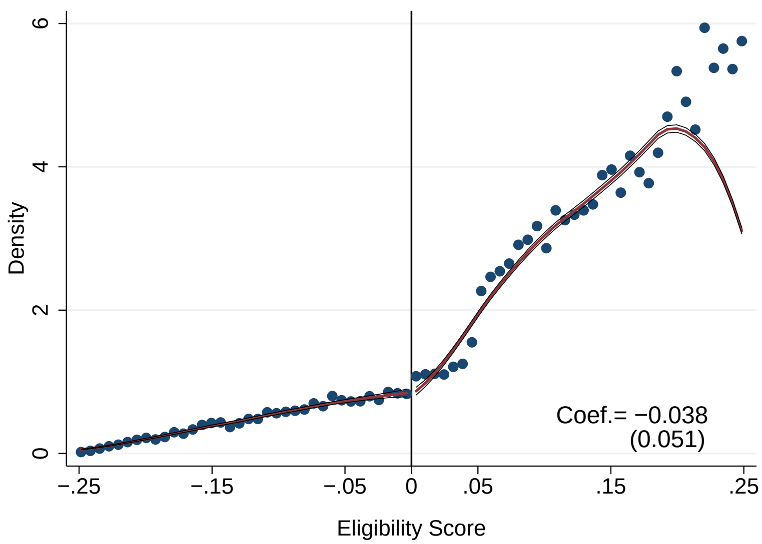
does not seem to be a major discontinuity in the fraction of applicants around the eligibility threshold, as manipulation of the eligibility score would imply. Panel B in that figure depicts the estimates corresponding to the McCrary test – i.e., the density of the eligibility score and a smoothed density estimator based on a local linear regression on both sides of the threshold. The formal test is implemented as a Wald test of the null hypothesis that there is no discontinuity in the density of the standardized eligibility score at the eligibility cutoff. The estimated discontinuity in the density is -0.038 – a 4.52% change relative to the density at the left of the cutoff point – with a standard error of 0.051. We cannot reject the null hypothesis of no discontinuity (p-value = 0.454).

Figure A.4.1: Distribution of the Assignment Variable and the McCrary Test

A. Eligibility Score (Assignment Variable)



B. McCrary's Density Test

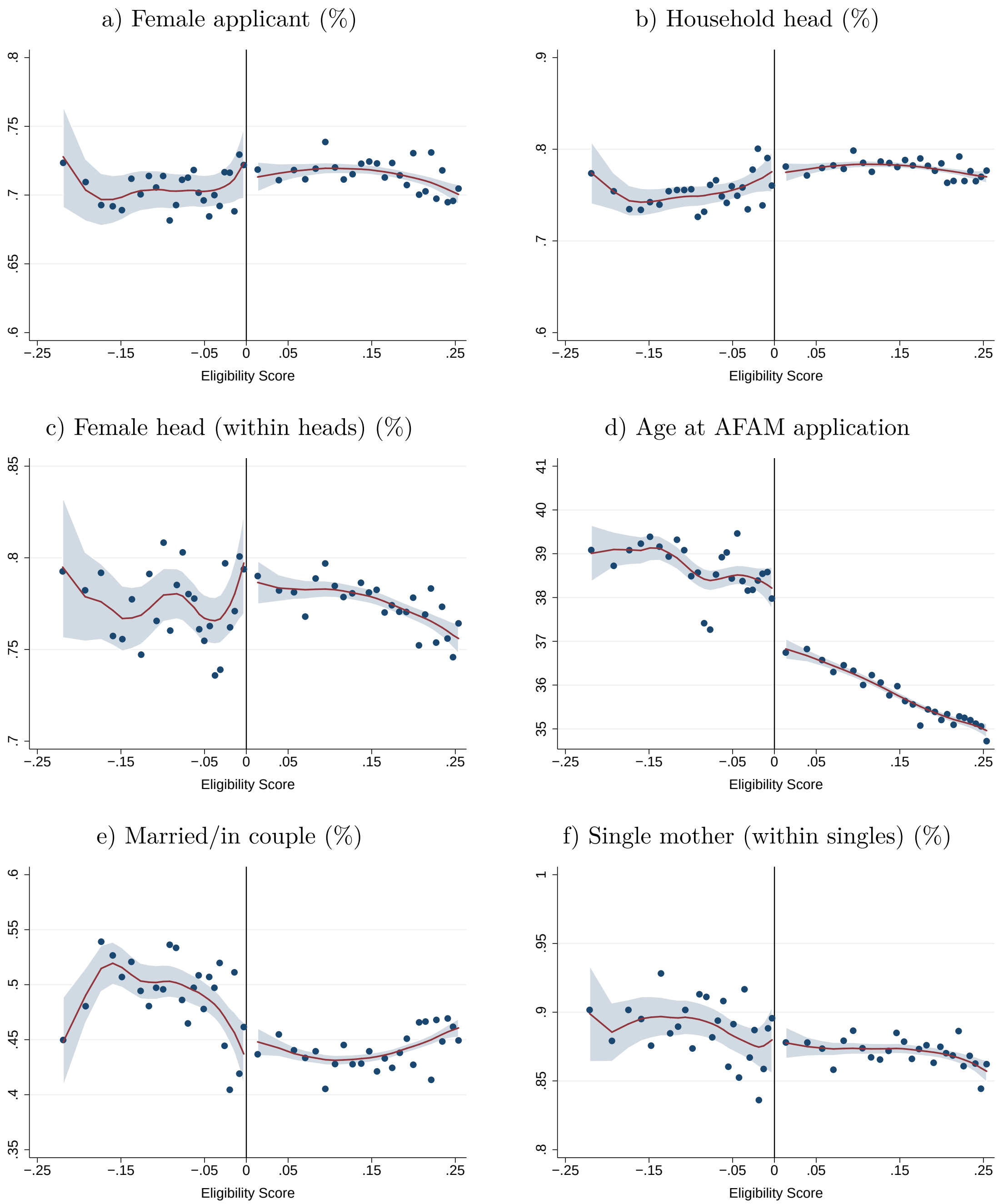


Notes: The sample corresponds to the population of applicant households with adults aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). Panel A shows the histogram of the eligibility score distribution in bins with a width equal to one percentage point. Panel B plots the eligibility score density in bins with a width of one-half of a percentage point. The solid red line plots predicted values from a local linear regression (with a width of one-half of a percentage point) with separate score trends estimated for either side of the eligibility threshold. The dashed lines show 95% confidence intervals. The bandwidth is optimally chosen and we use a rectangular kernel.

A.5 Additional Results: Figures

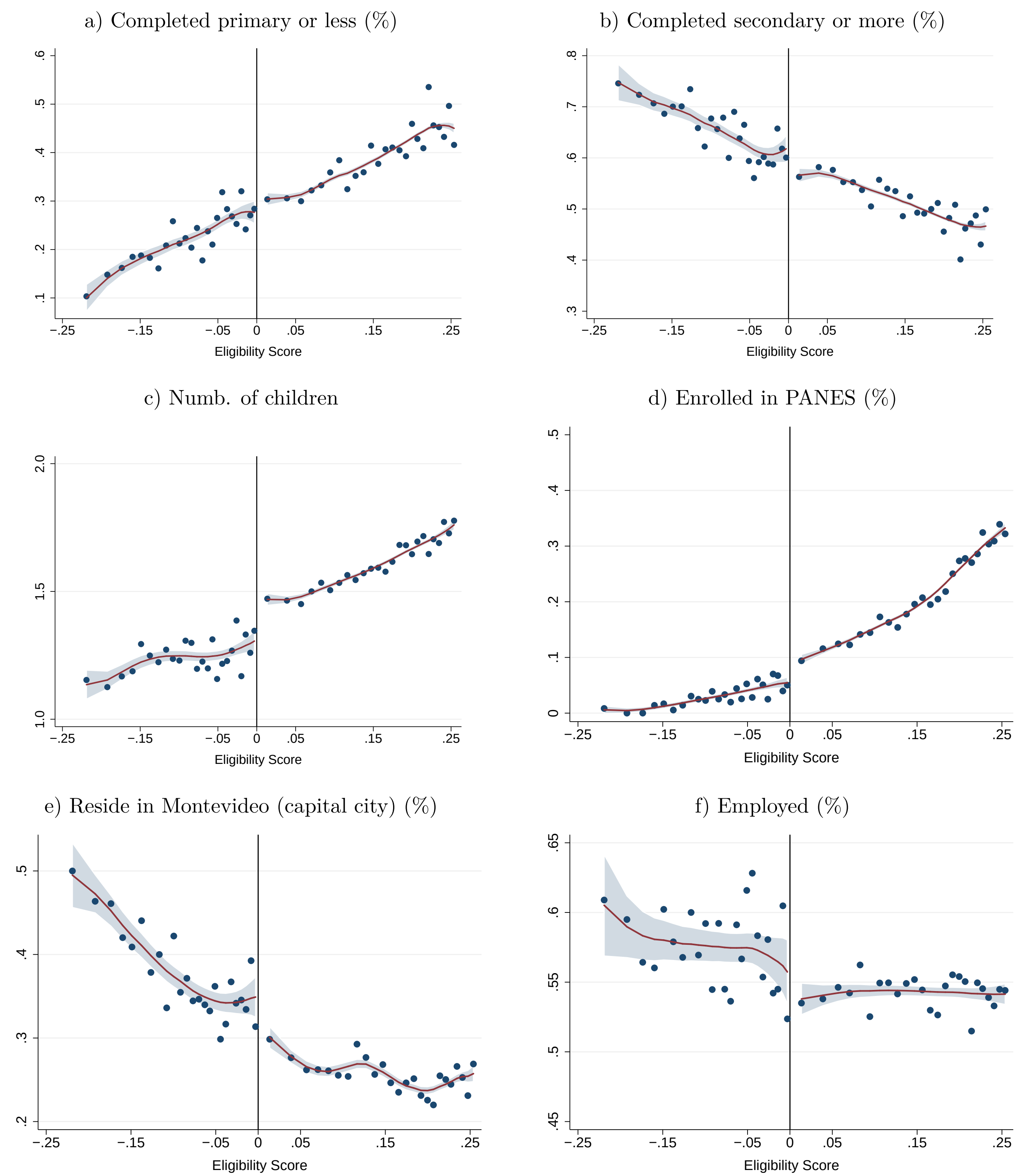
A.5.1 Balance Figures for Baseline Characteristics - Main Sample

Figure A.5.1: Covariates RD Plots



Notes: These figures plot pre-application characteristics against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one-half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold, along with the 95% confidence interval.

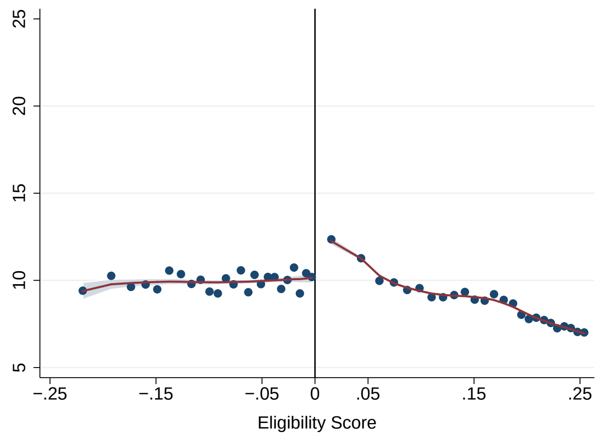
Figure A.5.2: RD Covariate Plots



Notes: These figures plot pre-application characteristics against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one-half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold, along with the 95% confidence interval.

Figure A.5.3: RD Covariate Plots

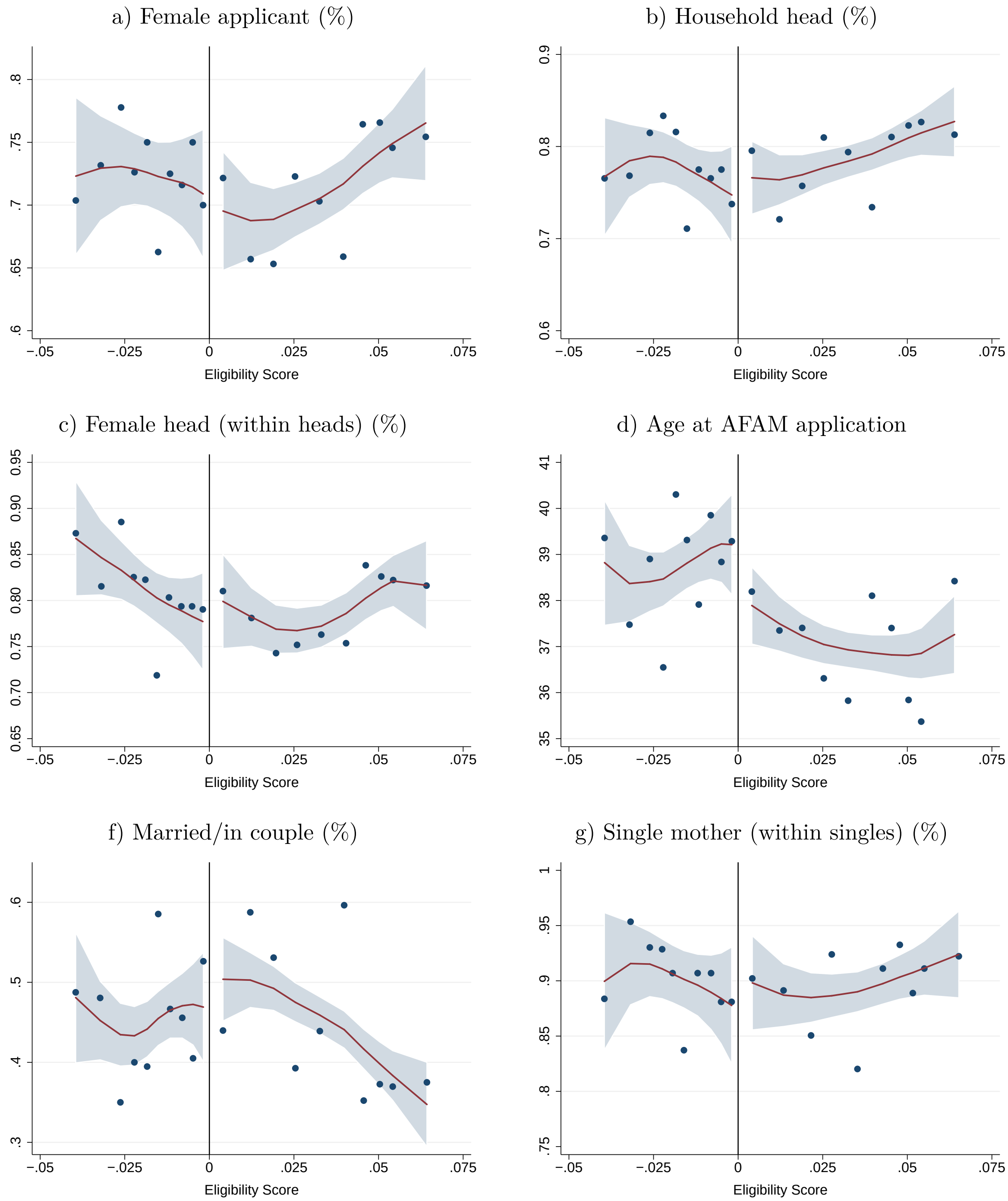
Application Date (Months since 01/2008)



Notes: These figures plot pre-application characteristics against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008-September 2010. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010). The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one-half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold, along with the 95% confidence interval.

A.5.2 Balance Figures for Baseline Characteristics - Follow-Up Sample

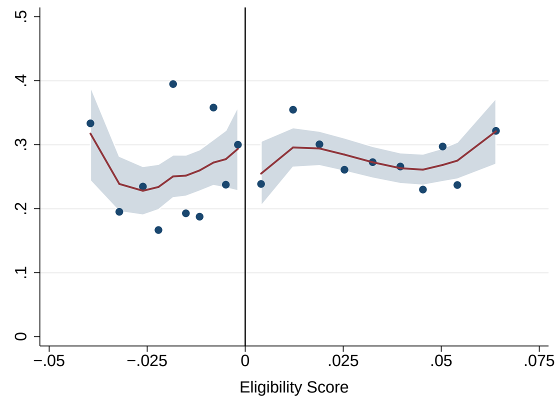
Figure A.5.4: Covariates RD Plots



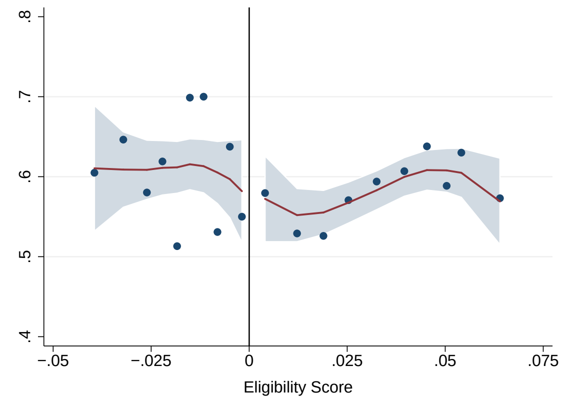
Notes: These figures plot pre-application characteristics against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008-September 2010, from households within the eligibility score range [−0.0426; +0.0727]. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the Follow-up Survey. The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one-half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold without additional covariates, along with the 95% confidence interval.

Figure A.5.5: Covariates RD Plots

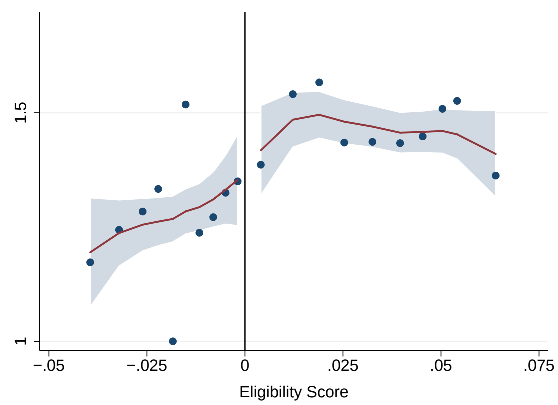
a) Completed primary or less (%)



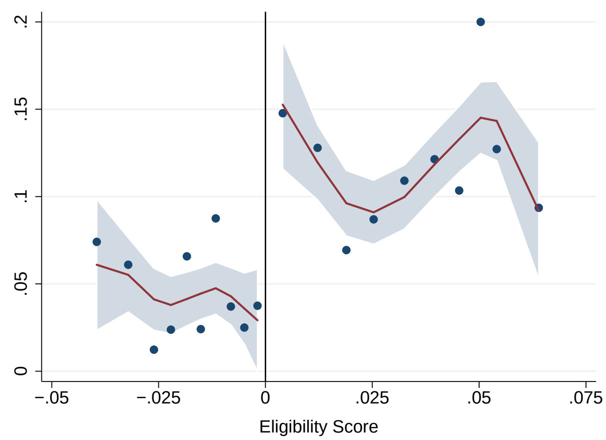
b) Completed secondary or more (%)



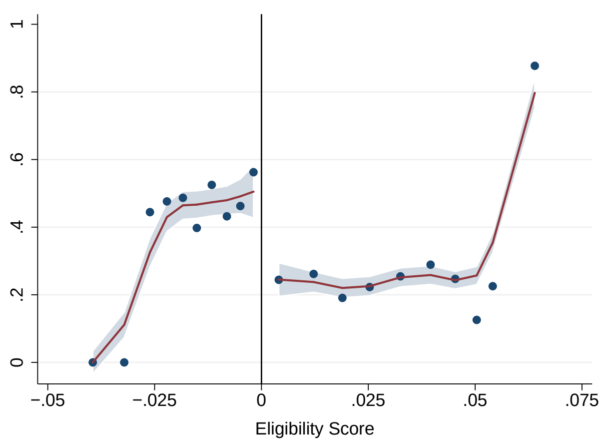
c) Numb. of children



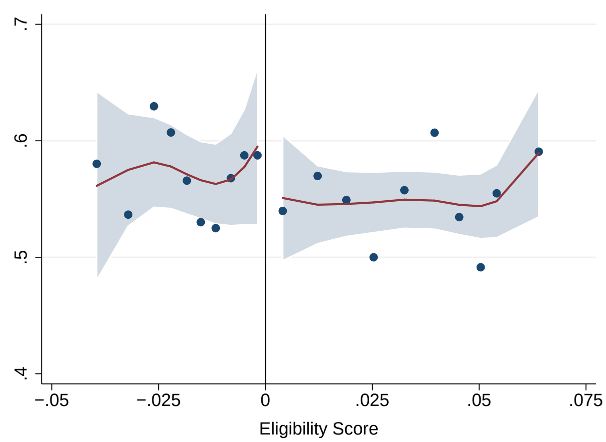
d) Enrolled in PANES (%)



e) Reside in Montevideo (capital city) (%)

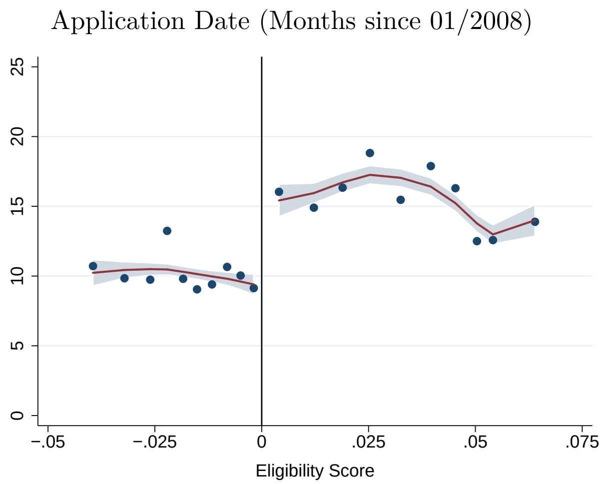


f) Employed before application (%)



Notes: These figures plot pre-application characteristics against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008- September 2010, from households within the eligibility score range [−0.0426; +0.0727]. The dataset corresponds to the *AFAM administrative records* (January 2008- September 2010) matched with the Follow-up Survey. The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one-half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold without additional covariates, along with the 95% confidence interval.

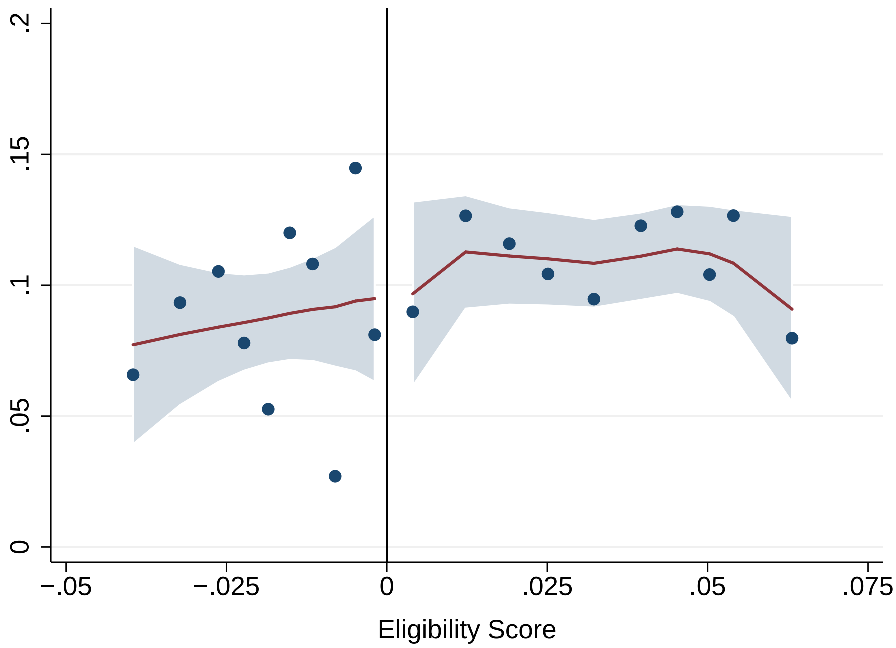
Figure A.5.6: Covariates RD Plots



Notes: These figures plot pre-application characteristics against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008- September 2010, from households within the eligibility score range $[-0.0426; +0.0727]$. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the Follow-up Survey. The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one -half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold without additional covariates, along with the 95% confidence interval.

A.5.3 Balance Figure for the Discrepancy Rate in Computation of Registered Employment

Figure A.5.7: RD Plot for Discrepancy Rate in Measurement of Registered Employment: SSA Records vs. Self-Declared Reporting (Follow-Up Survey)



Notes: This figure plots the discrepancy rate in measurement of registered employment against the eligibility score. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of application to AFAM during the period January 2008-September 2010, from households within the eligibility score range [−0.0426; +0.0727]. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the *SSA registered employment records* from the SSA's administrative records for the period January 2005-December 2012.The discrepancy variable is an indicator equal to 1 if the registered employment measured using *SSA registered employment records* does not coincide with self-reported employment status from the follow-up survey. The eligibility score is standardized so that the eligibility threshold is zero, with positive scores indicating individuals in eligible households and negative scores indicating individuals in ineligible households. Each point (blue circle) in the plot represents the average value of the outcome variable in eligibility score bins with a width of one-half of a percentage point. The red solid line plots estimated means from a local linear regression estimated at each side of the eligibility threshold without additional covariates, along with the 95% confidence interval.

A.6 Additional Results: Tables

A.6.1 Propensity to be a Registered Employee

Table A.6.1: Determinants of Propensity to be a Registered Employee

Dep. var.: prob. of registered employment	
Income Score	-0.149***
	(0.055)
Female applicants	0.002
	(0.008)
Age at application to AFAM	-0.002***
	(0.000)
Household head	0.019**
	(0.009)
Montevideo	0.045***
	(0.007)
Enrolled PANES	-0.032*
	(0.018)
Number of children	0.009**
	(0.004)
Married	0.005
	(0.009)
Married missing	0.040**
	(0.017)
Complete Primary or less	-0.007
	(0.011)
Secondary or more	0.009
	(0.010)
Employed	0.194***
	(0.008)
Months since application	-0.002***
	(0.001)
Registered 36 months pre	0.620***
	(0.009)
Constant	0.179***
	(0.024)
R-squared	0.37
Observations	223,416

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the *SSA registered employment records* for the period January 2005 to December 2012 from the SSA's administrative records (see Section 3.2 for a detailed description of the data). The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. All individual/household characteristics presented in this table are measured on the date of application or before, i.e. before the administrative decision on enrollment in the program. Huber-White robust standard errors are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

A.6.2 Heterogeneous Effects of AFAM by Socioeconomic Subgroups

Table A.6.2: Heterogeneous Effects of AFAM Eligibility by Socio-Demographic Sub-Groups

	Estimates (DD-RD)	BW	Comparison Mean	%Δ w.r.t. (1)	Diff. btw. Groups	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Household head	-0.067	0.058	0.475	-14.045	-0.004	219,552
	(0.019)			(4.015)	[0.922]	
	[0.001]			[0.001]		
Other HH. member	-0.063	0.060	0.439	-14.387		64,368
	(0.036)			(8.189)		
	[0.132]			[0.132]		
Female applicants	-0.064	0.055	0.423	-15.015	0.000	183,936
	(0.020)			(4.703)	[1.000]	
	[0.003]			[0.003]		
Male applicants	-0.064	0.071	0.574	-11.234		109,008
	(0.030)			(5.269)		
	[0.071]			[0.071]		
Two parent HH	-0.052	0.058	0.459	-11.241	0.021	120,048
	(0.027)			(5.862)	[0.559]	
	[0.091]			[0.091]		
Single parent HH	-0.073	0.055	0.468	-15.664		135,864
	(0.024)			(5.106)		
	[0.004]			[0.004]		
Single female	-0.087	0.064	0.457	-18.991	-0.093	147,672
	(0.023)			(5.060)	[0.235]	
	[0.000]			[0.000]		
Single male	0.006	0.053	0.567	1.022		15,984
	(0.075)			(13.187)		
	[0.744]			[0.744]		
Low education	-0.054	0.073	0.437	-12.289	0.015	114,240
	(0.030)			(6.773)	[0.684]	
	[0.092]			[0.092]		
Medium-high education	-0.069	0.057	0.476	-14.536		158,808
	(0.022)			(4.616)		
	[0.003]			[0.003]		
Aged 45 or less	-0.062	0.052	0.474	-13.041	0.011	192,360
	(0.020)			(4.274)	[0.767]	
	[0.006]			[0.006]		
Aged above 45	-0.073	0.087	0.446	-16.286		98,544
	(0.031)			(6.965)		
	[0.015]			[0.015]		

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Each row presents the estimates from the DD-RD model in equation (2) with covariates at time of application to the program for the corresponding subgroup, as in the notes to Table 1. Estimates from local linear regressions use a triangular kernel. The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. Column (1) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "BW" in Column (2) reports optimal bandwidths following the Calonico et al. (2014) procedure. "Comparison Mean" in Column (3) reports the average of the dependent variable for ineligible individuals within the score bandwidth. Column (4) reports the program's effect from column (1) as a percentage of the mean of the dependent variable for ineligible individuals at the cutoff (column 3). Column (5) reports the estimate of the difference in AFAM´s effect between groups and and the corresponding p-values (square parentheses). Column (6) reports the total number of observations.

A.6.3 Heterogeneous Effects of AFAM by Propensity to be Formal: Full Estimates

Table A.6.3: Effect of AFAM Eligibility on Registered Employment by Propensity to be a Registered Employee - All Individuals

	Estimates	BW	Comparison Mean	%Δ w.r.t. (1)	Obs.
	(1)	(2)	(3)	(4)	(5)
Panel A. Low prop. to be formal					
Elegible(RD Post)	-0.027	0.075	0.134	-19.965	113,568
	(0.017)			(12.477)	
	[0.148]			[0.148]	
Elegible(RD Pre)	-0.002	0.063	0.013	-14.253	136,937
	(0.003)			(23.809)	
	[0.484]			[0.484]	
Elegible(DD- RD)	-0.025	0.075	0.134	-18.572	113,568
	(0.017)			(12.478)	
	[0.182]			[0.182]	
Panel B. Medium prop. to be formal					
Elegible(RD Post)	-0.086	0.073	0.307	-28.197	118,056
	(0.025)			(8.239)	
	[0.000]			[0.000]	
Elegible(RD Pre)	0.003	0.088	0.106	2.527	235,357
	(0.006)			(5.722)	
	[0.622]			[0.622]	
Elegible(DD- RD)	-0.089	0.073	0.307	-29.062	118,056
	(0.025)			(8.239)	
	[0.000]			[0.000]	
Panel C. High prop. to be formal					
Elegible(RD Post)	-0.066	0.051	0.787	-8.331	86,496
	(0.026)			(3.342)	
	[0.019]			[0.019]	
Elegible(RD Pre)	-0.019	0.058	0.791	-2.399	157,435
	(0.017)			(2.120)	
	[0.236]			[0.236]	
Elegible(DD- RD)	-0.047	0.051	0.787	-5.926	86,496
	(0.026)			(3.342)	
	[0.104]			[0.104]	

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Rows report the estimates from the RD – using either pre (row 1) or post (row 2) application data – and DD-RD (row 3) models in equations (A.4.1, 2) with covariates at time of application to the program for the corresponding subgroup, as in the notes to Table 1. Estimates from local linear regressions use a triangular kernel. The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. Column (1) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "BW" in Column (2) reports optimal bandwidths following the Calonico et al. (2014) procedure. "Comparison Mean" in Column (3) reports the average of the dependent variable for ineligible individuals within the score bandwidth. Column (4) reports the program's effect from column (1) as a percentage of the mean of the dependent variable for ineligible individuals at the cutoff (column 3). Column (5) reports the estimate of the difference in AFAM´s effect between groups and and the corresponding p-values (square parentheses). Column (6) reports the total number of observations.

Table A.6.4: Effect of AFAM Eligibility on Registered Employment by Propensity to be a Registered Employee - Single Women

	Estimates	BW	Comparison Mean	%Δ w.r.t. (1)	Obs.
	(1)	(2)	(3)	(4)	(5)
Panel A. Low prop. to be formal					
Elegible(RD Post)	0.036	0.038	0.123	29.150	18,408
	(0.030)			(24.700)	
	[0.145]			[0.145]	
Elegible(RD Pre)	-0.003	0.075	0.010	-33.068	74,999
	(0.004)			(38.864)	
	[0.291]			[0.291]	
Elegible(DD- RD)	0.039	0.038	0.123	31.771	18,408
	(0.030)			(24.696)	
	[0.118]			[0.118]	
Panel B. Medium prop. to be formal					
Elegible(RD Post)	-0.122	0.056	0.250	-48.723	38,160
	(0.036)			(14.236)	
	[0.001]			[0.001]	
Elegible(RD Pre)	0.002	0.083	0.070	2.865	102,675
	(0.011)			(16.257)	
	[0.995]			[0.995]	
Elegible(DD- RD)	-0.124	0.056	0.250	-49.518	38,160
	(0.036)			(14.236)	
	[0.001]			[0.001]	
Panel C. High prop. to be formal					
Elegible(RD Post)	-0.147	0.052	0.715	-20.586	47,064
	(0.039)			(5.409)	
	[0.000]			[0.000]	
Elegible(RD Pre)	-0.051	0.083	0.653	-7.747	138,491
	(0.024)			(3.601)	
	[0.019]			[0.019]	
Elegible(DD- RD)	-0.098	0.052	0.715	-13.649	47,088
	(0.039)			(5.407)	
	[0.019]			[0.019]	

Notes: The sample corresponds to single mothers, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Rows report the estimates from the RD – using either pre (row 1) or post (row 2) application data – and DD-RD (row 3) models in equations (A.4.1, 2) with covariates at time of application to the program for the corresponding subgroup, as in the notes to Table 1. Estimates from local linear regressions use a triangular kernel. The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. Column (1) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "BW" in Column (2) reports optimal bandwidths following the Calonico et al. (2014) procedure. "Comparison Mean" in Column (3) reports the average of the dependent variable for ineligible individuals within the score bandwidth. Column (4) reports the program's effect from column (1) as a percentage of the mean of the dependent variable for ineligible individuals at the cutoff (column 3). Column (5) reports the estimate of the difference in AFAM´s effect between groups and and the corresponding p-values (square parentheses). Column (6) reports the total number of observations.

A.6.4 Robustness of the Effects of AFAM on Different Margins of Participation

Table A.6.5: Discrepancies in Measurement of Registered Employment When Using Administrative or Self-declared Information

	SSA records		
Self-declared F-U Survey	Non Employed / Informal	Registered	Total
Non Employed / Informal	1,183	152	1,335
Registered	94	974	1,068
Total	1,277	1,126	2,403

Notes: This table reports results from a cross-tabulation analysis between the variable registered employment measured by using information from the SSA's administrative records ("SSA records") and self-reported information from the follow-up survey ("Self-reported F-U Survey"). Registered employment is measured as an indicator variable equal to 1 if the SSA records (F-U Survey) indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, who were interview during the Follow-up Survey – i.e. from households within the eligibility score range $[-0.0426; +0.0727]$. The dataset corresponds to the *AFAM administrative records* (January 2008-September 2010) matched with the *SSA registered employment records* for the period January 2005 to December 2012 from the SSA's administrative records and with the Follow-up Survey (see Section 3.2 for a detailed description of the data).

Table A.6.6: Discrepancies in Measurement of Registered Employment When Using Administrative or Self-declared Information – Eligible Individuals

	SSA records		
Self-declared F-U Survey	Non Employed / Informal	Registered	Total
Non Employed / Informal	875	115	990
Registered	65	601	666
Total	940	716	1,656

Notes: This table replicates the analysis shown in Table A.6.5, but only for eligible individuals. See notes to Table A.6.5.

Table A.6.7: Discrepancies in Measurement of Registered Employment When Using Administrative or Self-declared Information – Ineligible Individuals

	SSA records		
Self-declared F-U Survey	Non Employed / Informal	Registered	Total
Non Employed / Informal	308	37	345
Registered	29	373	402
Total	337	410	747

Notes: This table replicates the analysis shown in Table A.6.5, but only for ineligible individuals. See notes to Table A.6.5.

A.6.5 Robustness of the Effects of AFAM on Registered Employment

- The sample of analysis includes observations with missing data in the imputed income variables (about 5%)

Table A.6.8: Effect of AFAM Eligibility on Registered Employment

	Estimates	BW	Comparison Mean	%Δ w.r.t. (1)	Obs.
	(1)	(2)	(3)	(4)	(5)
Panel A. No Covariates					
Eligible(RD Post)	-0.097	0.049	0.465	-20.782	225,216
	(0.020)			(4.255)	
	[0.000]			[0.000]	
Eligible(RD Pre)	-0.037	0.065	0.376	-9.749	514,337
	(0.016)			(4.319)	
	[0.033]			[0.033]	
Eligible(DD-RD)	-0.060	0.050	0.466	-12.926	225,744
	(0.020)			(4.245)	
	[0.006]			[0.006]	
Panel B. With Covariates					
Eligible(RD Post)	-0.093	0.052	0.466	-19.965	242,016
	(0.018)			(3.886)	
	[0.000]			[0.000]	
Eligible(RD Pre)	-0.030	0.074	0.376	-8.106	603,248
	(0.015)			(3.926)	
	[0.037]			[0.037]	
Eligible(DD-RD)	-0.063	0.052	0.466	-13.495	243,024
	(0.018)			(3.883)	
	[0.001]			[0.001]	

Notes: The sample corresponds to heads of households, and spouses of heads of households, aged 18 to 57 at the time of the AFAM application during the period January 2008-September 2010, from households within the eligibility score range $[-0.257; +0.257]$. The dataset corresponds to the *AFAM administrative records* matched with the *SSA registered employment records* for the pre-application (36 months) and post-application (24 months) period according to the SSA's administrative records (see Section 3.2 for a detailed description of the data). Panels A and B present the estimates from the RD – using either pre (row 1) or post (row 2) application data – and DD-RD (row 3) models in equations (A.4.1, 2) without and with socioeconomic covariates, respectively. Estimates from local linear regressions use a triangular kernel. The dependent variable is registered employment, measured as an indicator variable equal to 1 if the SSA records indicate that there are social security contributions from employment for the individual in a given calendar month, and zero otherwise. The covariates in the regressions in Panel B include gender, head of household status, age, marital status, educational level (in 3 categories), the number of children aged 0-17 in the household, whether the household was enrolled in the PANES program, residence in Montevideo (Uruguay's capital), and whether the individual was employed. All regressions include time and date-of-application fixed-effects. Column (1) reports "conventional estimate" coefficients and household-clustered standard errors (curved parentheses), and p-value "bias-corrected estimates" (square parentheses) following Calonico et al. (2014). "BW" in Column (2) reports optimal bandwidths following theCalonico et al. (2014) procedure. "Comparison Mean" in Column (3) reports the average of the dependent variable for ineligible individuals within the score bandwidth. Column (4) reports the program's effect from column (1) as a percentage of the mean of the dependent variable for ineligible individuals at the cutoff (column 3). Column (5) reports the total number of observations.