FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN
DEPARTAMENTO DE ECONOMÍA
UNIVERSIDAD DE LA REPÚBLICA URUGUAY

Estigma luego de la adhesión:

Evidencia de transferencias públicas de Uruguay

Rodrigo Nicolau Blanchet

Programa de Maestría en Economía de la Facultad de Ciencias Económicas, Universidad de la República.

Montevideo - Uruguay

Abril de 2023

FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN
DEPARTAMENTO DE ECONOMÍA
UNIVERSIDAD DE LA REPÚBLICA URUGUAY

Estigma luego de la adhesión:

Evidencia de transferencias públicas de Uruguay

Rodrigo Nicolau Blanchet

Tesis de Maestría presentada al Programa de Maestría en Economía de la Facultad de Ciencias Económicas, Universidad de la República, como parte de los requisitos para la obtención del título de Magíster en Economía.

Director de tesis:

Prof. Andrea Vigorito

Codirector de tesis:

Prof. Martin Leites

Director académico:

Prof. Rodrigo Ceni

Montevideo - Uruguay

Abril de 2023

INTEGRANTES DEL TRIBUNAL DE DEFENSA DE TESIS

_____

_____

_____

Montevideo - Uruguay

Abril de 2023

Agradecimientos

En primer lugar, quiero agradecer a mis tutores Andrea Vigorito y Martin Leites por su guía, paciencia (sobre todo paciencia) e innumerables consejos antes, durante y después de todo el proceso de esta tesis, y en mi vida de estudiante e investigador en formación. En segundo lugar, agradezco a Elena Cettolin, Patricio Dalton, Koen Decancq, Francisco Ferreira y Marco Manacorda por sus valiosos aportes y discusiones sobre los borradores preliminares de este documento. En tercer lugar, quisiera dar las gracias a los participantes en el Seminario de Economía Interna, la Charla de Microeconomía Aplicada y el Seminario de Estudiantes de Posgrado de la Universidad de Tilburg; el Seminario del Instituto de Economía; la Conferencia sobre Desarrollo Humano y Enfoque de Capacidades (HDCA); y la Conferencia Internacional de la Sociedad Adam Smith; por sus valiosos comentarios y sugerencias. En cuarto lugar, valoro las diversas discusiones mantenidas a lo largo de los dos últimos años en el grupo de investigación Ética, Justicia y Economía (FCEA - FHUCE) con Ana Fascioli, Yamila Lara, Helena Modzelewski, Juan Olano, Gustavo Pereira y Agustín Reyes que han nutrido a este trabajo. En quinto lugar, quisiera agradecer y reconocer a la Comisión Académica de Posgrados (CAP - CSIC - UDELAR), al Programa de Movilidad e Intercambio Académico (MIA - CSIC - UDELAR), y al Fondo Clemente Estable (FCE - ANII) y al Programa de la Maestría en Economía por su ayuda económica que hizo posible la redacción de este estudio y el trayecto por el programa de maestría. Por último, pero no por ello menos importante, agradezco a mi familia y amigos su constante cariño y apoyo. Cualquier error en este texto es de mi exclusiva responsabilidad.

“By necessaries I understand not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it indecent for creditable people, even the lowest order, to be without. A linen shirt, for example, is, strictly speaking, not a necessary of life. The Greeks and Romans lived, I suppose, very comfortably though they had no linen. But in the present times, through a greater part of Europe, a creditable day-labourer would be ashamed to appear in public without a linen shirt, the want of which would be supposed to denote that disgraceful degree of poverty which, it is presumed, nobody can well fall into without extreme bad conduct. Custom, in the same manner, has rendered leather shoes a necessary of life in England. The poorest credible person of either sex would be ashamed to appear in public without them."

Adam Smith, An Inquiry into the Nature and Causes of the Wealth of Nations, pp. 351-2

Resumen

El estigma de la asistencia social, es decir, la desutilidad de participar en políticas sociales per se, se ha estudiado principalmente en la literatura económica como un factor determinante de la adhesión incompleta de las políticas sociales gubernamentales, como las transferencias monetarias. Sin embargo, se sabe menos sobre lo que ocurre con la estigmatización de los individuos una vez que reciben efectivamente las prestaciones. Hasta el momento, esta cuestión sigue sin explorarse, ya que las evaluaciones de impacto de estos programas se han centrado principalmente en otros resultados de bienestar subjetivo y no en el estigma en sí. Este estudio busca abordar dicha brecha en la literatura mediante el análisis empírico de los efectos estigmatizantes de los dos programas públicos de transferencias públicas no contributivas más importantes de Uruguay: Asignaciones Familiares - Plan de Equidad y Tarjeta Uruguay Social. La estrategia de identificación emplea un Diseño de Regresión Discontinua (RDD, por su sigla en inglés) que explota un índice de vulnerabilidad que rige la asignación a cada programa. Los datos utilizados en este estudio provienen de registros administrativos (2008-2010) y de una encuesta de seguimiento detallada (2016-2018) que incluye preguntas específicas sobre sentimientos de vergüenza y humillación en contextos de pobreza. Los resultados sugieren que la participación en el programa aumenta los sentimientos de propensión a la vergüenza y humillación entre los beneficiarios entre 0,35 y 0,65 desvíos estándar, dependiendo de la especificación y del programa considerado. Estos efectos varían entre ambas políticas, lo que podría explicarse por diferencias en las características institucionales entre ellas. Tener en cuenta los efectos del estigma asociado a la asistencia social puede servir para informar el diseño de políticas públicas y mejorar potencialmente el bienestar de los beneficiarios.

Palabras clave

estigma; vergüenza; programas sociales; vergüenza; humillación; diseño de regresión discontinua; Asignaciones Familiares - Plan de Equidad; Tarjeta Uruguay Social; Uruguay.

Clasificación JEL: I38; J15; Z18

Abstract

Welfare stigma, i.e., the disutility coming from participating in social welfare programs per se, has been primarily studied by the economics literature as a determinant of incomplete take-up of social assistance policies, such as cash transfers. However, less is known about what happens to individuals' feelings of stigma after they effectively take-up the benefits. This question remains under-explored so far, as impact evaluations of these programs have mainly focused on other subjective well-being outcomes rather than stigma itself. This study looks to address such gap by empirically analyzing the stigmatizing effects of Uruguay's two largest non-contributory public cash transfer programs: Asignaciones Familiares - Plan de Equidad and Tarjeta Uruguay Social. The identification strategy employs a Regression Discontinuity Design (RDD) that exploits a vulnerability index which rules the assignment to each program. The data used in this study come from administrative records (2008-2010) and a detailed follow-up survey (2016-2018) that includes specific questions regarding shame and humiliation feelings in the context of poverty. The findings suggest that program participation increases self-reported feelings of shame and humiliation among beneficiaries between 0.35 and 0.65 SD, depending on the specification and on the program. These effects vary across both policies which might be explained by differences in institutional features between them. Taking welfare stigma effects into account can inform policy design and potentially improve the overall well-being of beneficiaries.

Keywords

welfare stigma; social welfare programs; shame; humiliation; regression discontinuity design; Asignaciones Familiares - Plan de Equidad; Tarjeta Uruguay Social; Uruguay.

JEL Codes: I38; J15; Z18

Contents

1 Introduction

Social welfare programs are commonly subject to criticism regarding both their supply and demand sides from those who admonish governmental social assistance policies. Their supply is usually criticized for fostering moral hazard, laziness, wasteful consumption, and permanent dependency on "handouts" (Baumberg et al., 2012). Regarding their demand, welfare beneficiaries are often morally questioned for not being able to be financially self-sufficient, and needing help from the government to get by (Walker, 2014; Lamont et al., 2016). Therefore, these type of policies may raise doubts about its own efficacy and might contribute to convey or reinforce negative beliefs concerning participants' deservingness and worth as citizens (Campbell, 2012).

Such negative beliefs can be observed in the Uruguayan context. Over the last two decades there has been a rise of animosity towards the poor and beneficiaries of social assistance in the public discourse. Within the country, there has been a striking increase in the predominance of the notion that people are poor "because they are lazy and lack of will".[1] Between Latin American and the Caribbean (LAC) countries, Uruguay seems to be the society where the belief that "social welfare recipients are lazy" is the most prevalent among the region.[2] This is in spite of the fact that Uruguay is one of the countries with the highest levels of public social spending and preferences for redistribution in the region. In addition, qualitatively evidence recovered from in-depth interviews and focus groups in Moreno et al. (2014) and Rivero (2020) also document the presence of these stigmatizing views about those on welfare, particularly regarding the "undeserving poor".[3]

This antipathy surrounding welfare participation may confer stigma towards beneficiaries, i.e., an attribute or stereotype that is deeply discrediting and that reduces those stigmatized to a devalued identity by tagging them as tainted individuals (Goffman, 1963). The economics literature has gathered interest in studying stigma on different contexts: social behaviour, norms and identity (Lindbeck et al., 1999; Hungerman, 2013; Bursztyn et al., 2020; Ghosal et al., 2022) ; labor market choices (Chakraborty et al., 2018; Osman and Speer, 2022); and take-up decisions of welfare programs (Currie, 2004; Kleven and Kopczuk, 2011; Bargain et al., 2012; Bhargava and Manoli, 2015; Ko and Moffitt, 2022). Within the latter, the concept of welfare stigma has been defined as the disutility or psychological cost arising from being on welfare *per se*, and its study has focused on its relation with incomplete take-up of social assistance policies in developed countries (Moffitt, 1983; Besley and Coate, 1992).[4]

However, welfare stigma not only can discourage welfare participation, but could also still negatively affect the -subjective- well-being of those individuals who actually participate and

[1] This is shown in Panel (a) of Figure A1 which presents data for Uruguay from the World Value Survey (WVS) between 1996 and 2011 about people's perceptions of why there are still people in need in the country.

[2] This is shown in Panel (b) of Figure A1 shows data for LAC countries from the Americas Barometer for 2012 about people's opinion on whether those who receive social assistance from government programs are lazy. Uruguay reports the highest average response among selected countries.

[3] Quotes extracted from these interviews are presented in Section 3.3.

[4] The literature has also identified more determinants of incomplete take-up other than welfare stigma, such as asymmetry of information and transaction costs (Currie, 2004; Bhargava and Manoli, 2015).

become recipients (Besley and Coate, 1992). This type of stigma can be (un)intentionally embodied in the design of welfare policy (targeting, means-testing, etc.) and can often be reinforced by humiliation or discrimination from other members and institutions of society such as taxpayers, neighbors, employers or media, having potential long lasting consequences in terms of access to opportunities, social status, dignity or agency (Link and Phelan, 2001; Spicker, 2011; Walker, 2014; Baumberg, 2015). While the role of welfare stigma in deterring take-up has been well documented, quantitative research on how such stigma operates after individuals receive welfare benefits and how could they be affected by it has mostly been scarce, especially for cash transfer programs and the LAC region.

Cash transfers have been widely implemented over the last two decades in LAC as one of the main social assistance policies designed to alleviate poverty. Between 1996 and 2016 cash transfers have quadrupled, as shown in Panel (a) of Figure A2. Over this period, cash transfers have helped to address their main goal: as seen in Panel (b) of Figure A2 extreme poverty and poverty have substantially decreased in the region over time (Fiszbein et al., 2009; Stampini and Tornarolli, 2012). Meanwhile, although not in every case, they have often been implemented by national governments and not by private non-governmental organizations, as could be the case for other low -and middle - income regions. This is reflected by the increase on public expenditure in these programs (Panel (c) of Figure A2). In addition, for the most part they have been framed outside the traditional social security system, i.e., more than half have been executed or lead by Social Development Ministries or Presidential Offices (Atuesta and Cecchini, 2017).

Altogether these different aspects set up an crucial environment in the region where welfare stigma can arise, or be fostered, from "tax payer resentment" (Besley and Coate, 1992) and may be grounded on "poor are lazy" stereotypes (Nussbaum, 2009; Walker, 2014; Lamont, 2018; Lamont et al., 2016; Lamont, 2019). Since these policies are often financed by taxpayers, the reciprocity between beneficiaries and other members of society could strain, fostering stigmatizing processes and negative views on the poor and those who need government's help (Stuber and Schlesinger, 2006). Moreover, cash transfers in the region have been designed by being targeted to the poor, which could potentially foster stigma in the context of societies with low preferences for redistribution (Coady et al., 2004). Lastly, while there is plenty of evidence of the effects of these programs on traditional economic dimensions such as income, health and education (Fiszbein et al., 2009; Stampini and Tornarolli, 2012), less is known about their limitations, specifically what happens with welfare stigma in the continent in spite of being a region where this type of stigma could emerge.

This study aims to address this gap by empirically analyzing the welfare stigmatizing effects of cash transfer programs, using the case of Uruguay's two main non-contributory transfers: *Asignaciones Familiares - Plan de Equidad* (AFAM-PE) and *Tarjeta Uruguay Social* (TUS).[5] The main characteristics of these policies are the following. AFAM-PE is a conditional cash transfer

[5]Uruguay is a small country in Latin America with 3.3 million inhabitants and is one of the most developed countries in the region according to different development indicators. According to World Banks's data, it ranks 3rd in GDP per capita, 3rd in life expectancy and also 3rd in the Human Development Index (HDI) developed by the United Nation Development Programme (UNDP) among Latin American and Caribbean countries.

targeted to socioeconomically vulnerable households with children under the age of 18 or pregnant women.[6] Its monthly amount equals U$49 (11% of the national minimum wage) for the first child beneficiary and its coverage is widespread, reaching approximately 15% of households (37% of those with children under 18) in 2017.[7] Meanwhile, TUS is an unconditional transfer delivered through a magnetic card that is restricted to purchases of basic necessity goods, targeted to the most vulnerable population of the country.[8] Its monthly amount equals U$33 (8% of the national minimum wage) for the first child beneficiary while it has a lesser coverage, reaching around 6% of households (15% of those with children under 18) in 2017.

Broadly, in programs such as AFAM-PE and TUS, shame and stigma can occur through different mechanisms and in opposite directions. Among others, the literature identifies the non-contributory character (in coexistence with other contributory policies), targeting, means-testing, form of delivery and, conditionalities required to beneficiary households as mechanisms that might foster shame and stigma (van Oorschot, 2002; Coady et al., 2004; Mkandawire, 2005; Stuber and Schlesinger, 2006; de Brauw and Hoddinott, 2011; Walker, 2014; Roelen, 2017, 2020). On the contrary, a monetary transfer could also reduce shame and stigma due to the fact that it increases household income and consumption, which could help individuals fill the material needs required to adequately participate in social life (Bastagli et al., 2016; Devereux, 2016; Molyneux et al., 2016).[9] This is why whether these policies negate or promote shame and stigma is an empirical question, which still remains open (Molyneux et al., 2016; Roelen, 2020).

To shed light on such question this study follows an identification strategy that takes advantage of the assignment rule to AFAM-PE and TUS, and relies on the discontinuity in the probability of participation in both programs according to a proxy means-test index: *Índice de Carencias Críticas* (ICC). A Regression Discontinuity Design (RDD) is thus used for the econometric analysis. The data come from administrative records (2008-2010) and a detailed follow-up survey (2016-2018) that includes specific questions regarding shame and humiliation feelings in the context of poverty developed by the Oxford Poverty and Human Development Initiative (OPHI) which provide in-depth approximate measures of stigma related feelings and emotions (Zavaleta, 2007). Based on these variables, internal shame and external humiliation composite indexes are constructed and used as personal and social stigma outcomes, respectively. While the former relates to personal feeling of own self-devalued identity and could be related to personal stigma, the latter involves the perception of identity devaluation coming from other people which relates to social stigma (Walker, 2014; Baumberg, 2015).

[6]The conditionalities require children from beneficiary households to regularly attend formal medical checks and for those aged 4 to 17 to assist to formal education.

[7]These data come from a combination of *Ministerio de Desarrollo Social* (MIDES) and *Instituto de Estadística* (INE) official statistics.

[8]These basic necessity goods include food, hygiene items, among other basic products, except cigarettes and alcoholic beverages.

[9]Here the potential mechanisms of shame and stigma in the context of poverty and welfare policies are briefly mentioned due to the fact that this work does not intend to disentangle which specific drivers could be explaining the effects found.

The main results suggest that welfare participation increases the shame and humiliation feelings of their recipients, although these effects vary between both programs. There is suggestive evidence that AFAM-PE only increases the feeling of internal shame, while it has no effect on external humiliation. AFAM-PE participation increases the score of the shame proneness index between 0.34 and 0.46 standard deviations (SD). Stronger evidence indicates that TUS increases both, shame and humiliation. TUS beneficiaries report on average higher feelings of shame and humiliation by between 0.44 and 0.67, and 0.42 and 0.65 SD, respectively. Results on shame are significant at the 95% confidence level in both programs, while results on humiliation in TUS are significant at the 99% level. These results imply that the effect of AFAM-PE on shame would increase the mean for the non-eligible group by 99% from 1.062 to 2.112, moving an individual from the 53th to the 70th percentile in the shame scale distribution in the AFAM-PE sample. Meanwhile, the effect of TUS on shame (humiliation) would increase the non-eligible mean by 150% (115%) from 1.333 (0.960) to 3.337 (2.072), moving an individual from the 48th (48th) to the 75th (65th) percentile in the shame (humiliation) scale distribution in the TUS sample.

Although disentangling the specific mechanisms that could be explaining these findings is out of the scope of this study, well grounded hypotheses can be taken from it by analyzing the differences in the effects of the two policies. There are some key distinctions regarding both programs. In comparison to AFAM-PE, TUS is targeted to more vulnerable households, it has a lesser coverage, and its usage is mainly restricted to food purchases. None of these features are directly visible publicly to other members of society, so, if any, one could expect differences regarding personal stigma. This is supported by the data since impact estimates on internal shame are of higher magnitude for TUS than AFAM-PE. Also, impact estimates on social stigma are not similar: TUS has a considerable and significant effect on external humiliation while AFAM-PE has not. Such difference could be explained by the fact that for the former the delivery is publicly visible since the food card is tagged, while for the latter the delivery is private. This explanation would be consistent with both, previous theoretical and laboratory evidence (Friedrichsen et al., 2018) regarding the public vs. private nature of the transfers, and qualitative data from the TUS program (Moreno et al., 2014) regarding perceived discrimination of TUS beneficiaries in grocery stores. Nonetheless, further research is needed to provide conclusive evidence regarding this potential welfare stigma mechanism.

This study aims to contribute to the literature in many ways. First, to the best of my knowledge, it is one of the first studies to provide causal evidence of welfare stigmatizing effects of cash transfers over their beneficiaries. So far, the available research has either been qualitative or had non-causal identification (Hochfeld and Plagerson, 2011; Gao and Zhai, 2017; Zhang, 2020; Han and Gao, 2020; Celhay et al., 2022; Della Guardia et al., 2022; Lapham and Martinson, 2022). Second, previous causal evidence has used broad well-being questions and interpreted them as stigma outcomes (Qi and Wu, 2018). In the present study, the availability of the specific questions used regarding shame and humiliation in contexts of poverty allows to go beyond broad well-being outcomes and provides in-depth approximations of stigma related feelings which are more precise measures for this issue. In this sense, this study complements previous work that has utilized this set of questions in cross-sectional studies for different countries and populations in LAC (Mills and Zavaleta, 2015; Hojman and Álvaro Miranda, 2018).

Third, this research adds to the growing literature about the effects of cash transfers programs on subjective well-being on developing countries (Haushofer and Fehr, 2014; Haushofer and Shapiro, 2016; Haushofer et al., 2020, 2019; Romero et al., 2021). On one hand, this literature has concentrated on other outcomes, mainly life satisfaction, finding different effects. On the other hand, these papers often analyze individual related outcomes, while this study provides evidence over a subjective well-being dimension that may involve more social dynamics outside the individual. In this sense, the present work provides new insights about a "missing" dimension of well-being which is relevant to the understanding of life in deprivation but also important to the people living in poverty themselves (Alkire, 2007). Furthermore, the studies of cash transfers and subjective well-being have mostly focused on low -and middle - income countries in continents such as Africa and Asia, but not in LAC countries that are located in a setting which could be more prone to the occurrence of welfare stigma due to the aforementioned regional characteristics.

Finally, this analysis adds a new dimension to the large body of evidence regarding LAC and particularly Uruguayan cash transfers policies, where extensive studies have been carried out analyzing multiple impacts, including subjective outcomes such as political support, empowerment and interpersonal trust, among others (Manacorda et al., 2011; Amarante et al., 2016; Failache et al., 2016; Bérgolo and Galván, 2018; Colombo et al., 2018; Bérgolo and Cruces, 2021; Ceni and Salas, 2021). Yet, as mentioned before, research regarding their impacts on subjective dimensions such as shame, stigma, humiliation or discrimination has mostly been incipient.

The rest of the paper is organized as follows. Section 2 provides an overview of the institutional context of the policies under analysis. Section 3 establishes the main theoretical framework and relates it to the existing body of academic literature in economics and other social sciences, such as sociology and social psychology. Section 4 describes the data utilized in this study, including the survey sample and the questionnaire from where the the key outcomes come from. Section 5 explains the identification strategy used to estimate the causal impact of the policies. The main results of the study are presented in Section 6. Finally, Section 7 concludes.

2 Institutional context

2.1 Background

The Uruguayan child allowances system, *Asignaciones Familiares*, has a long-lasting history dating back to 1943 (Law No. 10,449) as a contributory transfer. Initially, the system was linked to formal labor market participation by the parents (it was a benefit granted to all formal workers) and later on it required formal education attendance for the children. However, successive reforms, during a period of recession and increasing child poverty, in 1995 (Law No. 16,697), 1999 (Law No. 17,139), and 2004 (Law No. 17,758) have shifted the focus to expand to vulnerable households who were not covered by the formal social security system. While the 1995 reform introduced the components of means-testing and targeting for formal workers below a

certain household income threshold[10], the 1999 and 2004 reforms turned such benefits to a non-contributory scheme, creating the *Asignaciones Familiares - Hogares de Menores Recursos.*[11] This non-contributory structure was substantially expanded in 2005 after the creation of the *Plan de Atención Nacional a la Emergencia Social* (PANES) by the Law No. 17,869 (Colafranceschi and Vigorito, 2013).[12]

PANES was a temporary social assistance policy enacted by the Uruguayan national government in response to the escalating poverty rate resulting from the 2002 economic crisis, which nearly doubled by 2004 reaching almost 40% of the total population living under the official poverty line. The objective of PANES was to provide immediate financial relief to low-income households. The policy consisted of many different components, two of which were non-contributory cash transfers: *Ingreso Ciudadano* and *Tarjeta Alimentaria.* PANES was targeted at extreme vulnerable households. On one hand, *Ingreso Ciudadano* entailed monthly cash payments of approximately US$ 55 over a two-year period between April 2005 and December 2007. On the other hand, *Tarjeta Alimentaria* was launched later in April 2006 to provide a cash supplement of around US$ 22 delivered through an electronic food card to households with pregnant women or children under the age of 18 (Manacorda et al., 2011; Amarante et al., 2016).[13]

Both of these policies precede each of the two programs that are the focus of analysis in this study: *Asignaciones Familiares - Plan de Equidad* (AFAM-PE) and *Tarjeta Uruguay Social* (TUS), respectively. Both constitute the largest scale non-contributory cash transfers in the country and are part of a broader policy reform called *Plan de Equidad* (Law No. 18,227) implemented in January 2008. It follows PANES' footsteps in providing monetary relief to low-income households, now for a permanent period, a wider objective population, and a higher coverage rate (Lavalleja and Tenenbaum, 2020).

2.2 Asignaciones Familiares - Plan de Equidad

AFAM-PE is a non-contributory, conditional cash transfer that was launched in January 2008 and has been jointly administered by the *Ministerio de Desarrollo Social* (MIDES) and the *Banco de Previsión Social* (BPS). It essentially unified the objective population of the aforementioned 1999 and 2004 reforms, and PANES, into one consolidated program. The policy has two primary objectives. On the one hand, it provides monetary assistance to socioeconomically vulnerable household households with children under the age of 18 or pregnant women, intending to cover around 500,000 beneficiaries or 200,000 households (Law No. 18,227). This would be equivalent to the first income quintile of the country (which is substantially higher than PANES previous target around 8% of total households). As a result, its coverage is vast, approximately 15% of Uruguayan households (37% of households with children under 18) in 2017. On the other

[10]The threshold was equivalent to ten times the national minimum wage.

[11]Up to that point, transfer amounts were relatively low.

[12]For a more comprehensive understanding of the history and structure of the Uruguayan social security system and child allowances, please refer to Arim et al. (2009).

[13]The start was delayed due to technological constraints regarding the implementation of the electronic card (Arim et al., 2009).

hand, it seeks to promote school attendance and medical check-ups among the beneficiary children (Bérgolo et al., 2016). The primary recipient of the cash transfers is the caregiver of the beneficiaries, preferably a woman, who has priority in collecting payments (Law No. 18,227).

The program employs a means-test and proxy means-test approach to rule the targeting and reach its objective population. The eligibility of households is determined first by their formal household income.[14] Applicant households file a sworn affidavit stating their self-reported income which is then double-checked against the social security records. Then, a second check is carried by proxy means-test through the *Índice de Carencias Críticas* (ICC). The index ranges from 0 to 1 and considers various household-level socioeconomic indicators that reflect structural dimensions of well-being.[15] This set of indicators are easily observable and highly correlated with belonging to the first income quintile, and include education, household composition, housing infrastructure, and possession of durable goods. Only households with scores above a certain ICC threshold, which takes a value around 0.2, are eligible for the benefits (Lavalleja and Tenenbaum, 2020). MIDES collected this information upon request from households that applied to the programs at different ministry's offices. It is worth noting that the exact computation and inputs of the model should not be publicly know since this is classified information to the general public (MIDES, 2013; Moreno et al., 2014).[16]

Thirdly, the receipt of the transfer is conditional upon the beneficiaries' subsequent compliance with two requirements concerning minors in the household: regular medical check-ups and formal education attendance for children aged 4 to 17 years. However, it is important to note that the national authorities did not monitor compliance with these conditionalities until mid-2013 (Rossel et al., 2014).[17] The cash allowances are disbursed on a monthly basis through a regular bank account from the national public bank, *Banco de la República Oriental del Uruguay* (BROU), and an associated regular electronic debit card, which is shown in Figure A3. Altough at the beggining of the program these were withdrawn in regular payment and collection branches. The cash amount is annually adjusted by law through the official Consumer Price Index, and are determined by the following formula:

$$AFAMPE_i = M_p * (under18)^{0.6} + M_s * (under18sec)^{0.6}$$

Where $AFAMPE_i$ is the total amount of the monthly stipend received by a given household, M_p represents the transferred amount for each beneficiary under 5 years old (including those in gestation) or attending primary education, and M_s is the complementary amount granted for

[14] Since 2022 this means-test check was revoked, and the only the proxy-means test remains active.

[15] This eligibility index is computed using a probit model that estimates the likelihood of an applicant household belonging the lowest quintile of per capita income.

[16] For a more detailed description of the index please refer to MIDES (2013).

[17] In April 2013, after a prolonged parliamentary debate, the government decided to suspend payments to households failing to comply with the program's conditions. As a result, 42,263 payments were interrupted, which were able to be reinstated if households provided proof of educational attendance by presenting the corresponding certificates from schools (Rossel and Straschnoy, 2020).

each beneficiary attending secondary education. Moreover, $under18$ denotes the total number of children under 18 in the household, and $under18sec$ represents the total number of children under 18 attending secondary education. As of January 2017, the values for M_p and M_s are approximately U$S 49 ($ 1,404) and U$S 21 ($ 601), respectively. As a reference point, U$S 49 equals to around 11% of the national minimum wage in the country for that year. [18]

2.3 Tarjeta Uruguay Social

TUS is an non-contributory, unconditional cash transfer restricted to basic necessities purchases, which took over *Tarjeta Alimentaria* in 2008 and has been administered exclusively by MIDES.[19] The program's main objective is to cover basic food and hygiene necessities of the population in most extreme vulnerable economic conditions. Accordingly, it can only be used to purchase food, hygiene items, and other basic products, while it cannot be used to buy cigarettes, alcoholic beverages, or soft drinks. Its intended coverage is around 60,000 households, mostly with children (MIDES, 2012).[20]

The transfer is disbursed through an electronic food card, which is accepted only at roughly 1,000 authorized stores affiliated in a "*Red de Comercios Solidarios*" (as of 2017). As shown in Figure A3, the card has a distinctive appearance, highlighting that the provider of the card is the government through a social plan. Its amount depends on the number of children in the household, and it is divided into four brackets, as detailed in Table A1.[21] Opposite of AFAM-PE, TUS is capped at a fixed amount and does not offer any additional amount due to educational incentives. As of January 2017, the basic amount of the transfer is approximately U$S 33 ($ 955). As a reference point, U$S 33 equals to around 8% of the national minimum wage in the country. As it can be seen in Table A1, the amount of the benefit is doubled for the 30,000 households in the most vulnerable socioeconomic conditions, who receive a "Double TUS" while the rest receives a "Simple TUS".[22]

As is the case with AFAM-PE, TUS also rules its targeting through a proxy means-test approach, employing the same *Índice de Carencias Críticas* to determine household program eligibility. In this case, a government enumerator visits applicant households to recover the main eligibility information. Since the program is targeted to more vulnerable households, it follows that its

[18] The transfer amount increases marginally with the number of children and pregnant women in the household, taking into account economies of scale within households. Additionally, the higher amount granted to households with children attending secondary schools is intended to serve as an attendance incentive, as the program explicitly aims to encourage children to complete secondary education (Law No. 18,227).

[19] The term "unconditional" refers to the absence of quid pro quos for beneficiary households, yet it can be argued that the restriction of the transfer's usage to purchases of basic necessities goods only effectively imposes a form of conditionality.

[20] In recent years, the program has expanded to cover specific adult populations facing extreme vulnerability, including transgender individuals, homeless individuals, and those with chronic illnesses (Oviedo, 2019).

[21] It should be noted that, since 2012, households without children receive an amount equivalent to those with one children (Moreno et al., 2014).

[22] In addition, the value is adjusted annually or semi-annually based on the Food and Non-alcoholic Beverages Price Index and purchases made with the card are exempt from VAT.

eligibility cut-offs through the vulnerability index take higher values than AFAM-PE: 0.6 for the "Simple TUS" (regular cash amount) and approximately 0.8 for the "Double TUS" (doubled cash amount) (MIDES, 2013).[23] For the purpose of this study, both TUS values are considered together due to sample size restrictions and because the only difference between them is the total amount of cash received, which varies across households throughout the whole sample. Figure A4 exhibits the different eligibility thresholds and the approximate proportion of intended beneficiary households for each program over the ICC range.

Finally, it is worth noting that although the application process and the management of the AFAM-PE and TUS programs are independent from each-other, they quite overlapped over their beneficiaries, especially when considering households with children. According to estimates based on data from the Official National Household Survey for 2017, approximately 77% of households that receive TUS are also beneficiaries of AFAM-PE. This estimated proportion rises to 88% when considering only households with children.

3 Literature review

3.1 Theoretical framework

In the economics literature welfare stigma is defined as the "disutility" or 'psychological cost" of participating on welfare programs *per se* (Moffitt, 1983; Besley and Coate, 1992). Following Moffitt (1983), in the simplest model where welfare stigma arises just from participation itself, and does not depend on the amount of benefit received, it is defined as follows:

$$U = U(y + b \cdot p) - \phi \cdot p \tag{1}$$

Where an individual's utility depends on their private income, y, b is the amount of welfare benefits received, p is a binary indicator of participation in a welfare program, and ϕ is a disutility component which represents welfare stigma. Although this framework is general enough to be applied in most settings, the economics literature has mainly used it to explain why individuals decide against applying for social assistance programs (Yaniv, 1997; Blumkin et al., 2008). Therefore, to the best of my knowledge, there is currently no theoretical model that comprehensively addresses the effects of welfare stigma on individuals who decide to participate in such programs. This is why it is necessary to draw on other social sciences which have been studying stigma for decades, such as sociology and social psychology, to gather the most relevant concepts and relations related to the subject.

Goffman (1963) initially defined stigma as a negative characteristic of an individual that discredits, devalues, and alienates her from the rest of society. He further argued that stigma is a social phenomenon that arises from the interaction between individuals who are perceived as different, and the broader society in which they live. Stigmatized individuals are often subject

[23] TUS also does not have a means-tested formal income check.

to various forms of discrimination and exclusion, and may experience negative effects on their self-esteem, identity, and social relationships. Meanwhile, Link and Phelan (2001) added and emphasized the relational nature of stigma and its key dependence on societal preferences. The authors' definition of stigma is based on the concept of "label" rather than an "attribute", and define that stigma as the simultaneous occurrence of five components: labeling, stereotyping, separation, loss of status, and discrimination, which are mutually reinforcing and exercised through power relations. In the case of study in this work, welfare stigma results from being labeled as a recipient of governmental social assistance monetary transfers, which involves a dual discrediting symbolization: (i) of being poor, but also (ii) of being dependant on welfare (Rainwater, 1982). Possible costs from welfare stigma may include loss of status, reduction in interpersonal trust and self-respect, social isolation, shame humiliation, and discrimination (Link and Phelan, 2001; Spicker, 2011).

Stigma (and welfare stigma) may undertake different subcategories which present subtle but substantial differences between one another. Walker (2014) proposes a clear distinction that suggests various interrelated subcategories of stigma: personal stigma, social stigma, and institutional stigma. On one hand, personal or internal stigma is strongly linked to an individual's own set of beliefs, under the concept that claiming social benefits is equivalent to personal failure (Baumberg, 2015). This emphasizes the fact that applying for the program, as well as being a recipient of the program, can be stigmatizing for the individual. Personal stigma relates to a personal feeling of own self-devalued identity, and can be related to the suffering of shame (Walker, 2014; Baumberg, 2015; Roelen, 2020).

Shame and stigma are intrinsically linked: there are deep feelings of shame involved (feelings of worthlessness, powerlessness, feeling small) if one is stigmatised, and both are concerned with personal and other's evaluations of self (Zavaleta, 2007). As described by Tangney and Tracy (2011), shame is a distressing and enduring emotion that can be regarded as a consequence of stigmatization and has been associated with distress. It involves a personal evaluation of the self, which implies failing according to one's own standards (Tracy and Robins, 2004). Walker (2014) argues that stigma can arise from the feeling of personal failure of poor people to meet the prevailing social norm in modern societies of being financially self-sufficient by participating in the workforce, especially if shamed individuals share this norm.

However, analyses of the self when stigma is involved might be affected by a larger extent by other's evaluations of ourselves rather than by our own, highlighting the role that social conditions, norms and values play in the process of stigmatization (Zavaleta, 2007). Therefore, stigmatization also involves the perception of identity devaluation coming from other members of society (Baumberg, 2015). This is what is referred to as social stigma, i.e., the process by which the rest of society labels, stereotypes, and distances beneficiaries based on discriminatory attitudes, thoughts, and actions on the part of the stigmatizing group, which, in this context, are linked to issues of merit, responsibility, and need (Walker, 2014; Lamont et al., 2016). Social stigma posses a divisive nature which differentiates between the acceptable and the unacceptable, between "them vs. us", from where its relation to humiliation comes from since being humiliated is often in relation to others (Zavaleta, 2007; Walker, 2014).

Humiliation is usually linked to an act or event that results in the feeling of being undermined by other people or institutions that are in a condition of relative power (Lindner, 2007; Mills and Zavaleta, 2015). It inherently interactional and has been related to various negative psychological consequences such as low self-esteem, discrimination and oppression (Hartling and Luchetta, 1999). Since poverty is believed to be evidence of laziness, poor people and welfare participants are thus questioned because of their believed lack of effort and willpower, making them a target for stigmatization in most modern societies, especially since they are believed to cost money to taxpayers (Nussbaum, 2009; Lamont, 2018). Institutions and other members of the community such as taxpayers, neighbors, employers, media, and society in general can foster stigma and humiliation on welfare participants (Walker, 2014; Baumberg, 2015).

Finally, institutional stigma is reflected in the design and implementation of public policies, as they may instigate stigma through targeting, indirect means-testing, the treatment of political operators, and the nature and delivery channel of the transfer (Roelen, 2020). Welfare stigma can be intentionally or unintentionally embodied in the design of welfare policies. In terms of the framework from Walker (2014), stigma related to take-up can be considered within this type of institutional stigma.

In the present study, only the first two subcategories of (welfare) stigma are considered: personal and social stigma. Operalization will be discussed further in Section 4.3, but each of these subcategories will follow the aforementioned framework and be measured with their closest associated feeling: shame and humiliation, respectively. Meanwhile, institutional stigma, which would mostly relate to the take-up discussion, cannot be analyzed under this setting since all individuals that conform the sample of the study have decided to apply to the programs and all have taken part through the institutional administrative process.

3.2 Previous literature

Previous empirical research has shown that stigma, along with other factors such as information asymmetry and transaction costs, can have a negative impact on the take-up rate of welfare policies in the United States, with varying grades of importance across programs (Stuber and Kronebusch, 2004; Stuber and Schlesinger, 2006; Kaye et al., 2013; Manchester and Mumford, 2010; Bhargava and Manoli, 2015; Brizmohun and Duffy, 2016).

Experimental studies using laboratory games have also provided evidence in favour of the welfare stigma hypothesis regarding incomplete take-up in social welfare policies. Friedrichsen et al. (2018) present novel causal evidence of social image concerns reducing the take-up of social benefits by manipulating the visibility of the take-up decision.[24] They run a Specifically, by making the claiming of social benefits to be public or private information they find that individuals are less likely to take-up a publicly visible transfer and try to avoid being inferred by other people as low-skilled, which they call ability stigma, and being willing to live off others, which they call free-rider stigma. Concordantly, participants are also more willing to vote for a private transfer rather a public one.

[24]The experiment was carried in Berlin, Germany, between 2015 and 2016 with online recruited participants.

Beyond the non take-up discussion where it is well documented that welfare stigma is a determinant of adherence to social assistance policies, the evidence regarding what occurs after beneficiaries enroll in these programs has been scarce. Most of the available evidence regarding this issue comes from the The Minimum Living Security System (MLSS) or "Dibao" program in China, which provides basic cash benefits for those living below the poverty line. Yet, findings on stigma have been inconclusive and dependent on the country's region, i.e. urban versus rural areas. For instance, Han and Gao (2020) find no evidence of stigma for rural areas in China, suggesting a positive effect of the program on the perceived social status, future prospects, and life satisfaction of beneficiaries. In contrast, they find evidence of stigma in urban areas, where beneficiaries tend to be less optimistic about their future and less happy compared to non-beneficiary peers. This finding is supported by qualitative research conducted by Zhang (2020), which reveals that female beneficiaries in urban areas face stigma, social exclusion, discrimination, and isolation.

Concomitantly, Qi and Wu (2018) analyze the effects of the Dibao program on well-being and psychological health and find that it has negative impacts on satisfaction, happiness, self-confidence, and interpersonal trust. It also increases the likelihood of feeling depressed, hopeless, helpless, and the perception of difficulty in accomplishing things. The authors argue that these stigma-related effects hinder the purpose of the poverty alleviation program. Overall, although with limited external validity since these studies are all from the same country, they suggest that the effects on stigma vary depending on the program's location and design. While some programs may have positive effects on the well-being and prospects of beneficiaries, others may create stigma-related effects that hinder their effectiveness.

Additional evidence of the presence of stigma in welfare programs is documented in (Celhay et al., 2022). They explore how under-reporting of recipiency of the Supplemental Nutrition Assistance Program (SNAP) in New York is related to stigma. The authors argue that widespread adherence to these programs should lead to less stigma, and they find evidence supporting their hypothesis by comparing recipiency reporting in household surveys versus administrative records. The intuition behind this indirect measure of welfare stigma is that if welfare stigma were to exist, then beneficiaries of SNAP in low-participation areas should be more likely to under-report their recipiency status, as it would be more shameful for them if that were to be known by others in their neighbouring areas. The authors find suggestive evidence of such hypothesis, although the results are not causal and rely on indirect measures of stigma.

In Latin America, Chong et al. (2009) find evidence that participation in social programs negatively affects interpersonal trust, which they attribute to the presence of stigma in these types of policies. In Uruguay, previous studies have estimated take-up rates to the aforementioned main non-contributory cash transfer programs in the country, such as PANES and AFAM-PE, which are similar with previous findings in the literature: non-take up is around one fifth of eligible population (Burdín and de Melo, 2009; Dean and Vigorito, 2015). Regarding take-up and stigma, Ghazarian (2020) examines the presence of stigma in AFAM-PE using a cross-sectional dataset. She finds that non take-up in AFAM-PE is correlated with socioeconomic variables, but not cultural (or stigma) variables. However, the author notes that this does not necessarily

rule out the presence of stigma in the adherence to this program, as the data used may not accurately capture this type of phenomenon.

Finally, there are several papers for Uruguay that study multiple impacts of the main non-contributory cash transfers (Manacorda et al., 2011; Amarante et al., 2016; Failache et al., 2016; Bérgolo and Galván, 2018; Bérgolo et al., 2016; Colombo et al., 2018; Rivero et al., 2019; Bérgolo and Cruces, 2021). Regarding TUS, the evidence is incipient. Colombo et al. (2018) use administrative records of the program and find that the policy has positive effects on the consumption of durable goods, formal employment and improvement in housing materials. In this same direction, Rivero et al. (2019) also find a positive effect on purchases of housing materials.

Regarding AFAM-PE, the evidence is much more extensive. The program evaluations show that the program does not significantly impact the medical control of children or the present or future fertility of mothers (Bérgolo et al., 2016; Rivero et al., 2019). Moreover, there is no effect on the nutritional status of children (Rivero et al., 2019). Regarding educational performance, while no substantive effects are found on school lag, a positive effect is found in terms of attendance at the secondary school level (Bérgolo et al., 2016; Rivero et al., 2019).

Considering potential undesired effects of the program, these evaluations indicate that the allowances do not significantly impact employment or the number of hours worked by adults in the household. However, there is evidence consistent with the hypothesis that the transfer has a positive impact on informal work, especially among female recipients (Failache et al., 2016; Bérgolo et al., 2016; Bérgolo and Galván, 2018; Rivero et al., 2019; Bérgolo and Cruces, 2021). If this information was widely known, social stigma towards those who receive transfers but are not employed should not be present since the beneficiaries do not decrease their workload but rather work to a greater extent in the informal sector. However, it is unlikely that this information is knwonn by the general population, so stigma may still be present.

Furthermore, these papers also analyze other effects related to beneficiaries' perceptions and opinions. For example, Rivero et al. (2019) finds that the program has no effect on the interpersonal trust. The authors also find no significant effects in a large set of responses on qualitative evaluations, except for a better evaluation of the government and the institutions that grant the allowances (MIDES and BPS). In addition, they study whether AFAM-PE affects the ability of feeling of appearing in public without discomfort, finding that beneficiaries report feeling more uncomfortable in public due to their appearance in comparison to non-eligible individuals. Finally, Bérgolo and Galván (2018) provide suggestive evidence that AFAM-PE increases women's empowerment within the household by improving their perceived ability to make decisions inside the family.

3.3 Hypotheses

Building on the institutional context details, the theoretical framework, and the empirical background, the hypotheses of this study are that: (i) both programs have welfare stigmatizing effects over their beneficiaries, and specifically personal stigmatizing effects, but that (ii) only TUS has social stigmatizing effects over its beneficiaries.

The justification for the first hypothesis comes from the fact that both programs are non-contributory and targeted to the poor, both of which are of central importance to the generation of stigma. In the economics literature the application of targeting instruments has been often contrasted with universal programs, arguing that the latter do not generate stigma (Mkandawire, 2005). While universal programs are conceived as a provision based on rights, targeted programs are perceived as charity intended for those who cannot economically sustain themselves (Walker, 2014). In this sense, while targeting aims to provide monetary assistance to those who truly need it, it also generates a precise division between the poor and non-poor, marking its recipients and reinforcing negative stereotypes associated with poverty, and questioning the deservingness of help by the poor (Devereux, 2016; Lamont et al., 2016). Anecdotal evidence of negative beliefs towards the poor and those on welfare that include these negative stereotypes about them can be found in Rivero (2020):

> *"I think that those who really need it are not helped. There are single mothers, who have four hundred children and get the allowance, but you work, you have your wage, and your daughter is not entitled to it. (...) They help the scoundrel, not the poor"*
>
> *"The State helps those who do nothing, not to the poor. You don't have to do formal work. If you are a single mother and you do not have a job (...) you go to MIDES, you apply, and they just give you money on the side"*
>
> *"MIDES gives them money not to work, they do not teach them how to work, or give them a plot to work the land (...). People are getting used to not work"*

Regarding the second hypothesis, there are some distinctions between both programs. In comparison to AFAM-PE, TUS is framed outside the traditional social security system, is targeted to more vulnerable households, it has a lesser coverage, and its usage is mainly restricted to food purchases. Furthermore, the delivery of the transfer for TUS the delivery is publicly visible since the food card is tagged by displaying that it belongs to a government social plan. Meanwhile, the delivery for AFAM-PE is not publicly visible since it uses a regular debit card. This comparison can be seen in Figure A3 and is a key difference between both programs. As was reviewed before, there is strong evidence towards people rejecting publicly visible transfers and trying to avoid signalling their recipiency status to others, either in a lab context (Friedrichsen et al., 2018) and in a real life cross-sectional study (Celhay et al., 2022). Consistent with this hypothesis, and as testimonies documented in Moreno et al. (2014) point out, TUS beneficiaries report feeling discriminated and treated unjustly in the groceries stores where they use the food card. Some translated testimonies are:

> *"The people [from the store] look at you as if you were stealing... as if he was discriminating you."*
>
> *"If the supermarket is very crowded, they say no, and you have to wait several days."*
>
> *"They make a face [grocery store employees], like, I don't know what... they tell you: oh no, are you going to pay with the food card?"*

"The guy [grocery store employee] looks at you as if you were stealing...as if he is discriminating against you."

"Discrimination is there... It's not that they treat you badly, but yes, you can feel it."

4 Data

4.1 The ESAFAM survey

The data used in this study comes from two unified sources of information: governmental administrative records and a follow-up field survey.[25] The administrative data was obtained from MIDES and BPS and contain crucial information about applicant (eligible and ineligible) households regarding program participation from 2008 to 2018: exact ICC score of each household, participation period and detailed baseline socioeconomic characteristics for all household individuals that were used for the calculation of the ICC.[26] The baseline period over which the empirical strategy is designed and from where the baseline administrative data is obtained lies between 2008 and 2010, at the early stages of the programs. Later on, the follow-up survey sample of AFAM-PE and TUS beneficiaries and non-beneficiaries was drawn from the universe of these administrative records. [27]

The *Encuesta de Seguimiento de AFAM* (ESAFAM) follow-up survey was designed specifically for the evaluation of the AFAM-PE program using a Regression Discontinuity Design.[28] Thus, the survey is overrepresented in the vicinity of the eligibility threshold, and its sample representative of those households eligible and non-eligible around this cut-off.[29] Consequently, the bandwidth selection was optimized within the survey sample to balance the need for a small interval to ensure comparability with the need for a sufficient sample size. (Amarante and Vigorito, 2011). In addition, the survey was extended in order to be able to also evaluate the TUS program. However, in this case the sample and bandwidth selections for TUS simply reflect the range of the vulnerability index where eligible TUS households are available on one side, and non-eligible TUS households, but eligible AFAM-PE households are available on the other. It is worth noting that given the survey design, either non-eligible or eligible TUS applicants available in the sample are also all eligible for AFAM-PE. Nonetheless, as noted before, this is

[25] Data access was possible due to the framework agreement for cooperation and technical assistance between the *Instituto de Economía* (IECON) and the *Ministerio de Desarrollo Social* (MIDES).

[26] It should be noted that when the program began, automatic applications were accepted for households already included in the PANES administrative records, both eligible and ineligible. Therefore, the AFAM-PE records contain PANES baseline data that was not updated for the computation of the ICC and date back to a date prior to the beginning of the program.

[27] Figure A5 outlines the main timeline of this study. It focuses on medium and long term impacts, measuring the outcomes between 6 and 10 years later, after participation in the program for the first time.

[28] The survey was designed jointly by the *Instituto de Economía* (IECON) and the *Instituto de Estadśtica* (IESTA) both from *Universidad de la República* (UDELAR).

[29] The sample was selected using a stratified random sampling design, where strata were defined by program status (eligible or ineligible) and region of residence (Montevideo or the rest of the country) (Bérgolo et al., 2016).

the case for 88% of households with children that are beneficiary of TUS in 2017, according to own estimates using the Official National Household Survey (ECH 2017).

Two different waves of the ESAFAM panel survey have been carried out. The first one was conducted between September 2011 and April 2012, carrying out 3,832 interviews in total. Meanwhile, the second wave was conducted between May 2016 and November 2018, managing to recover 1,734 households (45% of first wave's interviews). After taking into account missing variables, 995 and 604 individual responses are left for the AFAM-PE and TUS analysis, respectively. For AFAM-PE 66% of households are eligible for the program (660 out of 995 cases), while for TUS 51% of the sample is eligible (307 out of 604 observations). Estimates available in Rivero et al. (2019) suggest that attrition in the sample is at least uncorrelated with the ICC and observable variables.[30]

The surveys were conducted in-person by trained interviewers using a structured questionnaire and interviewees were intended to be the program applicants of the households. The survey contains a rich set of questions about different aspects of the household and its members, such as employment, income, education, expenditures and consumption, opinions, attitudes, and beliefs of the interviewees. Specifically, the Opinions and Expectations module of the survey contains a survey module developed by the Oxford Poverty and Human Development Initiative (OPHI) which includes internationally comparable measures of shame and humiliation feelings in the context of poverty, providing "*direct measures of people's experiences of shame, humiliation, stigma and discrimination* (Zavaleta, 2007, p. 3). Thus, this survey module will provide the main outcomes of interest of this study.

4.2 The OPHI survey module on shame and humiliation

The OPHI survey module on shame and humiliation was developed by (Zavaleta, 2007) to identify deprivation in a key "missing dimension" of poverty and well-being - the ability to go about without shame, which is often overlooked in the analyses of deprivation and well-being (Alkire, 2007).[31] To be able to avoid shame and humiliation can be closely linked to one's personal dignity and self-respect, all aspects that can influence people's capacity to carry on a life they value and have reasons to value (Alkire, 2007; Zavaleta, 2007). As mentioned before, both of these emotions are key to the understanding of welfare stigma since since they are amongst the possible ways this type of stigma can manifest (Spicker, 2011; Walker, 2014).

The module includes two scales that are meant to measure the aforementioned stigma related feelings and emotions: shame and humiliation. Although both are often used interchangeably as negative emotions that refers to the self, they are different from each other. While shame involves an individual, personal assessment, humiliation inherently requires interaction with others (Zavaleta, 2007; Mills and Zavaleta, 2015). The former relates to personal feeling of own self-devalued identity and could be related to personal stigma, the latter involves the perception

[30] For additional detailed information about the survey, please refer to Rivero et al. (2019).

[31] Alkire (2007) refers to "missing dimension" as those understudied features that are not only relevant to the analysis of deprivation but also important to those living in poverty themselves.

of identity devaluation coming from other people which relates to social stigma (Walker, 2014; Baumberg, 2015). This is why the present study shame is considered a measure of personal, while humiliation is considered a measure of social stigma.

Shame

On one hand, the module includes a set of items regarding shame proneness which refers to an the tendency of individuals to experience the emotion of shame in response to specific daily negative events (Tangney and Dearing, 2002; Tangney and Tracy, 2011). Shame is the result of a self-judgment that often leads to the conclusion of failure, involving the belief that one deserves to feel shame for failing their own personal standards (Hartling and Luchetta, 1999). Such as is the case with failing the aforementioned shared norm of financial self-sufficiency. The indicators conforming this scale come from the the Personal Feelings Questionnaire-2 (PFQ-2) developed by Harder and Zalma (1990), which is one of the one of the most established measures of shame proneness in the psychological field (Tangney and Dearing, 2002; Zavaleta, 2007).

In the shame proneness scale, the following question is asked for an extensive list of shame related emotions. "*For each of the following listed feelings please place a number from 1 to 4, reflecting how frequent the feeling is for you*". The list of items includes: feeling self-conscious, ridiculous, embarrassed, humiliated, laughable, stupid, childish, feelings of blushing, helpless or paralyzed, and feeling disgusting to others. The possible answers for each question range from 0 = Rarely or Never to 3 = Always or almost always. These questions try to elicit this long-term disposition to shame, giving a better assessment of their level of shame in their daily life experiences instead of capturing shameful isolated events (Zavaleta, 2007)

Humiliation

On the other hand, the module includes indicators regarding external humiliation which commonly refers to an act or event that results in feeling undermined by other people or institutions that are in a condition of a higher relative power (Lindner, 2007; Mills and Zavaleta, 2015). Since humiliation is inherently an interactional phenomena, these indicators highlight that an interaction between two parties is taking place (Zavaleta, 2007). The indicators included in this group were drawn from existing surveys that collect humiliation indicators in regard to external events, such as the European Social Survey, among others (Zavaleta, 2007).

In the external humiliation scale, the question asked is: *"Have you felt that you have been [...] during the last three months?"*. In this case, the list of items contains: feeling treated without respect, unfairly, and with discrimination. While the first two items are included to capture values affecting interactions among individuals, the third one attempts to collect actions or events in particular scenarios of everyday life (Zavaleta, 2007). The answering options in this case are the same as in the previous scale. This set also uses a frequency response scale instead of binary responses, although in this case the time-frame of the question dates back to three months, which leads the scale to capture mid-term experiences of humiliation and discrimination.

Descriptive statistics

Table A2 presents the questions and individual items corresponding to each dimension, including basic descriptive statistics. Figures A6, A7, and A8 display the proportion of responses for each individual item, differentiating between the AFAM-PE and TUS samples, and eligible and non-eligible groups within each program. The distributions of most items from the internal shame and external humiliations scales developed in Zavaleta (2007) are left-skewed, with the majority of responses being zero (around 80% on most items), and a low proportion of positive responses. This suggests that either (i) few people experience stigma, or (ii) the questions included in the scales may be eliciting intense stigma emotions, instead of light feelings, that are difficult for participants to reveal to an unfamiliar interviewer. Furthermore, this set of figures shows that respondents in the TUS sample, who are experiencing greater levels of deprivation, report a higher frequency in most items, consistent with what should be expected, theoretically (Walker, 2014) and empirically (Hojman and Álvaro Miranda, 2018). Also, those who are eligible in each program's sample provide a higher frequency of non-zero answers.

4.3 Outcome variables

In order to reduce data and test the composite indexes are created: one for shame proneness and another for external humiliation. To construct these linear indexes, it is necessary to determine which items best represent the latent constructs and how to weight them. Consequently, a Principal Component Analysis (PCA) is used to select the items for each scale. The PCA is described in more detail in Annex B. As a summary of this process, the set of items retained are those that explain the highest proportion of the variance of the principal component (62%) while keeping fairly high values of goodness-of-fit indexes such as the Kaiser-Meyer-Olkin (KMO) and Cronbach Alpha instruments. Also, following Yin and Etilé (2019) only those indicators that present loadings higher than 0.3 in the principal component are kept. Additionally, the item regarding the emotion of feeling stupid is dropped since it does not directly correlate to the context of poverty and welfare recipiency in which the questionnaire is being used. This is reflected by the fact that this variable's inter-item correlation is never higher than 0.14 with respect to the rest of the items. Moreover, after dropping this variable, both the average inter-item correlation and the proportion of the variance explained by the principal component (62%) increase. As a result, the indicators included in the internal shame proneness scale are Self-conscious, Ridiculous, Embarrassed, Humiliated, Laughable, and Helpless. As for the external humiliation scale, all three available items are used: Disrespect, Unfairness, and Discrimination.

On the other hand, three possible alternatives to weight the individual items can be considered: using the weights from the PCA (Filmer and Pritchett, 2001), computing the unweighted average of the items (Hoynes et al., 2016), or simply adding up the values of the items (Yin and Etilé, 2019). The additive indexes are chosen for the sake of simplicity, since the choice of weighting method does not qualitatively change the main results. Finally, to facilitate the comparison across the different indexes and programs under analysis, each of these two scales are standardized with mean equal to zero and standard deviation equal to a unit, in order to interpret the results in terms of standard deviations. In the end, two different indexes, Shame

z-score and Humiliation z-score are considered in the analysis.

Figure 1 shows the distribution of these indexes in their summarized version by plotting their kernel density functions for both AFAM-PE and TUS samples. Following their individual components, the composite shame and humiliation measures also present a left-skewed distribution which also is similar between both indexes. Table A3 shows the main descriptive statistics for both of these aggregated outcomes in the two different versions: with and without standardization. Two things can be observed by looking at the non-standardized version of the outcomes: (i) the average value of both shame and humiliation scales approximately doubles from those who are non-eligible for AFAM-PE beneficiaries to TUS eligible households, while those who are eligible for AFAM-PE but not for TUS report higher values on average than those who are above the ICC distribution but are still non-eligible for the TUS, although closer to this second cut-off. Moreover, a simple difference in means tests between each program suggest that those who are non-eligible for these policies report less shame and humiliation than those who are.

Finally, in order to assess some of the limitations of the main shame and humiliation scales and analyze whether the results are consistent with the presence of stigma or some other underlying factor, other subjective outcomes regarding perceptions and opinions of individuals are considered. These will be discussed further in Section 6.2. For now, only their questions are displayed and their main summary statistics are presented in Table A4.

- Perceived position in income distribution: "*Imagine a scale from 1 to 10, where in 1 are the poorest and in 10 the richest people, where do you place yourself?*".
- Recipients should feel ashamed of themselves: "*Do you agree with the statement that people who receive AFAM-PE (or TUS) should be ashamed of themselves?*". Possible answers are 0 (No) or 1 (Yes).
- Grade of support towards the program:
    - For AFAM-PE: "*Do you think that AFAM-PE benefits should be provided less in cash and a part should be given through a food card? (It is always the same money)*" The answer scale goes from 1 (strongly disagree) to 5 (strongly agree).
    - For TUS: ' *'Do you think that TUS is a...?*". The possible responses range from 1 (very bad benefit) to 5 (very good benefit).
- Ashamed of appearance: "*Have you thought about not attending or have you not attended a work, family or social event during the last month because you felt you did not have the clothes or appearance required for that venue?*". Possible answers are 0 (No) or 1 (Yes).
- Life satisfaction: "*On a scale of 1 to 10, where 1 is very dissatisfied and 10 is very satisfied, how satisfied are you in relation to to your life in general*"

5 Empirical strategy

5.1 Regression Discontinuity Design

As stated before, the objective of this study is to analyze the causal effect of cash transfers on the welfare stigma perceived by their beneficiaries. In the realm of social sciences, identifying a causal effect is a challenging task because social scientists cannot conduct experiments where the same individual is observed in two different scenarios: one in which they are exposed to an intervention (treatment), and another where they are not affected (control). This means that the direct counterfactual of an individual being exposed to a given treatment is not observable, and thus researchers need a source of exogenous variation to build a proper counterfactual for the treated group. The gold standard way to achieve this among the methods that allow for causal inference is to randomize the treatment between two groups and later compare their means regarding the outcomes. However, in this case that is not feasible since the programs under study, AFAM-PE and TUS, were not assigned randomly. However, they were assigned following an eligibility rule which can be exploited as a running variable to infer causality with a Regression Discontinuity Design (RDD).

In the absence of randomization, RDD is particularly an advantageous method to use when the intervention is assigned based on a continuous variable that has a clear and meaningful threshold. Such is the case in AFAM-PE and TUS, where their assignment criteria through the eligibility index, ICC, which determines the probability of participating in each program, allows to implement this quasi-experimental design. The main assumption behind this design is that eligible and ineligible households are very similar around the cut-off point, i.e. everything should be continuous around the this threshold except for the treatment and the outcome. This design is valid if the following requirements are met: the treatment has a clear-cut jump at the cut-off, the distribution of individuals is continuously dense around the threshold, and no other changes (such as other interventions) take place at the cut-off. Therefore, households within a close vicinity of this threshold will be very similar to each other. For a sufficiently small interval around the assignment threshold, if the above conditions are met, then the RDD should be as valid an alternative as a randomized experiment (Lee and Lemieux, 2009).

It is worth noting that from now on, rather than using the simple score of the ICC, the normalized score of the ICC will be used (denoted ICC*) to take into account that such index takes two different threshold values for different regions of Uruguay: one for Montevideo (capital city - and department - of the country) and another one for the rest of the country.[32] Thus, the eligibility threshold value is normalized to zero, such that $ICC^* = ICC - \overline{ICC}$, where $\overline{ICC}$ is the value of the corresponding cut-off point, which for AFAM-PE is approximately 0.2 and for TUS is approximately 0.6. Hence, each policy has its own threshold at zero in the ICC*.

Regarding bandwidth selection, as noted before, it differs according to each program. For AFAM-PE, since the survey was specifically designed to evaluate its impacts, bandwidth se-

[32] This difference in the values allows to take into account disparities regarding cost of life, infrastructure and access to different services between both regions.

lection was already optimized within the survey sample to balance the need for a small interval to ensure comparability with the need for a sufficient sample size to ensure statistical power (Amarante and Vigorito, 2011). Hence, the maximum possible bandwidth in the survey is already quite close to the AFAM-PE eligibility cut-off: it takes values of [-0.047; 0.073] in its normalized version. Meanwhile, for TUS, the bandwidth simply reflects the range of the vulnerability index where eligible TUS households are available on one side, and non-eligible TUS households but eligible AFAM-PE households are available on the other. This results in a wider bandwidth compared to AFAM-PE, and its values belong to the interval [-0.400 ; 0.400] considering its own normalized version. Both of these reasons are why it is not possible to go further away from these bandwidths. However, robustness checks of the main estimates are provided using narrower ranges in order to assess whether the results are sensible to such bandwidth.

5.2 RDD identification requirements

The first requirement for identification underlying this design implies that eligibility should be a strong determinant of actual treatment (Bernal and Peña, 2011), following:

$$\lim_{icc^* \uparrow \overline{ICC^*}} T_i(ICC^* = icc^*) \neq \lim_{icc^* \downarrow \overline{ICC^*}} T_i(ICC^* = icc^*) \tag{2}$$

The validity of this design relies on the existence of a "jump" in the probability of treatment assignment at the eligibility cut-off point, as indicated by Equation (2) where T_i denotes the actual treatment indicator. In other words, there should be a discontinuity in treatment assignment at the cut-off point. If this jump is perfect, meaning that eligibility strictly determines treatment, a Sharp RDD is appropriate. However, if the discontinuity is not perfect, and eligibility strongly influences the probability of treatment assignment but not deterministically, a Fuzzy RDD is needed. In the sharp specification, impact OLS estimates are obtained by regressing the outcome variable against the eligibility indicator alongside a function of the running variable. Meanwhile, the fuzzy specification requires to instrument the actual treatment with eligibility. Hence, impact 2SLS estimates are obtained by regressing the actual treatment against the eligibility indicator as the first stage regression.

Figure 2 shows first stage estimates of each program's effective participation against their respective ICC* score. The dots represent the share of program participants over ten bin intervals at each side of the respective cut-offs. The lines represent linear fits. Panel (a) displays results for the AFAM-PE sample. while Panel (b) displays results for the TUS sample for [-0.047; 0.073] and [-0.400 ; 0.400] bandwidths of each standardized running variable, respectively. A couple of points stand out. There is a jump in the probability of participating in both programs at the cut-off points, although these jumps are not sharp. The values of these jumps are around 0.75 and 0.63, for AFAM-PE and TUS respectively. Regarding the AFAM-PE program, the point cloud in the left side of the graph is lifted up due to the fact that some households that were originally not eligible for AFAM-PE in the baseline period ended up being eligible in subsequent years. Regarding the TUS program, the fact that the point cloud to the right sinks down may be explained

by different phenomena. While there may be some targeting issues, some non-compliance from applicant households may also be possible: either potentially eligible households did not apply for this second program or they declined participation. Nevertheless, the higher the ICC* score the more likely a household is effectively take part in the program, which is consistent with purposes of the policy design.

The second requirement for a valid RDD identification is non-manipulation of the eligibility instrument and implies that program applicants and governmental agents should not be aware of its computation and cutoff point. Otherwise, they could strategically change their answers in order to become eligible, and this could introduce selection bias. This would be evidenced by a bunching or clustering of ICC* scores next to the right of the cut-off points. Figure 3 shows graphically the verification of the assumption of no manipulation by displaying the McCrary (2008) density test, and there is no clear cut evidence for the index to have been manipulated in either program. Analytically, the estimated log difference in the heights at the cut-off is -0.186 with a standard error of 0.226, and 0.152 with a standard error of 0.248, for AFAM-PE and TUS, respectively. Therefore, the hypotheses of continuity is not rejected. The same conclusion is reached when considering other manipulation tests. The p-values of the tests for each program are 0.336 and 0.646 for the density test proposed by Cattaneo et al. (2018) and 0.999 for the density test proposed by Bugni and Canay (2021).

The third requirement for valid identification in a RDD is to ensure that ineligible households are a valid counterfactual for the eligible, i.e., both groups should be very similar to each other at each side of the eligibility cut-off to ensure that differences at the jumps are due to program participation. This requirement should be satisfied by fulfilling the local continuity requirement, which indicates that both the outcome variable and the covariates must be continuous around the eligibility threshold (Angrist and Pischke, 2009).[33] This means that these variables should not present discontinuities prior to treatment. Formally, this requirement for identification can written as follows:

$$\lim_{icc^* \uparrow \overline{ICC^*}} E[Y(0)/\ \overline{ICC^*} = icc^*] = \lim_{icc^* \downarrow \overline{ICC^*}} E[Y(0)/\ \overline{ICC^*} = icc^*] \tag{3}$$

$$\lim_{icc^* \uparrow \overline{ICC^*}} E[Y(1)/\ \overline{ICC^*} = icc^*] = \lim_{icc^* \downarrow \overline{ICC^*}} E[Y(1)/\ \overline{ICC^*} = icc^*] \tag{4}$$

Where Y(0) denotes an outcome in the absence of treatment and Y(1) denotes an outcome in the presence of treatment. This property can be tested by checking for imbalance of covariates at the cutoff through running RDD estimates for each covariate as the outcome of interests, following the estimation procedure described in Equation 8.[34] Table 1 displays these results for

[33] Here only the continuity for the covariates is shown since the set of outcome variables for the baseline period was not collected.

[34] In detail explanation of the estimation procedure is presented in the next paragraphs.

the following baseline covariates sex, age, region, years of education, log of income per capita, employment, and the amount of household size; and one contemporary covariate, ethnicity, due to its unavailability in the administrative data at the baseline. The table contains coefficients corresponding to the parameter of interest, which represents the effect of each program on each considered outcome. Columns (1) and (2) show impact estimates for the AFAM-PE program, while columns (3) and (4) show estimates for the TUS program. Furthermore, the last row of this table shows the value and statistical significance of an F-Statistic test that checks for all coefficients being jointly not different from zero for each specification. The result of this test cannot reject this null hypothesis, and overall these results do not allow to rule out for either program the null that those who are eligible and not eligible are not significantly different between each other. This presents a challenge for the identification of the causal impacts of the program with this design.[35]

However, in both programs these overall differences seem to be driven by one specific variable. Looking at Columns (2) and (4) which contain quadratic fit specifications, only one variable remains statistically significant in each program: Sex in AFAM-PE and Region in TUS. This can be seen in Figures 4 and 5, where the RDD graphs show sparse scatter plots of the covariates as outcomes with no clear jumps at the cut-offs, with the exception of these two variables. This also appears to be backed up by the fact that if each variable is taken out from the F-Test, then the test becomes non-significant statistically, taking values of 6.41 (p-value = 0.49) and 8.31 (p-value = 0.31) in the AFAM-PE and TUS samples, respectively. Therefore, one way of testing whether these differences cause significant changes in the main results would be to re-run these estimates using only individuals belonging to one of the subgroups. e.g. only using woman in AFAM-PE and only using the rest of the country region in TUS, and afterwards compare if these estimates are similar to the ones obtained with the main specifications.[36] If the results hold, then these discontinuities could be less worrisome for the interpretation of the main results. Indeed, the results seem to be robust for these sub-sample analyses as shown in Section 6.3. It is worth noting that for the subgroup sub-sample to be valid, the rest of the RDD requirements should also hold. This is shown in Figures A9 and A10. Both, the jumps at the cut-offs and the McCrary density tests hold for each program. Also, Table 2, accompanied by Figures A11 and A12, display the alternative balance tests. Now, considering the quadratic specifications, none of the covariates considered has a coefficient with a statistically significant level over 95%, while the F-Statistics for the joint test are not statistically significant anymore.

[35]These discontinuities are in line with what previous impact evaluations of the programs which use the same sources of information find (Bérgolo et al., 2016; Failache et al., 2016; Bérgolo and Galván, 2018; Rivero et al., 2019). In studies with analysis at the individual level these discontinuities often appear, while in studies with analysis at the household level these discontinuities often disappear, which would suggest that these discontinuities are not strong.

[36]I choose these sub-samples due to sample size. For AFAM-PE, approximately 90% of the sample are woman, while for TUS almost 74% of households do not reside in the capital city.

5.3 RDD econometric specifications

Based on the previous analysis, impact estimates of each program are achieved by comparing individuals right above and below the thresholds of the standardized eligibility index. As a benchmark measure which serves as an additional robustness check, an OLS parametric estimation of a Sharp RDD specification is performed, which would have solely been used if the jumps at the cut-offs were deterministic. Following Lee and Lemieux (2009), such reduced form estimation approach is done through:

$$Y_i = \alpha_1 + \beta_1 D_i + \gamma_1 f(ICC_i^*) + \gamma_1' D_i \times f(ICC_i^*) + \delta_1 X_i + \epsilon_i \tag{5}$$

Where Y_i is the outcome variable considered, $f(ICC_i^*)$ is a polynomial function of the normalized eligibility score, X_i is a vector of control variables, ϵ_i is the error term, and D_i is the eligibility indicator, which takes a value of one if the score of the ICC^* is higher or equal to zero and zero otherwise, such as:

$$D_i = \begin{cases} 1 & \text{si } ICC^* \geq 0 \\ 0 & \text{si } ICC^* < 0 \end{cases} \tag{6}$$

However, the jumps at the cut-offs are not sharp, but fuzzy: these take values of approximately 0.75 and 0.63, for AFAM-PE and TUS respectively. Therefore, in second place a Fuzzy RDD specification is estimated by using 2SLS, which basically scales the results obtained and adjusts them for the aforementioned non-compliance around the thresholds. In this approach, effective participation in the program is instrumented with the probability of effectively participating, i.e., being eligible to the program. Thus, this first stage can be written as:

$$T_i = \alpha_0 + \lambda_0 D_i + \gamma_0 g(ICC_i^*) + \gamma_0' D_i \times g(ICC_i^*) + \delta_0 X_i + \eta_i \tag{7}$$

Where T_i is the treatment indicator to be instrumented and the rest follows the same structure as Equation 5. First stage estimates are also obtained for the polynomial interactions between the eligibility indicator and the polynomial functions of the ICC* score. Then, estimates of the second stage are obtained using the following equation:

$$Y_i = \alpha_2 + \beta_2 \widehat{T}_i + \gamma_2 h(ICC_i^*) + \gamma_2' \widehat{T_i \times h(ICC_i^*)} + \delta_2 X_i + \mu_i \tag{8}$$

Where $\widehat{T}_i$ is the effective treatment variable that is instrumented by the eligibility indicator D_i, while the the rest follows the same structure as Equations 5 and 7.[37]

[37] In this setting, the IV estimator obtained through the second stage corresponds to the Wald estimator, which could just be computed by diving the β_1 obtained via the naive OLS by the λ_0 estimated through the first stage of the 2SLS estimation.

The parameter of interest in both Sharp and Fuzzy specifications is β, which represents the effect of the policy on the outcome variables to be considered. Four model specifications that vary according to the polynomial degree of the ICC* function (linear, linear with controls, quadratic, and quadratic with control) are tested for robustness purposes. The control variables included in the main regressions are sex (woman = 1), age, region (Montevideo = 1), and ethnicity (white = 1). In all specifications the errors are clustered by the ICC* score.

Finally, to further assess robustness local polynomial semi-parametric estimates are implemented, following Cattaneo et al. (2018). These estimates have the advantage of computing bias-corrected estimates and robust standard errors, following Calonico et al. (2014). In addition, this approach can also provide calculations of optimal bandwidths of the running variable for specific outcomes. However, given the constrains of: (i) being already close to the thresholds, and (ii) sample size, these estimates are computed with the bandwidths already fixed. The second stage estimation follows the next equation:

$$Y_i = \beta_3 \mathrm{I}[\widehat{ICC_i^* \geq 0}] + \rho_3 + \rho_3' ICC_i^* + \rho''_3 ICC_i^* \widehat{\times \mathrm{I}[ICC_i^* \geq 0]} + \delta_3 X_i + _i \tag{9}$$

Where $\mathrm{I}[\widehat{ICC_i^* \geq 0}]$ is the first-stage predicted indicator for eligibility and again β_3 represents the effect of crossing the cutoff on outcome Y_i.

Before continuing to the next section, additional details about the internal validity of the research design must be discussed. Until now, the internal validity discussion has focused on meeting the RDD requirements' criteria. However, since the preferred specification involves the Fuzzy set up, the internal validity also depends on the adequacy of the IV estimator. For the IV approach to be valid, two conditions must be met in this setting. First, the instrument must strongly affect the probability of treatment. Second, the exclusion restriction, i.e. the instrument should affect the outcome exclusively via its effect on the treatment, must be satisfied. As mentioned before, both first stages have highly strong jumps at the cut-off and also F-Test values. For AFAM-PE, the F-Test for the first stage for the treatment equals 366.67, while for the first stage of the interaction between treatment and the running variable, the F-Test takes a value of 270.05. For TUS, these statistics are lower but sufficiently strong still, 75.13 and 76.51, respectively. These first stages values are quite strong and satisfy the minimum heuristic rules from Stock and Yogo (2005). Moreover, these first stage regressions satisfy the testing for weak instruments and over-identification tests. Regarding the exclusion restriction, it is unlikely to be violated given that the construction, weights and cut-offs of the running variable were not publicly known by either applicants of government interviewers, and as it's seen in Figure 3 it has not been manipulated according to the data.

Regarding the external validity, in a RDD setting it is essential to note that the programs' impact estimates are only local and cannot be generalized to the full universe of the beneficiary

$$\beta_2 = \frac{E[Y|D=1] - E[Y|D=0]}{E[T|D=1] - E[T|D=0]} \approx \frac{\beta_1}{\lambda_0}$$

population, resulting in limited external validity for the design (Bernal and Peña, 2011). Consequently, the main estimated effects in this study correspond to local average treatment effects (LATEs) and cannot be extrapolated to the entire eligible population of the programs. An additional implication of this feature implies the following. Since most households who receive TUS also receive AFAM-PE in this study, the impact of TUS would actually be "on top" of the impact of AFAM-PE. However, due to these separate effects being LATE, they cannot be added up together to compute a common effect of TUS plus AFAM-PE. The AFAM-PE effect may not be homogeneous across across the ICC* distribution.

6 Results

6.1 RDD impact estimates on shame and humiliation

This section begins by presenting the main estimates of the effect of the AFAM-PE and TUS programs on the perceived shame and humiliation feelings of their recipients. Figure 6 presents graphical evidence of these effects, following the OLS specification of Equation 5. It shows Sharp RDD impact estimates for the main outcome z-scores indices: shame and humiliation for both programs. The dots represent the mean value of the shame and humiliation z-score indexes over ten bin intervals at each side of the ICC* cut-off, which is outlined with the vertical bar at zero. The lines represent linear and quadratic fits of the standardized eligibility score, respectively. Panels (a) and (c) show estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while Panels (b) and (d) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. These are the highest bandwidths available in each survey sample.

On one hand, AFAM-PE and TUS beneficiaries just above the eligibility thresholds report higher proneness to feel ashamed versus non-beneficiaries just below the thresholds as shown in Panels (a) and (b). Graphically, the effect for TUS in this case looks cleaner and of higher magnitude. On the other hand, while the exhibit for AFAM-PE does not present significant effects of the program on the proneness to feel humiliated, the exhibit for TUS does, showing that TUS beneficiaries report feeling more humiliated than their non-beneficiary counterparts. Furthermore, it can be observed that the linear fits better approximate the scatter clouds, while quadratic fits get really "jumpy" around the thresholds. For this reason, the linear estimates (with controls) will be the preferred specification from now onwards. Table 3 shows these results in more detail in regression form (Panel A). Moreover, it also shows the results following the Fuzzy RDD specified in Equation 8, which adjusts the estimates for the non-compliance around the thresholds using 2SLS (Panel B). It displays coefficients corresponding to the parameter of interest β that represent the effect of each program on the considered outcomes, i.e. the jumps at the cut-off. Columns (1) to (4) show impact estimates for AFAM-PE while columns (5) to (8) show estimates for TUS. A number of aspects stand out.

First, both programs have a significant effect at the 95% confidence level on the shame z-score index. Following the sharp linear specification with controls of Columns (1) and (3), participa-

tion in AFAM-PE and TUS, on average, increases the reported shame proneness of individuals by 0.34 and 0.44 standard deviations, respectively. Adjusting these coefficients for the evidenced non-compliance around the thresholds scales up these estimates to 0.46 and 0.67, respectively, both still statistically significant at the 95% level. These magnitudes are not negligible and would suggest both programs increase their beneficiaries' perceived personal stigma. For instance, the effect of AFAM-PE on shame would increase the mean for the non-eligible group by 99% from 1.062 to 2.112, moving an individual from the 53th to the 70th percentile in the shame scale distribution in the AFAM-PE sample. Meanwhile, the effect of TUS on shame would increase the non-eligible mean by 150% from 1.333 to 3.337, moving an individual from the 48th to the 75th percentile in the shame scale distribution in the TUS sample, as can be seen in Figure 1.

Second, AFAM-PE has no statistically significant effect on the external humiliation reported by the individuals. Meanwhile TUS has a positive effect on this index. TUS increases the proneness of feeling humiliation by between 0.42 and 0.65 standard deviations, a magnitude of similar value to that of the shame proneness index. This result is statistically significant at the 99% level. In this case, the effect of TUS on humiliation would increase the non-eligible mean by 115% from 0.960 to 2.072, moving an individual from the 48th to the 65th percentile in the humiliation scale distribution in the TUS sample, as can be observed in Figure 1.

Third, the main results are quite robust across specifications, even though estimates for TUS increase in their magnitude but decrease in their statistical significance when considering quadratic specifications. However, this may be due to the fact that survey sample size for this program is quite small, 560 observations, which may not be enough for a quadratic fit to properly adjust to the scatter points. Something similar occurs for AFAM-PE. This reinforces the preference for considering the linear estimates the preferred econometric specification. In addition, concerns could be raised about the construction method chosen for the main outcomes. To assess these, estimates were replicated with different aggregation alternatives: using the weights from the PCA to compute the weighted average, rotating such weights, and simply adding the items without standardization. Results turn out to be robust to the method of construction and are similar in terms of standard deviations magnitudes and statistical significance.[38]

Fourth, while the effects on shame are similar across programs, albeit the effect of TUS is higher, the difference in the reported feeling of humiliation is quite striking, although it goes in line with the previously raised hypothesis. Both programs share similar institutional features and surrounding public discourses, that were detailed in previous sections and might foster internal stigmatizing feelings such as shame. However, while not the only one, a key difference between both is that the TUS food card makes beneficiaries more visible to others in the public sphere. This could potentially make them more likely to be exposed to discrimination situations where external stigmatizing feelings, such as humiliation, may arise more frequently, which would go in the same line as the qualitative evidence of Section 3.3 and the literature reviewed regarding welfare stigma and the visible nature of recipiency status (Friedrichsen et al., 2018; Celhay et al., 2022). However, this explanation cannot be directly tested so it remains a hypothesis that should guide future research and not be considered a concrete finding of this study.

[38] These estimates can be provided upon request from the reader.

Although such hypothesis cannot be directly tested, additional explorations can be made in order to further asses its likelihood. If TUS households were indeed more likely to be exposed to this kind of humiliating situations due to the food card, individuals just above the TUS thresholds would report higher frequencies for external feelings rather than internal ones. This can be further analyzed by looking at the individual items that comprise the shame proneness scale. For instance, while the feelings of self-conscious and embarrassed may be linked to internal self-conscious emotions (Tracy and Robins, 2004; Tangney and Tracy, 2011), feeling humiliated might be emotions of a more external nature (Zavaleta, 2007; Mills and Zavaleta, 2015).

Table 4 shows RDD impact estimates for the individual items comprising both shame and humiliation scales, following the fuzzy linear specification of Equation 8.[39] Outcomes are again expressed in standard deviations in order to facilitate magnitude comparisons with the main variables. It additionally displays Romano-Wolf multiple hypotheses adjusted p-values in bold, since this exercise analyzes multiple outcomes at the same time, which could potentially lead to find significant results by chance (Clarke et al., 2020). Estimates and levels of significance vary across items and programs. In first place it is worth noting that statistically significant magnitudes are similar to those of the main results in terms of standard deviations, which provides additional reassurance about the estimates using the composite outcomes. In second place, TUS has positive significant effects on feeling humiliated, laughable and helpless or paralyzed, all items that allude to scenarios of external nature.[40] Meanwhile, AFAM-PE's effect on shame is driven by the item of feeling self-conscious, a deep internal feeling, since it is the only statistically significant individual result of this program.

Given the last result, following Hoynes et al. (2016), a drop most significant item out test is implemented in order to check whether the main result regarding shame is sensible to the exclusion of its most relevant item, which differs in each program. Thus, estimates are again computed excluding the item self-conscious in the case of AFAM-PE and the item helpless for the TUS analysis. For TUS the effect is quite robust (0.61 point estimate with significance at the 95% level). For AFAM-PE, the effect decreases in magnitude from 0.46 to 0.37, and loses some statistical significance, which is now at a 90% confidence level (p-value = 0.088). This was expected, since the only individual item that is significant in the AFAM-PE program analysis is removed from the scale. Nevertheless, the statistical significance of the AFAM-PE effect on the reported shame proneness is not vanished, which means that the other items still contribute to the overall scale to some extent, while the self-conscious item strengthens the composite index. This same exercise can also be applied with the items that were left out of the main shame proneness scale, for which Table A5 shows their results. In both cases, the most relevant item is feeling disgusting. The same result is reached: while for TUS the effect is still quite robust (0.64 point estimate significant at the 95% level), for AFAM-PE the effect is not as robust although it retains some statistical significance (0.40 point estimate and p-value = 0.063).[41]

[39] Figures A13 and A14 show the same analysis in graphical form for AFAM-PE and TUS, respectively. Additionally, Table A5 reports results for those items that were excluded from the main scale.

[40] Unsurprisingly, for the individual items of the humiliation scale, TUS has statistically significant effects on feeling treated unfairly and with discrimination, while AFAM-PE has none.

[41] Similar conclusions are reached if we also add the item referring to feeling stupid: 0.59 point estimate and

6.2 Further analysis of other perceptions

This subsection discusses whether the shame and humiliation effects previously found are indeed due to stigma or if there are other possible reasons other than participation in the programs that could make beneficiaries feel more ashamed and humiliated. In first place, one possible way these difference may not be explained by welfare stigma is if being non-eligible to the program made non-eligible respondents feel less poor. Since their participation was rejected, they could feel they do not need it, and thus that they are not as poor as those who were admitted. Subsequently, if they felt less poor, they could feel less shame or humiliation, and program participation would not be the only cause of the difference between eligible and non-eligible groups. The question available in the ESAFAM survey about the perceived position of individuals in the income distribution can be used to assess this issue. Panel A of Table 5 displays estimates for this variable. No statistically significant differences can be found between eligible and non-eligible from either program, discarding this alternative explanation.

In second place, it could also be plausible that non-eligibility makes people who were not accepted be more proud for not being considered poor enough to need governmental assistance, thus leading to feel less shame and humiliation. Another plausible interpretation would be that non-eligible resent those eligible because the former did not get into the program while the latter did. A possible way to check this is testing whether non-beneficiaries are more likely to believe that AFAM-PE or TUS beneficiaries should be ashamed of receiving benefits from these programs. A reasonable assumption is that they did not think that was the case since they all applied to these programs. Therefore, if differences were to exist in this regard, they should be explained by eligibility, and therefore eligibility would not exclusively affect shame and humiliation through program participation but also through non-eligible pride or resentment towards those eligible. The question about whether recipients should feel ashamed of themselves allows to elucidate the matter in question.[42] Panel A of Table 5 reports results for this outcome. Again, eligible and non-eligible from either program display differences in their answers to this question, which would disregard this alternative mechanism.

In third place, it's not only possible that non-eligible individuals resent beneficiaries, but also that they resent the program itself due to their rejection which could make them feel more proud, and report less shame and humiliation for other reasons program eligibility directly. Do non-eligible individuals report less support towards the policies? Panel A of Table 5 shows results for two available questions in the ESAFAM survey that directly or indirectly reference the grade of support towards these policies. No significant differences can be found between eligible and non-eligible responses one more time, which leads to reject this alternative reason for the effects on shame and humiliation.

In addition, shame and humiliation could not only be caused by participating in the transfer programs but also from other sources arising in everyday life. One of them which has a question about it available in the ESAFAM survey is presenting oneself in public without shame coming

0.021 p-value for TUS, and 0.37 point estimate and 0.081 p-value for AFAM-PE.

[42] Those being assessed in the AFAM-PE (TUS) analysis were asked about AFAM-PE (TUS).

from one's appearance. Panel A of Table 5 shows estimates regarding this dimension. The result for AFAM-PE is quite noisy since the impact estimate is statistically non-significant with the linear specification but statistically significant at the 95% level with the quadratic one. A possible hypothesis on why this could the case is that these individuals are not extremely poor, so they might frequently visit places where there is a higher likelihood of them being judged upon their looks. Nonetheless, this result is not robust across specifications. On the other hand, the result for TUS is close to zero across every specification. On average, TUS beneficiaries are not reporting feeling more ashamed than their counterparts because of their clothes or appearance in public events. Something else may as well be explaining their higher perceived shame proneness and external humiliation, which can be their visible welfare participation.

Lastly, shame and humiliation are important predictors of life satisfaction as is reported in Hojman and Álvaro Miranda (2018). More importantly, previous studies have linked welfare stigma as an explanation of why social assistance programs may reduce the subjective well-being of their beneficiaries, usually measured with life satisfaction scales (Gao and Zhai, 2017; Qi and Wu, 2018). Therefore, if stigma is indeed prevalent it could affect the entirety of individual's well-being. Do the shame and humiliation effects translate to a decrease in the subjective well-being of beneficiaries? Suggestive answers to these questions can be provided by leveraging the life satisfaction question within the ESAFAM survey. Panel B of Table 5 reports results for this question. Considering the original version, no significant differences are found between eligible and non-eligible in either program across any specification. Two aspects are worth mentioning. First, although it's not possible to rule out that coefficients are distinct from zero, these are all negative, hinting that if there was any significant effect it would be negative. Second, magnitudes are substantially higher for TUS than for AFAM-PE, which would be consistent with the finding of such program having deeper stigmatizing effects. Does this difference accentuate in more extreme cases?

A simple way to further check the previous question is splitting those individuals who report being satisfied with life above or equal the variable median (which is 7 in AFAM-PE and 6 in TUS), and those who report their answers below the median in two groups. Following such alternative, while eligible and non-eligible of AFAM-PE do not present differences in life satisfaction either, TUS eligible individuals are 25%–29% more likely to be below the median of satisfaction with life responses, considering the preferred specification. These effects are significant at the 95% level. This last set of results is not robust enough to draw any conclusions with respect to whether program participation (and feeling shame and humiliation) leads to feeling less satisfied with life. Nonetheless, there is a non-negligible difference between the both programs over this dimension in the second exercise. A plausible hypothesis could be that this difference is explained by the different stigma levels on these programs. However, going beyond this assumption and proving this relation escapes the scope of this study.

6.3 Robustness checks

This subsection presents a set of robustness checks carried out to further asses the internal validity of the RDD design. As already mentioned, a potential threat to identification lies in the

fact that treatment and control groups present a significant discontinuity on a key covariate in each case: while in the case of AFAM-PE the variable in question is the sex indicator, in the case of TUS the problematic variable is the region of residency of the household. Such a lack of homogeneity with respect to the comparison groups may affect the causal identification of the RDD, given that the results found could be driven by the discontinuity in sex (region) of the individuals rather than participation in the program. For this reason, the main results are re-estimated but restricting the sample to women (non-capital residents) in order to compare two more homogeneous groups, women with respect to women, and Montevideo (capital) with Interior (non-capital), and check if such results still hold. Figure 7 compares the parameters of interest β obtained from the full sample and sub-sample Equation 8 linear estimates including control variables. The results turn out to be quite similar between both alternatives for each of the variables and programs considered. For TUS, the results fairly hold, although considering the humiliation outcome, its statistical significance slightly decreases compared to the main estimates. However, it is worth noting that the sample size for this analysis is even lower, so there could be an issue of power in regard to these regressions. For AFAM-PE, the effect on humiliation remains statistically non-significant, while the effect on shame proneness slightly increases. All in all, results seem to be robust to this discontinuity checks, strengthening the validity of the main estimates.

Next, a series of robustness tests that follow the conventional RDD literature are presented. First, placebo tests are performed to check whether the effects found actually arise at the programs' eligibility thresholds or whether they also occur at other points of the index close to the cut-off point. If the former occurs, the causal effect of the policy would be adequately identified, while if the latter occurs, it opens the possibility that the effect found could be due to another factor in addition to the program, hence, invalidating the causal interpretation of the results. Along these lines, estimates for the main outcome variables are computed again but now these consider varying cut-off values. The following alternative cut-offs are considered for AFAM-PE range from [-0.020, -0.019, ..., -0.011 -0.010] to [0.01, 0.011. ... 0.019, 0.020], while for TUS these values range from [-0.20, -0.19, ..., -0.11 -0.10] to [0.10, 0.11 ... 0.19, 0.20]. Figure 8 shows estimates for the main outcome variables according to these alternative cut-offs, using the preferred specification: a linear Fuzzy RDD with controls, as described in Equation 8. For none of the alternative cut-offs other than zero the parameters are significantly different from zero. This result reinforces the causal interpretation of the main estimates, since the only significant change in the shame and humiliation reported by the individuals come exclusively at the eligibility thresholds of the policies.

Another canonical robustness check for the validity of the RDD design consists of moving the selected bandwidth across the running variable and see whether this affects the results. As previously mentioned, the survey sample does not allow to go further away from the policies' thresholds, but nonetheless it is possible to get closer to them. Table 6 displays the results from this test. It presents alternative linear fuzzy estimates considering shorter bandwidths for the running variable. Panel A shows results for a trimmed running variable with 0.01 (0.1) cuts from each tail for the AFAM-PE (TUS) sample. Meanwhile, Panel B shows results for 0.015 (0.15) cuts and Panel C shows results for 0.02 (0.2) for the AFAM-PE (TUS) sample. The main

results not only hold, but the table also shows that the narrower the bandwidth the higher the estimates obtained. All new estimates are also statistically significant at the 95% confidence level. For AFAM-PE, the effect on shame increases from 0.46 in the main estimates to 0.57 and 0.55 for a 0.01 cut and 0.02 cut bandwidth, respectively. For TUS, the effect on shame remains the same for a 0.1 cut bandwidth, and increases from 0.67 in the main estimates to 0.88 a for a 0.2 cut bandwidth. The biggest increase happens to the TUS effect on humiliation, which while it decreases from 0.65 to 0.61 for a 0.1 cut bandwidth, it increases to 0.93 for a 0.2 cut, almost a one standard deviation effect, which is considerably high. This set of results is consistent with the RDD set-up: the closer to the cut-off the bigger the effect of the policy. This is further assessed in Figure A15 which shows coefficients of interest for additional cut-off trimmings.

In addition, donut RDD estimates are performed in order to test whether observations that are immediately around the thresholds introduce noise into the estimation. For example, in the different panels of Figure 6 it can be seen that some of the bins next to the cut-offs don't necessary follow the scatter clouds and may be driving the estimates. Instead of trimming the tails of the running variable, a Donut RDD consists in trimming the center of this variable, by leaving out of the regression those observations right next to the threshold. Table 7 presents this robustness check. Panel A shows results for a trimmed running variable with 0.001 (0.01) cuts right next to the threshold of AFAM-PE (TUS). Meanwhile, Panel B shows results for 0.005 (0.05) cuts for the AFAM-PE (TUS) sample. As it was the case before, these results show that the estimated coefficients either remain the same or even increase, still at a 95% confidence level. Observations right below and above the cut-off, if noisy, make the main estimates a sort of lower bound for the significant effects of these policies. Again, results turn out to be robust.[43]

Finally, RDD estimates can also be computed using a non-parametric approach that exploits local polynomial estimates to better approximate the effect at the cut-off and give more weight to those observations nearest to it. To check whether parametric and non-parametric estimates are similar to each other, the analysis is rerun following Equation 9 based on Cattaneo et al. (2018). Table 8 presents these non-parametric results for the preferred fuzzy linear specification. Bias-corrected estimates of β and robust standard error are reported, following Calonico et al. (2014). Panel A shows results using a uniform kernel specification, while Panel B shows results with a triangular kernel, and Panel C with a epanechnikov kernel. Although statistical significance slightly varies across different kernels (p-values are around 0.05), the magnitude of the effects of the policies increase in all cases. AFAM-PE impact estimates on the reported shame and humiliation go from 0.46 and 0.14 to 0.64 and 0.38, respectively. For the shame proneness scale, that's 40% increase. Meanwhile, TUS impact estimates on these two outcomes increase from 0.67 and 0.65 to as high as 0.8 and 0.83, which represent jumps of 19% and 22%, respectively. These results are consistent with those that use narrower bandwidths from Table 6. The stronger the weight assigned to those observations closer to the cut-offs, the higher the magnitude of the impacts of the policies. This provides additional support for the validity of the RDD results, and robustness to the preferred more cautious linear parametric estimates.

[43] Figure A16 shows the same analysis in graphical form for the 0.005 and 0.05 cuts.

7 Final remarks

This thesis studied the welfare stigma effects of cash transfers, analyzing the case of the main cash transfer programs in Uruguay: AFAM-PE and TUS. The main results suggest that welfare participation increases the shame and humiliation feelings of their recipients compared to a group of similar non-recipients who also applied to the programs. However, these effects vary between both policies. There is suggestive evidence that AFAM-PE only increases the feeling of internal shame, while it has no effect on the feeling of external humiliation. AFAM-PE participation increases the score of the shame proneness index between 0.34 and 0.46 standard deviations. Stronger evidence indicates that TUS increases both, shame and humiliation. TUS beneficiaries report on average higher feelings of shame and humiliation by between 0.44 and 0.67, and 0.42 and 0.65 standard deviations, respectively. These results imply that the effect of AFAM-PE on shame would increase the mean for the non-eligible group by 99% from 1.062 to 2.112, moving an individual from the 53th to the 70th percentile in the shame scale distribution in the AFAM-PE sample. Meanwhile, the effect of TUS on shame (humiliation) would increase the non-eligible mean by 150% (115%) from 1.333 (0.960) to 3.337 (2.072), moving an individual from the 48th (48th) to the 75th (65th) percentile in the shame (humiliation) scale distribution in the TUS sample. These findings are also quite robust to a series of different validity checks.

Overall, these results provide support for the two hypotheses raised in this study. First, both programs have welfare stigmatizing effects over their beneficiaries and specifically personal stigmatizing effects, as the two increase the internal feeling of shame proneness perceived by their beneficiaries. In addition, this effect is stronger for TUS than for AFAM-PE, which highlights the higher stigmatizing capacity of such program. This interpretation gains ground when considering that only the TUS program has social stigmatizing effects over its beneficiaries, which was the second hypothesis of the present work. Only TUS raises the external feeling of humiliation reported by their recipients, while AFAM-PE has a statistically non-significant effect. This set of results suggest the presence of mild welfare stigma in AFAM-PE, mainly through personal stigma, and moderate to high welfare stigma in TUS, through both personal and stigma channels and also with higher intensity.

Disentangling which mechanisms could be behind these effects is out of the feasible scope of this study, although some hints for future research can be sketched out by comparing both programs. A possible set of explanations for the stigmatizing effects and their differences across programs could rely on program design. For instance, differences in the impact's magnitude on the internal feeling of shame could lie in the fact that AFAM-PE is less stigmatizing because the program has a longer trajectory and a vaster coverage than TUS, which has been framed outside the traditional long-lasting social security system of the country. Regarding the higher external humiliation effect of TUS, a possible explanation for this difference could arise from the fact that the TUS food card makes beneficiaries more visible to others in the public sphere. While the delivery of AFAM-PE is not publicly visible by others, the TUS magnetic card is tagged by displaying that it belongs to a government social plan. This could potentially make them more likely to be exposed to discrimination situations where external stigmatizing feelings, such as humiliation, may arise more frequently, for instance when using the card for

shopping in grocery stores, as told by some testimonies from beneficiaries presented in Section 3.3. Furthermore, this would go in line with literature reviewed regarding welfare stigma and the visible nature of recipiency status (Friedrichsen et al., 2018; Celhay et al., 2022). Nonetheless, this explanation cannot be directly tested so it remains a hypothesis that should guide future research and not be considered a concrete finding of this study.

Future research should try to elucidate which specific channels could be driving the different findings of this study. One direct venue for future research is to provide clear-cut evidence of whether the TUS food card's visibility feature is such a mechanism of transmission of stigma. A possible way to assess this would be to select a random sample of TUS beneficiaries and offer them a magnetic card that looks like a common debit card instead of the current TUS. In addition, theoretical economic models are needed in order to better understand how shame and stigma originate, how they can be shaped by social assistance and what is the net social welfare costs and benefits of these programs under the presence of shame and humiliation.

In this sense, the present study could benefit from adding some sort of back of the envelope welfare calculation that allows a sketch answer about the net welfare effects for the beneficiaries of the policies under study. Also, it could interesting to use other type of less subjective outcomes or additional variation to give more precise answers to the research question. One possibility could be using truthful declaration of recipiency as an outcome variable and assess whether reporting is heterogeneous across household location, following Celhay et al. (2022). I.e., those surrounded by a higher proportion of beneficiaries in their neighbourhood should theoretically report less stigma by under-reporting in a lesser proportion. Lastly, it would be interesting to study heterogeneities across years of recipiency or exits from the program, which would allow to explore whether stigma dissipates over time, especially when individuals leave the program. This is a possible strategy to undertake here since the administrative records allow to identify who continues (and does not) receiving the transfers over the years.

Finally, policy makers should carefully consider the possibility that governmental policies designed to alleviate poverty may have non-desired side effects regarding shame, humiliation and stigma perceived by their beneficiaries, as well as fostering stigmatizing political discourses surrounding them. As Bertrand et al. (2006) highlight, the framing of public policies is a factor of great importance for fulling achieving their objectives, independently of their design. Two possible policy recommendations can be raised from this study in this regard. On one hand, masking the TUS food card and making it a non-visible transfer for others could help avoid external humiliating and discriminating encounters, as those described in the qualitative evidence from Section 3.3. What also could be even more efficient both in terms of stigma and administrative costs would be for AFAM-PE and TUS to combine in an unique instrument which only differs in the amount transferred to beneficiaries as is proposed in Riella et al. (2018). On the other hand, framing benefits as a "right" people are entitled to, especially when children are involved, could help to alleviate internal shame feelings. A policy framing that seeks to enhance the agency and autonomy of its beneficiaries could contribute to prevent stigmatizing rhetorics and help cash transfer programs to achieve their main goal: alleviate the scarcity faced by a considerable part of the region's population.

References

Alkire, S. (2007). The missing dimensions of poverty data: Introduction to the special issue. *Oxford Development Studies*, 35(4):347–359. 5, 16

Amarante, V., Manacorda, M., Miguel, E., and Vigorito, A. (2016). Do cash transfers improve birth outcomes? evidence from matched vital statistics, program, and social security data. *American Economic Journal: Economic Policy*, 8(2):1–43. 5, 6, 13

Amarante, V. and Vigorito, A. (2011). Una propuesta para la evaluación de impacto del programa asignaciones familiares-plan de equidad y la tarjeta alimentaria del mides. Reporte t'ecnico, FCEA-UDELAR, MIDES , Montevideo, Uruguay. 15, 21

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. 22

Arim, R., Cruces, G., and Vigorito, A. (2009). *Programas sociales y transferencias de ingresos en Uruguay: los beneficios no contributivos y las alternativas para su extensión*. Number 146 in Serie Politicas sociales. CEPAL, Santiago de Chile. 6

Atuesta, B. and Cecchini, S. (2017). Conditional cash transfer programmes in latin america and the caribbean: Coverage and investment trends. Social Policy Series 224, Division of the Economic Commission for Latin Americaand the Caribbean. 2, 60

Bargain, O., Immervoll, H., and Viitamäki, H. (2012). No claim, no pain. Measuring the non-take-up of social assistance using register data. *The Journal of Economic Inequality*, 10(3):375–395. 1

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., and Pellerano, L. (2016). Cash transfers: What does the evidence say? a rigorous review of programme impact and of the role of design and implementation features. Overseas development institute. 3

Baumberg, B. (2015). The stigma of claiming benefits: a quantitative study. *Journal of Social Policy*, 45(2):181–199. 2, 3, 10, 11, 17

Baumberg, B., Bell, K., and Gaffney, D. (2012). Benefits stigma in britain. Project report, Turn2us, Elizabeth Finn Care. 1

Bérgolo, M. and Cruces, G. (2021). The anatomy of behavioral responses to social assistance when informal employment is high. *Journal of Public Economics*, 193:104313. 5, 13

Bérgolo, M., Dean, A., Perazzo, I., and Vigorito, A. (2016). Evaluación impacto del programa Asignaciones Familiares - Plan de Equidad y la Tarjeta Alimentaria del MIDES. Ministerio de desarrollo social (MIDES). 7, 13, 15, 23

Bernal, R. and Peña, X. (2011). *Guía práctica para la evaluación de impacto*. Universidad de los Andes, Colombia. 21, 26

Bertrand, M., Mullainathan, S., and Shafir, E. (2006). Behavioral economics and marketing in aid of decision making among the poor. *Journal of Public Policy & Marketing*, 25(1):8–23. 34

Besley, T. and Coate, S. (1992). Understanding welfare stigma: Taxpayer resentment and statistical discrimination. *Journal of Public Economics*, 48(2):165–183. 1, 2, 9

Bhargava, S. and Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review*, 105(11):3489–3529. 1, 11

Blumkin, T., Margalioth, Y., and Sadka, E. (2008). The role of stigma in the design of welfare programs. *CESifo Working Paper Series*. 9

Brizmohun, R. and Duffy, P. A. (2016). Do personal attitudes about welfare influence food stamp participation? (333-2016-14271):23. Paper 9133. 11

Bugni, F. A. and Canay, I. A. (2021). Testing continuity of a density via g-order statistics in the regression discontinuity design. *Journal of Econometrics*, 221(1):138–159. 22, 45, 68

Burdín, G. and de Melo, G. (2009). Prestaciones sociales en uruguay: ¿por qué hay beneficiarios que no las solicitan? *Revista CEPAL*. 12

Bursztyn, L., Egorov, G., and Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American Economic Review*, 110(11):3522–48. 1

Bérgolo, M. and Galván, E. (2018). Intra-household behavioral responses to cash transfer programs. evidence from a regression discontinuity design. *World Development*, 103:100 – 118. 5, 13, 23

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326. 25, 32, 58

Campbell, A. L. (2012). Policy makes mass politics. *Annual Review of Political Science*, 15(1):333–351. 1

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *Stata Journal*, 18(1):234–261. 22, 25, 32, 45, 58, 68

Celhay, P. A., Meyer, B. D., and Mittag, N. (2022). Stigma in welfare programs. Working Paper 30307, National Bureau of Economic Research. 4, 12, 14, 27, 34

Ceni, R. and Salas, G. (2021). Transfer program enforcement and children's time allocation. *Review of Economics of the Household*, 19(4):1099–1137. 5

Chakraborty, T., Mukherjee, A., Rachapalli, S. R., and Saha, S. (2018). Stigma of sexual violence and women's decision to work. *World Development*, 103:226–238. 1

Chong, A., Nopo, H., and Ríos, V. (2009). Do Welfare Programs Damage Interpersonal Trust? Experimental Evidence from Representative Samples for Four Latin American Cities. *SSRN Electronic Journal*. 12

Clarke, D., Romano, J. P., and Wolf, M. (2020). The romano–wolf multiple-hypothesis correction in stata. *The Stata Journal*, 20(4):812–843. 28

Coady, D., Grosh, M., and Hoddinott, J. (2004). *Targeting of Transfers in Developing Countries : Review of Lessons and Experience*. Number 14902 in World Bank Publications. The World Bank. 2, 3

Colafranceschi, M. and Vigorito, A. (2013). Uruguay: evaluación de las políticas de transferencias. La estrategia de inclusión y sus desafíos. In Rofman, R., editor, *HACIA UN URUGUAY MÁS EQUITATIVO. LOS DESAFÍOS DEL SISTEMA DE PROTECCIÓN SOCIAL*. Banco Mundial, Montevideo, Uruguay. 6

Colombo, K., Lohner, G., and Ramirez-Diaz, E. (2018). The impact of a cash transfer program on life outcomes: Evidence from uruguay. 5, 13

Currie, J. (2004). The take up of social benefits. Working Paper 10488, National Bureau of Economic Research. 1

de Brauw, A. and Hoddinott, J. (2011). Must conditional cash transfer programs be conditioned to be effective? the impact of conditioning transfers on school enrollment in mexico. *Journal of Development Economics*, 96(2):359 – 370. 3

Dean, A. and Vigorito, A. (2015). El acceso a programas de transferencias de ingreso de la población de menores recursos en Uruguay. Un análisis en base a la Encuesta de panel INE-MIDES-UDELAR. Serie Documentos de Trabajo, DT 01/2015, Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Montevideo, Uruguay. 12

Della Guardia, A., Lake, M., and Schnitzer, P. (2022). Selective inclusion in cash transfer programs: Unintended consequences for social cohesion. *World Development*, 157:105922. 4

Devereux, S. (2016). Is targeting ethical? *Global Social Policy*, 16(2):166–181. 3, 14

Failache, E., Giaccobasso, M., and Ramírez, L. (2016). Transferencias de ingresos y mercado de trabajo: el impacto de Asignaciones Familiares Plan de Equidad sobre la informalidad laboral. Serie Documentos de investigación estudiantil, DIE 01/2016, Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Montevideo, Uruguay. 5, 13, 23

Filmer, D. and Pritchett, L. H. (2001). Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India*. *Demography*, 38(1):115–132. 18, 80

Fiszbein, A., Schady, N., Ferreira, F. H., Grosh, M., Keleher, N., Olinto, P., and Skoufias, E. (2009). *Conditional Cash Transfers : Reducing Present and Future Poverty*. Number 2597 in World Bank Publications - Books. The World Bank Group. 2

Friedrichsen, J., König, T., and Schmacker, R. (2018). Social image concerns and welfare take-up. *Journal of Public Economics*, 168:174–192. 4, 11, 14, 27, 34

Gao, Q. and Zhai, F. (2017). Public Assistance, Economic Prospect, and Happiness in Urban China. *Social Indicators Research*, 132(1):451–473. 4, 30

Ghazarian, M. (2020). *La no adhesión a las AFAM-PE: un análisis con énfasis en los determinantes socioculturales y psicológicos.* Tesis de Maestría, Universidad de la República, Facultad de Ciencias Económicas y de Administración, Montevideo, Uruguay. 12

Ghosal, S., Jana, S., Mani, A., Mitra, S., and Roy, S. (2022). Sex Workers, Stigma, and Self-Image: Evidence from Kolkata Brothels. *The Review of Economics and Statistics*, 104(3):431–448. 1

Goffman, E. ([1963]). *Stigma; Notes on the Management of Spoiled Identity*. Englewood Cliffs, N.J. : Prentice-Hall, [1963]. 1, 9

Han, H. and Gao, Q. (2020). Does Welfare Participation Improve Life Satisfaction? Evidence from Panel Data in Rural China. *Journal of Happiness Studies*, 21(5):1795–1822. 4, 12

Harder, D. H. and Zalma, A. (1990). Two promising shame and guilt scales: A construct validity comparison. *Journal of personality assessment*, 55(3-4):729–745. 17

Hartling, L. M. and Luchetta, T. (1999). Humiliation: Assessing the impact of derision, degradation, and debasement. *The Journal Primary Prevention*, 19(4):259–278. 11, 17

Haushofer, J., Chemin, M., Jang, C., and Abraham, J. (2020). Economic and psychological effects of health insurance and cash transfers: Evidence from a randomized experiment in kenya. *Journal of Development Economics*, 144:102416. 5

Haushofer, J. and Fehr, E. (2014). On the psychology of poverty. *Science*, 344(6186):862–867. 5

Haushofer, J., Reisinger, J., and Shapiro, J. (2019). s your gain my pain? effects of relative income and inequality on psychological well-being. Working paper. 5

Haushofer, J. and Shapiro, J. (2016). The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya*. *The Quarterly Journal of Economics*, 131(4):1973–2042. 5

Hochfeld, T. and Plagerson, S. (2011). Dignity and stigma among south african female cash transfer recipients. *IDS Bulletin*, 42(6):53–59. 4

Hojman, D. A. and Álvaro Miranda (2018). Agency, human dignity, and subjective well-being. *World Development*, 101:1–15. 4, 18, 30

Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-run impacts of childhood access to the safety net. *American Economic Review*, 106(4):903–34. 18, 28

Hungerman, D. M. (2013). Substitution and stigma: Evidence on religious markets from the catholic sex abuse scandal. *American Economic Journal: Economic Policy*, 5(3):227–53. 1

Kaye, L., Lee, E., and Chen, Y. Y. (2013). Barriers to food stamps in new york state: A perspective from the field. *Journal of Poverty*, 17(1):13–28. 11

Kleven, H. J. and Kopczuk, W. (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy*, 3(1):54–90. 1

Ko, W. and Moffitt, R. A. (2022). Take-up of social benefits. Working Paper 30148, National Bureau of Economic Research. 1

Lamont, M. (2018). Addressing recognition gaps: Destigmatization and the reduction of inequality. *American Sociological Review*, 83(3):419–444. 2, 11

Lamont, M. (2019). From 'having' to 'being': self-worth and the current crisis of american society. *The British Journal of Sociology*, 70:660–707. 2

Lamont, M., Silva, G. M., Welburn, J. S., Guetzkow, J., Mizrachi, N., Herzog, H., and Reis, E. (2016). *Getting Respect: Responding to Stigma and Discrimination in the United States, Brazil, and Israel*. Princeton University Press. 1, 2, 10, 14

Lapham, J. and Martinson, M. L. (2022). The intersection of welfare stigma, state contexts and health among mothers receiving public assistance benefits. *SSM - Population Health*, 18:101117. 4

Lavalleja, M. and Tenenbaum, V. (2020). Mecanismos de focalización de los programas de transferencias monetarias en América Latina: el caso del Uruguay. Estudios y Perspectivas – Oficina de la CEPAL en Montevideo 43, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL). 6, 7

Lee, D. S. and Lemieux, T. (2009). Regression discontinuity designs in economics. Working Paper 14723, National Bureau of Economic Research. 20, 24

Lindbeck, A., Nyberg, S., and Weibull, J. W. (1999). Social norms and economic incentives in the welfare state. *The Quarterly Journal of Economics*, 114(1):1–35. 1

Lindner, E. G. (2007). In times of globalization and human rights: Does humiliation become the most disruptive force? *Journal of Human Dignity and Humiliation Studies*, 1(1):5–30. 11, 17

Link, B. G. and Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology*, 27(1):363–385. 2, 10

Manacorda, M., Miguel, E., and Vigorito, A. (2011). Government transfers and political support. *American Economic Journal: Applied Economics*, 3(3):1–28. 5, 6, 13

Manchester, C. F. and Mumford, K. (2010). How costly is welfare stigma? separating psychological costs from time costs. Purdue university economics working papers, Purdue University, Department of Economics. 11

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714. The regression discontinuity design: Theory and applications. 22, 45, 68

MIDES (2012). Informe Tarjeta Uruguay Social. Reporte técnico, Ministerio de Desarrollo Social (MIDES). 8

MIDES (2013). ¿Qué es el Índice de Carencias Críticas? Reporte técnico, Ministerio de Desarrollo Social (MIDES). 7, 9, 62

Mills, C. and Zavaleta, D. (2015). *Shame, Humiliation and Social Isolation: Missing Dimensions of Poverty and Suffering Analysis*. Springer Netherlands, Dordrecht. 4, 11, 16, 17, 28

Mkandawire, T. (2005). Targeting and universalism in poverty reduction. *Social Policy and Development Programme Paper Number 23*. 3, 14

Moffitt, R. (1983). An economic model of welfare stigma. *The American Economic Review*, 73(5):1023–1035. 1, 9

Molyneux, M., Jones, W. N., and Samuels, F. (2016). Can cash transfer programmes have 'transformative' effects? *The Journal of Development Studies*, 52(8):1087–1098. 3

Moreno, M., Rojo, V., Rosa, C. D., Real, M., Chiossi, J., and Aguerre, E. (2014). Evaluación cualitativa del programa Tarjeta Uruguay Social: Informe Final. Reporte técnico, Dirección Nacional de Evaluación y Monitoreo (DINEM) - MIDES. 1, 4, 7, 8, 14

Nussbaum, M. (2009). Hiding from humanity: Disgust, shame, and the law. *Hiding from Humanity: Disgust, Shame, and the Law*. 2, 11

Osborne, J., C. A. . K. J. (2008). Best practices in exploratory factor analysis. In Osborne, J., editor, *Best practices in quantitative methods*. SAGE Publications, Inc. 80

Osman, A. and Speer, J. D. (2022). Stigma and Take-Up of Labor Market Assistance: Evidence from Two Field Experiments. Working paper. 1

Oviedo, R. (2019). Un análisis normativo de las políticas de transferencias monetarias directas en Uruguay: Asignaciones Familiares del Plan de Equidad y Tarjeta Uruguay Social. Serie Documentos de investigación estudiantil, DIE 06/2019, Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Montevideo, Uruguay. 8

Qi, D. and Wu, Y. (2018). Does welfare stigma exist in china? policy evaluation of the minimum living security system on recipients' psychological health and wellbeing. *Social Science Medicine*, 205:26 – 36. 4, 12, 30

Rainwater, L. (1982). 2 - stigma in income-tested programs. In Garfinkel, I., editor, *Income-Tested Transfer Programs*, pages 19 – 65. Academic Press. 10

Riella, R., Rijo, L., Urraburu, J., and Vera, S. (2018). Impuestos y transferencias monetarias dirigidas a niños y adolescentes. Oficina de Planeamiento y Presupuesto, Dirección de Presupuestos, Control y Evaluación de la Gestión. 34

Rivero, A., Katkzkowicz, N., Parada, C., and Vigorito, A. (2019). Evaluación cuantitativa del impacto de Asignaciones Familiares - Plan de Equidad y Tarjeta Uruguay Social. Informe de la ronda 2016/18. Ministerio de desarrollo social (MIDES). 13, 16, 23

Rivero, L. (2020). Anexo 6: Informe final de grupos focales. In *Una propuesta para la medición del bienestar multidimensional en Uruguay. Informe final. CONVENIO MIDES - INE - MEF - OPP- FAFCEA*. Ministerio de Desarrollo Social (MIDES), Montevideo, Uruguay. 1, 14

Roelen, K. (2017). Shame, Poverty and Social Protection. IDS Working Paper 489, Brighton: IDS. 3

Roelen, K. (2020). Receiving social assistance in low- and middle-income countries: Negating shame or producing stigma? *Journal of Social Policy*, 49(4):705–723. 3, 10, 11

Romero, J., Esopo, K., McGuire, J., and Haushofer, J. (2021). The effect of economic transfers on psychological well-being and mental health. Working paper. 5

Rossel, C., Courtoisie, D., and Marsiglia, M. (2014). Programas de transferencias, condicionalidades y derechos de la infancia: apuntes a partir del caso del Uruguay. Políticas Sociales 215, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL). 7

Rossel, C. and Straschnoy, M. (2020). ¿Cuánto pueden condicionar las condicionalidades? Evidencia sobre las asignaciones familiares de Argentina y Uruguay. *Latin American Research Review*, 55(1):16–30. 7

Spicker, P. (2011). *Stigma and Social Welfare*. 2, 10, 16

Stampini, M. and Tornarolli, L. (2012). The growth of conditional cash transfers in latin america and the caribbean: Did they go too far? 2

Stock, J. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York. 25

Stuber, J. and Kronebusch, K. (2004). Stigma and other determinants of participation in tanf and medicaid. *Journal of Policy Analysis and Management*, 23(3):509–530. 11

Stuber, J. and Schlesinger, M. (2006). Sources of stigma for means-tested government programs. *Social Science Medicine*, 63(4):933 – 945. 2, 3, 11

Tangney, J. and Tracy, J. (2011). Self-Conscious Emotions. UBC Emotion & Self Lab, British Columbia. 10, 17, 28

Tangney, J. P. and Dearing, R. L. (2002). *Shame and Guilt*. Guilford Press, New York, NY, US, 1st edition. 17

Tracy, J. L. and Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, 15(2):103–125. 10, 28

van Oorschot, W. (2002). Targeting welfare : On the functions and dysfunctions of means-testing in social policy. In P. Townsend, D. Gorden (Eds.), World poverty: new policies to defeat an old enemy. pages 171 – 193. (Studies in poverty, inequality and social exclusion). Policy Press. 3

Walker, R. (2014). *The shame of poverty*. Oxford University Press. 1, 2, 3, 10, 11, 14, 16, 17, 18

Yaniv, G. (1997). Welfare fraud and welfare stigma. *Journal of Economic Psychology*, 18(4):435–451. 9

Yin, R. and Etilé, F. (2019). Measuring identity orientations for understanding preferences: A french validation of the aspects-of-identity questionnaire. 18, 80, 81, 82

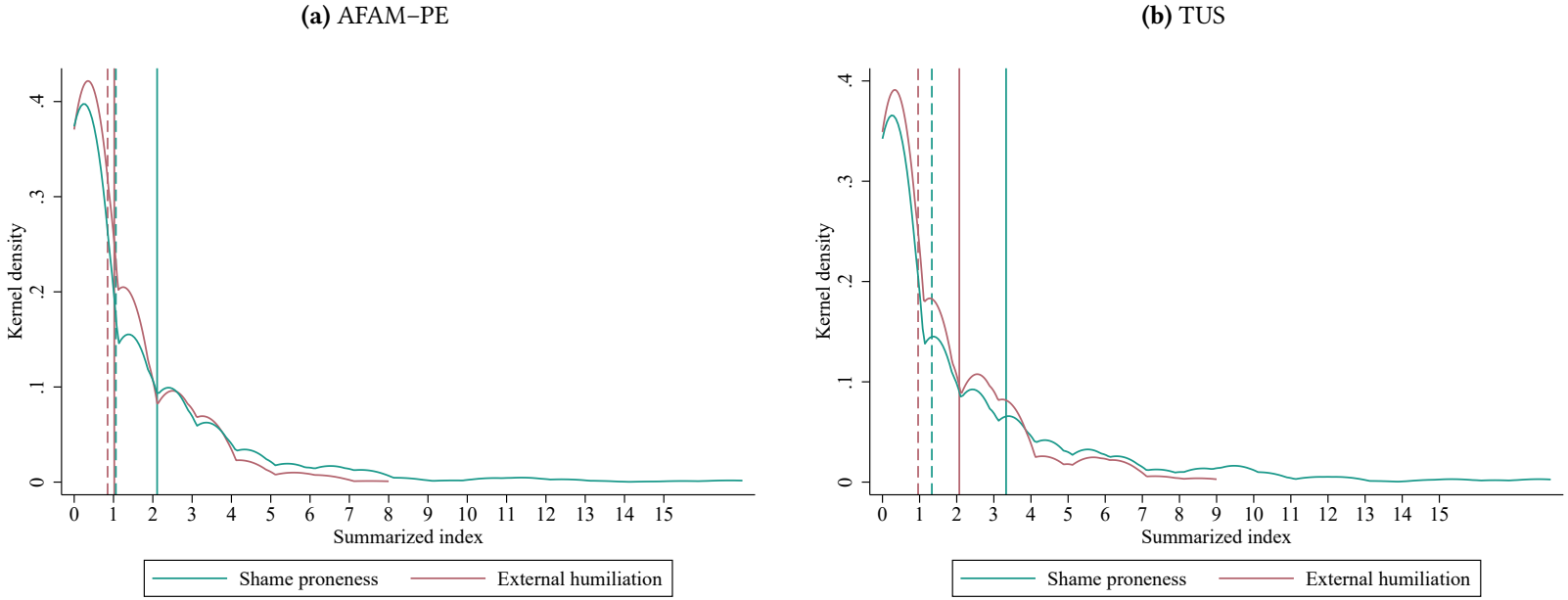
Zavaleta, D. (2007). The ability to go about without shame: A proposal for internationally comparable indicators of shame and humiliation. *Oxford Development Studies*, 35(4):405–430. 3, 10, 16, 17, 18, 28, 76, 80

Zhang, H. (2020). Welfare stigma of social assistance on women recipients in urban china. *Journal of Social Service Research*, 46(4):538–552. 4, 12

Figures and Tables

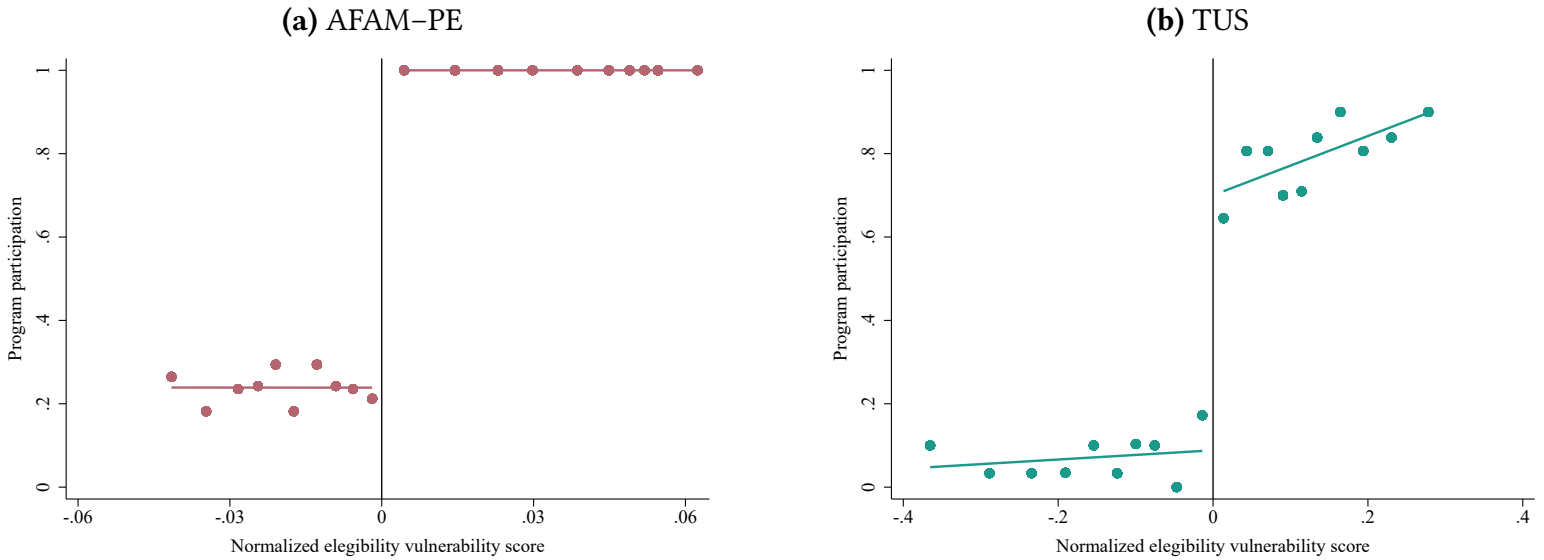
Figures

Figure 1:
Kernel Densities of Main Aggregated Outcomes



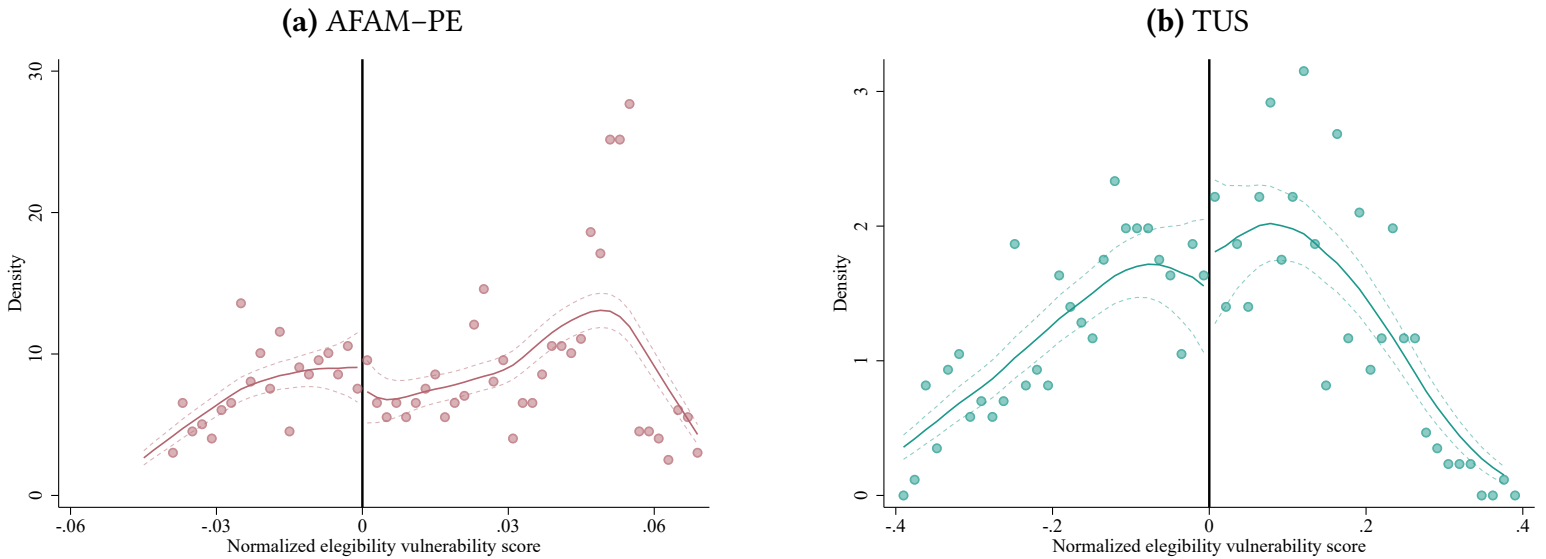
Notes: These figures show epanechnikov kernel densities of shame (green) and humiliation (red) proneness indexes. Panel (a) displays results for the AFAM-PE sample, while Panel (b) shows results for the TUS sample. The dashed vertical lines represent the average value for the non-eligible group while the solid vertical lines represent the sum of this non-eligible mean plus the estimated effect of the program in each case. These lines reflect the estimated push caused by each program across the distribution of the indexes. For instance, the effect of AFAM-PE on shame would increase the mean for the non-eligible group by 99% from 1.062 to 2.112, moving an individual from the 53th to the 70th percentile in the shame scale distribution in the AFAM-PE sample. Meanwhile, the effect of TUS on shame (humiliation) would increase the non-eligible mean by 150% (115%) from 1.333 (0.960) to 3.337 (2.072), moving an individual from the 48th (48th) to the 75th (65th) percentile in the shame (humiliation) scale distribution in the TUS sample. Table A3 reports the full summary statistics for these indexes for each program separately, and for each of the summarized and standardized versions.

Figure 2:
First Stage Estimates: Participation and Eligibility According to the Vulnerability Score



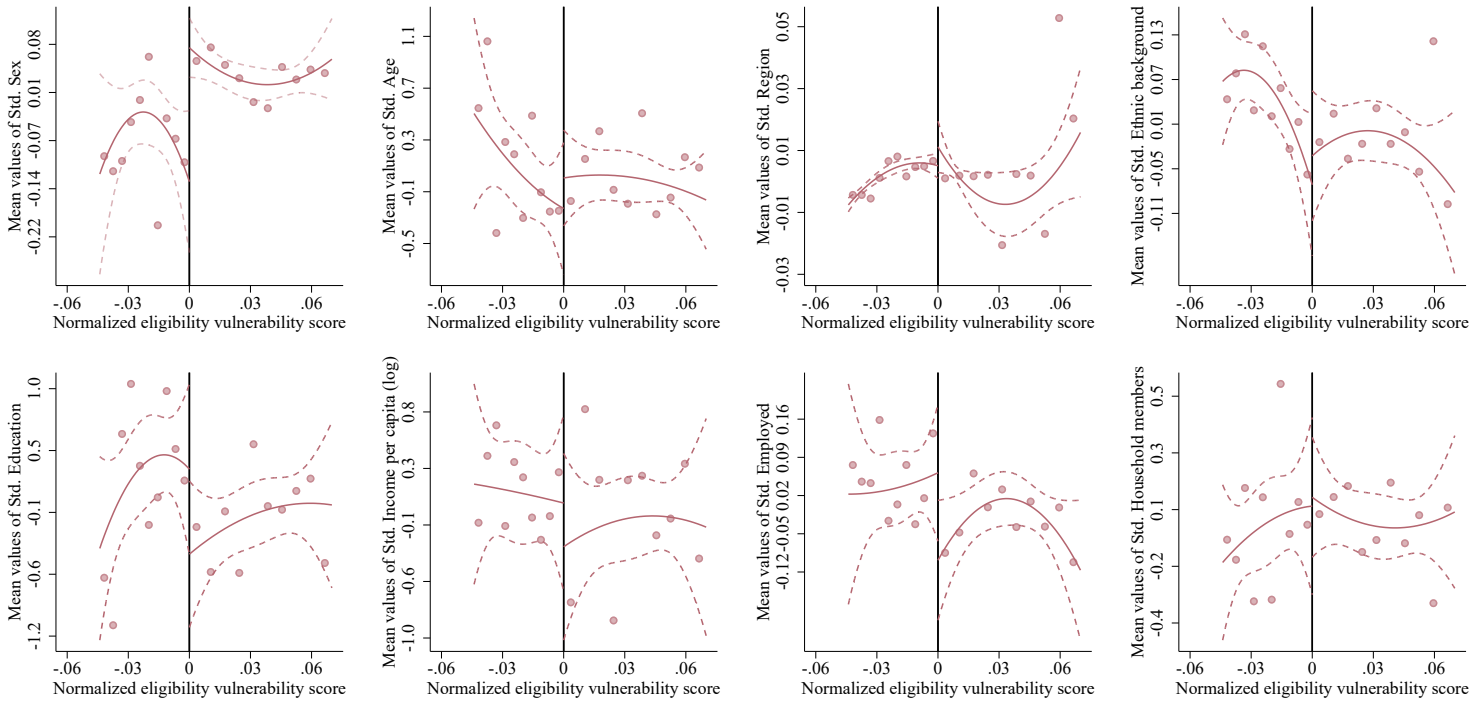
Notes: These figures show first stage estimates of each program's effective participation according to the standardized eligibility vulnerability score, *Índice de Carencias Críticas* (ICC*). The dots represent the share of participants in the program over ten bin intervals at each side of the ICC* cut-offs at zero, represented by the vertical black lines. The lines correspond to linear estimated fits. Panel (a) displays results for the AFAM-PE sample, while Panel (b) displays results for the TUS sample for [-0.047; 0.073] and [-0.400 ; 0.400] bandwidths of each standardized running variable, respectively. The value of the jump at the cut-off which reflects the probability of effectively participating in the program according to its eligibility score is around 0.75 and 0.63, for AFAM-PE and TUS respectively. The F-Test values for these first stages are 366.67 and 75.13, respectively. Figure A9 shows the same graph for the sub-sample analyses.

Figure 3:
McCrary Density Test: Bunching Around the Eligibility Threshold



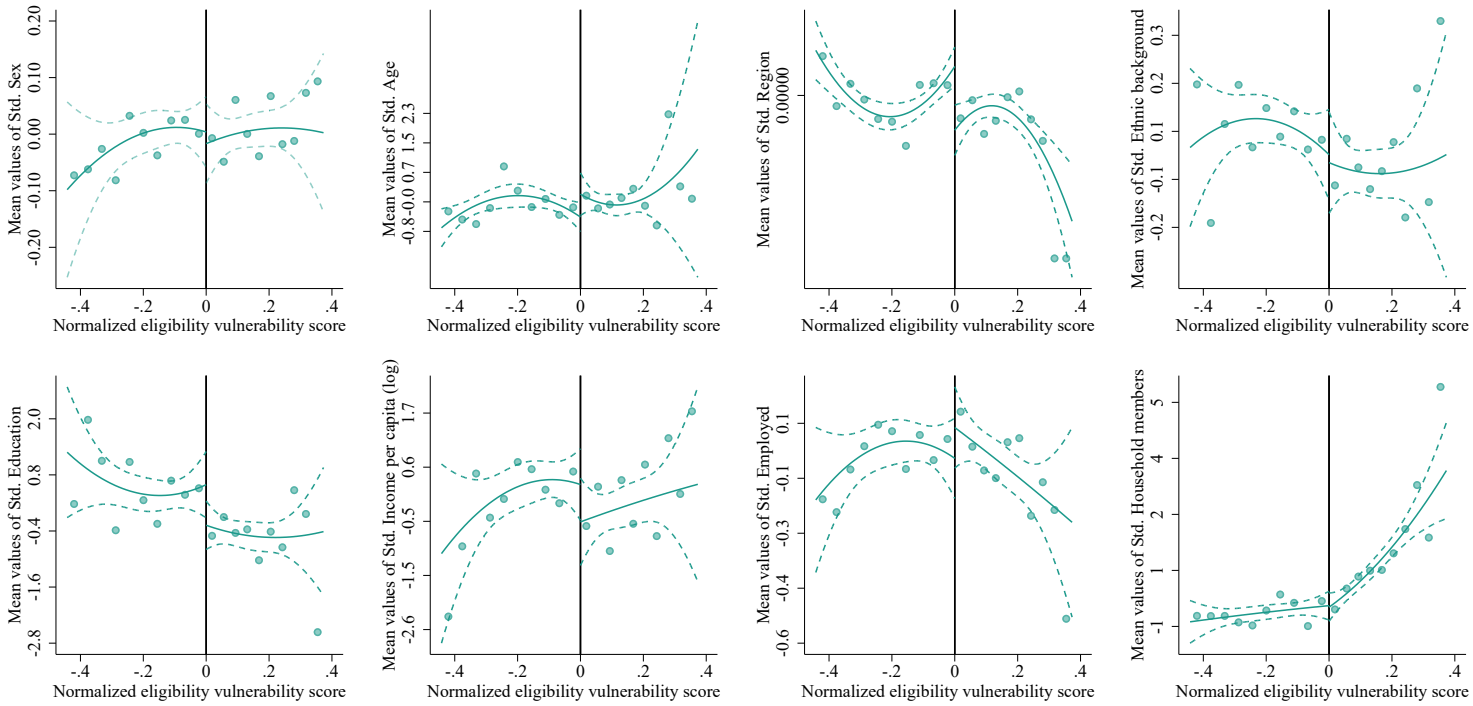
Notes: These figures show the McCrary (2008) density test to check for the presence of systematic bunching at the right side of the cut-off of each program, represented by the vertical black lines at zero. The solid thick line indicates the density estimate, computed from a local linear regression with separate trends for each side of the threshold. The dashed lines represent 95% confidence intervals. Panel (a) displays results for the AFAM-PE sample, while Panel (b) displays results for the TUS sample for [-0.047; 0.073] and [-0.400 ; 0.400] bandwidths of each standardized running variable, respectively. The estimated log difference in the heights at the cut-off is -0.186 with a standard error of 0.226, and 0.152 with a standard error of 0.248, for AFAM-PE and TUS, respectively. Therefore, the hypotheses of continuity is not rejected. The same conclusion is reached when considering other newer manipulation tests: the p-values of the tests for each program are 0.336 and 0.646 for the one proposed by Cattaneo et al. (2018) and 0.780 and 0.999 for the one proposed by Bugni and Canay (2021). Figure A10 shows the same graph for the sub-sample analyses.

Figure 4:
Balance Test Around the Eligibility Threshold
AFAM–PE



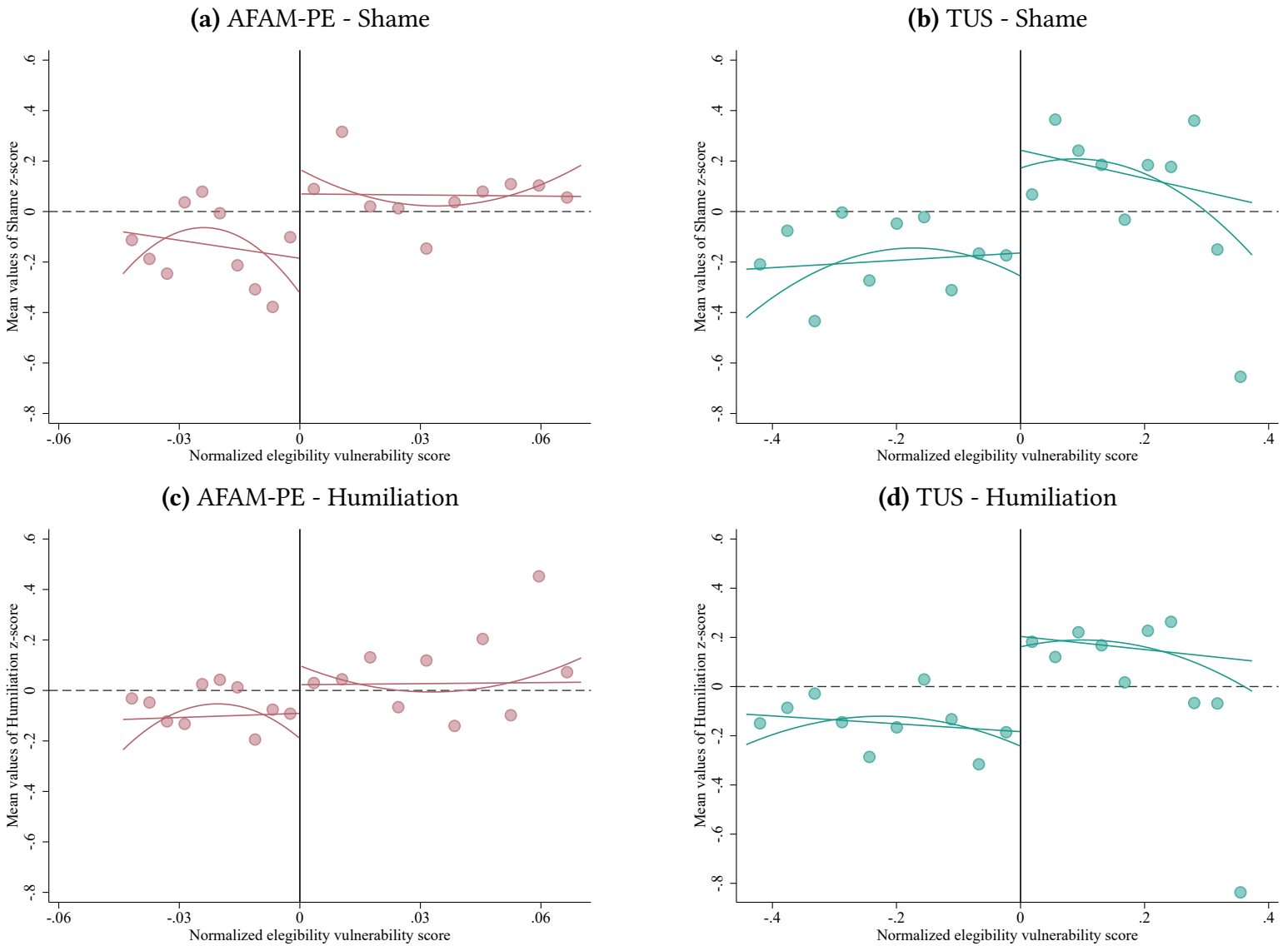
Notes: These figures exhibit graphical evidence of the balance test displayed in Table 1 for the AFAM-PE program, which checks for systematic imbalances of covariates at each side of the cut-off. It shows RDD estimates using baseline covariates as dependent variables for a [-0.047; 0.073] bandwidth of its standardized running variable. The dots represent their mean standardized value over ten bin intervals at each side of the ICC* cut-off, which is marked with the vertical bar at zero. The thick line represents a quadratic estimated fit. The dashed lines represent 90% confidence intervals. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1), each being taken out when in the presence of collinearity, e.g sex is taken out when it's being evaluated as an outcome. Standard errors clustered by ICC*. Figure A11 shows an alternative balance test for a women sub-sample only where imbalances are less prevalent.

Figure 5:
Balance Test Around the Eligibility Threshold
TUS



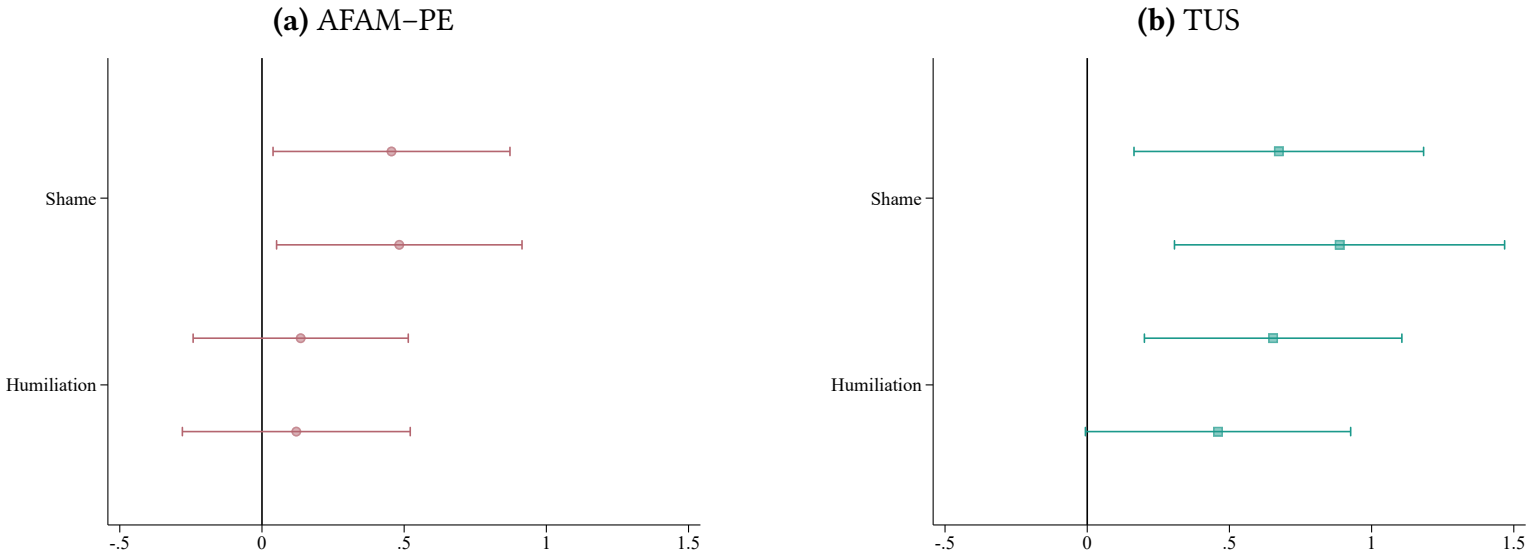
Notes: These figures exhibit graphical evidence of the balance test displayed in Table 1 for the TUS program, which checks for systematic imbalances of covariates at each side of the cut-off. It shows RDD estimates using baseline covariates as dependent variables for a [-0.040; 0.040] bandwidth of its standardized running variable. The dots represent their mean standardized value over ten bin intervals at each side of the ICC* cut-off, which is marked with the vertical bar at zero. The thick line represents a quadratic estimated fit. The dashed lines represent 90% confidence intervals. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1), each being taken out when in the presence of collinearity, e.g sex is taken out when it's being evaluated as an outcome. Standard errors clustered by ICC*. Figure A12 shows an alternative balance test for a non-capital residents sub-sample only where imbalances are less prevalent.

Figure 6:
RDD Impact Estimates of AFAM-PE and TUS on Shame and Humiliation



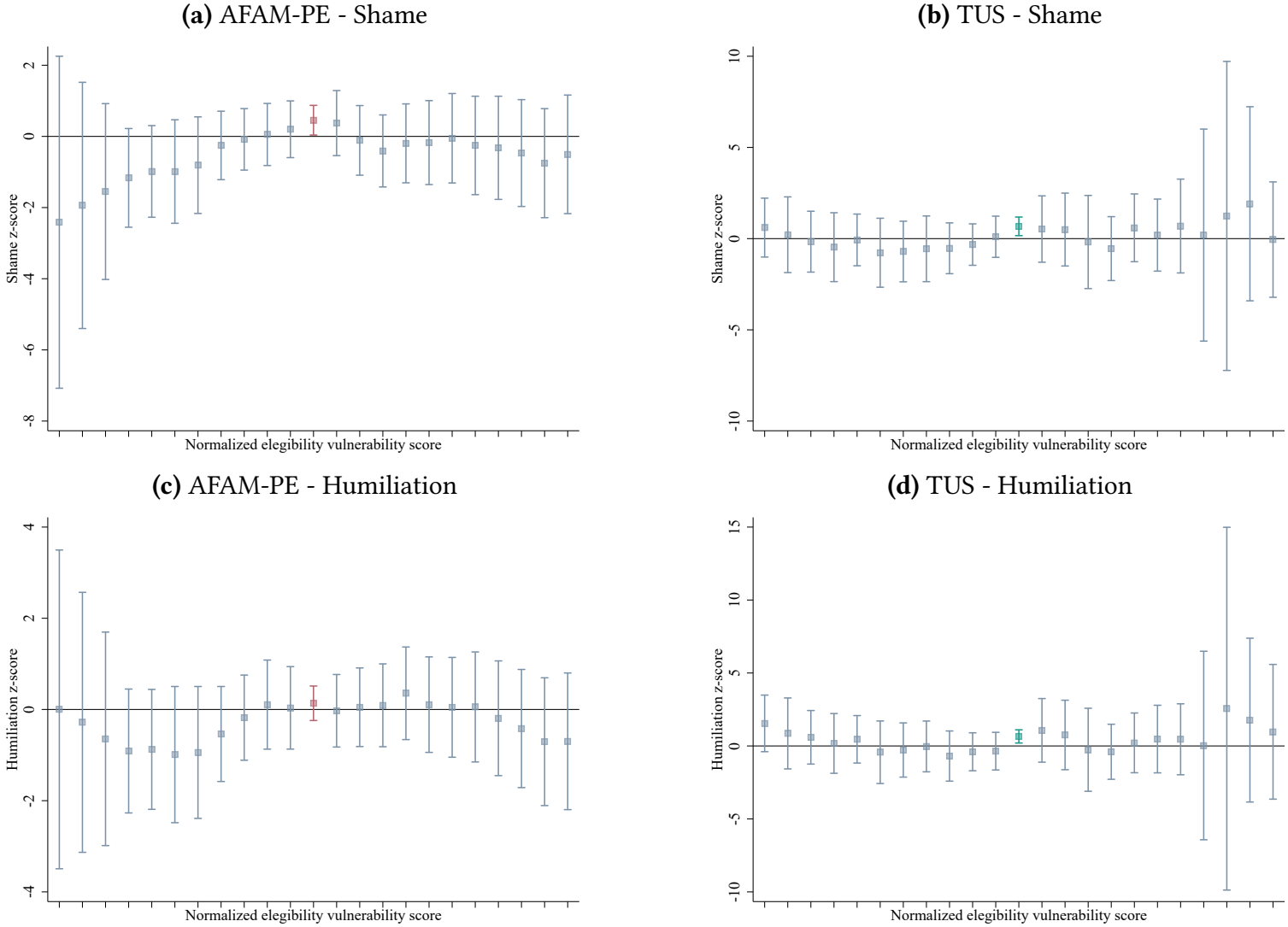
Notes: These figures show parametric sharp RDD impact estimates for the main outcome z-scores indices of shame and humiliation. The dots represent the mean value of the shame (humiliation) z-score index over ten bin intervals at each side of the ICC* cut-off, which is marked with a vertical bar at zero. The lines represent linear and quadratic estimated fits, respectively. Panels (a) and (c) show estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while Panels (b) and (d) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC*. Table 3 shows the same analysis in regression form.

Figure 7:
Robustness Test: Sub-Sample Analyses
Sub-Sample vs. Full Sample RDD Impact Estimates



Notes: These figures show robustness check for sub-sample analyses. It plots the main coefficients of interest β, which represent the effect of each program on each outcome. These are obtained through the regressing the linear specification with controls of Equation 8, and compares results for the full sample versus estimates for the sub-sample, which only include women for AFAM-PE and non-capital city residents for TUS. The first and third coefficients in each panel correspond to the full sample estimates, while the second and fourth correspond to the sub-sample analyses. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1), depending on the sub-sample. Standard errors clustered by ICC*. 95% confidence intervals are displayed.

Figure 8:
Robustness Test: Placebo Estimates



Notes: These figures show placebo estimates for the main outcome variables of shame and humiliation, for each program. They plot the coefficients of interest β, which represent the effect of each program on each outcome. The alternative cut-offs considered for AFAM-PE range from [-0.020, -0.019, ..., -0.011 -0.010] to [0.01, 0.011. ... 0.019, 0.020], while for TUS range from [-0.20, -0.19, ..., -0.11 -0.10] to [0.10, 0.11 ... 0.19, 0.20]. These are estimates are displayed in gray, while the ones with a cut-off at zero, which are the impact estimates of the programs, are highlighted in color. Estimates are obtained through regressions following the preferred specification: a linear Fuzzy RDD with controls, as described in Equation 8. Panels (a) and (c) show estimates for the AFAM-PE program while Panels (b) and (d) show estimates for the TUS program. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC*. 95% confidence intervals are displayed.

Tables

Table 1:
Balance Tests

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Sex (woman = 1)	0.165**	0.268***	-0.058	-0.035
	(0.063)	(0.091)	(0.062)	(0.104)
Age	-6.117***	3.268	-4.549*	-4.691
	(2.047)	(2.859)	(2.503)	(4.029)
Region (Montevideo = 1)	-0.606***	0.030	0.084	0.600**
	(0.109)	(0.154)	(0.107)	(0.238)
Ethnicity (white = 1)	-0.038	0.038	-0.126	-0.041
	(0.070)	(0.102)	(0.113)	(0.186)
Education (years)	-1.312**	-1.026	-1.149*	-1.638*
	(0.540)	(0.775)	(0.594)	(0.972)
Per capita household income (log)	-0.387	-0.488	-1.662**	-1.331
	(0.605)	(0.817)	(0.813)	(1.294)
Labor status (employed = 1)	-0.113	-0.218	0.054	0.173
	(0.095)	(0.133)	(0.123)	(0.208)
Household size (#)	-0.064	0.053	-0.134	0.080
	(0.229)	(0.332)	(0.380)	(0.591)
Control variables	Yes	Yes	Yes	Yes
Polynomial degree	1	2	1	2
F-Statistic	58.24***	19.50**	16.20**	20.00**

Notes: This table tests for systematic imbalances of covariates at each side of the cut-offs by running RDD estimates for baseline covariates as dependent variables, following Equation 8. All coefficients correspond to the parameter of interest β, which represent the effect of each program on each considered outcome. In this case, these display the results of separate regressions to test for discontinuities around the eligibility thresholds. Columns (1) and (2) show impact estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while columns (3) and (4) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. Control variables include: sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1), each being taken out when in the presence of collinearity, e.g region is taken out when it's being evaluated as an outcome. The last row of this table displays the value and statistical significance of an F-Statistic test that checks whether all coefficients are jointly not different from zero for each specification. Standard errors clustered by ICC* are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.10. Figures 4 and 5 show the same analysis in graphical form for each program. Table 2 reports an alternative balance test for sub-samples where imbalances are less prevalent, women for the AFAM-PE analysis and non-capital residents for the TUS analysis.

Table 2:
Alternative Balance Tests
Sub-sample Analyses

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Sex (woman = 1)	-	-	-0.098	-0.149
			(0.067)	(0.124)
Age	-6.769***	-4.722	-4.592*	-3.449
	(2.219)	(3.011)	(2.663)	(4.576)
Region (Montevideo = 1)	-0.571***	0.051	-	-
	(0.119)	(0.169)		
Ethnicity (white = 1)	0.042	0.084	-0.019	-0.089
	(0.078)	(0.110)	(0.117)	(0.191)
Education (years)	-1.429**	-1.311	-0.530	-0.683
	(0.591)	(0.839)	(0.598)	(0.950)
Per capita household income (log)	0.135	0.321	-1.262	-0.483
	(0.674)	(0.894)	(0.800)	(1.352)
Labor status (employed = 1)	-0.113	-0.151	0.030	0.237
	(0.095)	(0.152)	(0.130)	(0.227)
Household size (#)	-0.014	0.050	-0.373	0.136
	(0.109)	(0.374)	(0.398)	(0.617)
Control variables	Yes	Yes	Yes	Yes
Polynomial degree	1	2	1	2
Sub-sample	Women	Women	Non-capital	Non-capital
F-Statistic	41.97***	9.32	10.63	5.37

Notes: This table reports an alternative balance tests for sub-samples of the main analysis. It tests for systematic imbalances of covariates at each side of the cut-offs by running RDD estimates for baseline covariates as dependent variables, following Equation 8. All coefficients correspond to the parameter of interest β, which represent the effect of each program on each considered outcome. In this case, these display the results of separate regressions to test for discontinuities around the eligibility thresholds. Columns (1) and (2) show impact estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while columns (3) and (4) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. Control variables include: sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1), each being taken out when in the presence of collinearity, e.g region is taken out when it's being evaluated as an outcome. The last row of this table displays the value and statistical significance of an F-Statistic test that checks whether all coefficients are jointly not different from zero for each specification. Standard errors clustered by ICC* are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.10. Figures A11 and A12 show the same analysis in graphical form for each program.

Table 3:
RDD Impact Estimates of AFAM-PE and TUS on Shame and Humiliation

	AFAM-PE				TUS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: *Sharp RDD estimates*								
Shame z-score	0.255*	0.342**	0.490**	0.518**	0.407**	0.437***	0.428*	0.456*
	(0.165)	(0.168)	(0.234)	(0.229)	(0.191)	(0.213)	(0.234)	(0.242)
Humiliation z-score	0.113	0.101	0.289	0.257	0.387***	0.418***	0.404*	0.398*
	(0.144)	(0.159)	(0.234)	(0.229)	(0.147)	(0.148)	(0.209)	(0.205)
Panel B: *Fuzzy RDD estimates*								
Shame z-score	0.337*	0.456**	0.631**	0.684**	0.624**	0.674***	0.759*	0.773*
	(0.191)	(0.213)	(0.299)	(0.305)	(0.254)	(0.260)	(0.421)	(0.410)
Humiliation z-score	0.151	0.136	0.372	0.344	0.605***	0.654***	0.714*	0.675*
	(0.181)	(0.193)	(0.273)	(0.282)	(0.230)	(0.231)	(0.382)	(0.352)
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial degree	1	1	2	2	1	1	2	2
Observations	917	917	917	917	560	560	560	560

Notes: This table shows parametric RDD impact estimates for the main outcome z-scores indices: shame and humiliation. All coefficients correspond to the parameter of interest β, which represent the effect of each program on each considered outcome. Panel A reports OLS estimates following the Sharp RDD specification of Equation 5. Panel B reports 2SLS estimates following the Fuzzy RDD specification of Equation 8. Columns (1) to (4) show impact estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while columns (5) to (8) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC* are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.10. Figure 6 shows the same analysis in graphical form.

Table 4:
RDD Impact Estimates of AFAM-PE and TUS on Individual Item Outcomes

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Panel A: *Shame items*				
Self-conscious	0.395	0.497	0.441	0.482
	(0.172)**	(0.177)***	(0.256)*	(0.258)*
	(0.172)	**(0.177)****	**(0.256)**	**(0.258)**
Ridiculous	0.056	0.153	0.012	0.035
	(0.186)	(0.202)	(0.232)	(0.241)
	(0.186)	**(0.202)**	**(0.232)**	**(0.241)**
Embarrassed	0.247	0.334	0.469	0.490
	(0.182)	(0.196)*	(0.248)*	(0.252)*
	(0.182)	**(0.196)**	**(0.248)**	**(0.252)**
Humiliated	0.163	0.205	0.460	0.544
	(0.168)	(0.178)	(0.240)*	(0.239)**
	(0.168)	**(0.178)**	**(0.240)**	**(0.239)***
Laughable	0.276	0.369	0.636	0.640
	(0.170)	(0.191)*	(0.228)***	(0.234)***
	(0.170)	**(0.191)**	**(0.228)****	**(0.234)****
Helpless	0.154	0.203	0.632	0.677
	(0.198)	(0.221)	(0.271)**	(0.273)**
	(0.198)	**(0.221)**	**(0.271)***	**(0.273)****
Panel B: *Humiliation items*				
Disrespect	0.232	0.119	0.366	0.367
	(0.174)	(0.179)	(0.263)	(0.262)
	(0.174)	**(0.179)**	**(0.263)**	**(0.262)**
Unfairness	0.009	0.049	0.499	0.563
	(0.190)	(0.201)	(0.249)**	(0.249)**
	(0.190)	**(0.201)**	**(0.249)**	**(0.249)***
Discrimination	0.028	0.115	0.433	0.472
	(0.183)	(0.200)	(0.228)*	(0.229)**
	(0.183)	**(0.200)**	**(0.228)**	**(0.229)***
Control variables	No	Yes	No	Yes
Polynomial degree	1	1	1	1
Observations	917	917	560	560

Notes: This table shows parametric RDD impact estimates for the individual items comprising the main aggregated shame and humiliation scales. Items are standardized with mean zero and standard deviation of one. Estimates follow the linear fuzzy specification of Equation 8. Columns (1) and (2) show impact estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth, while columns (3) and (4) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of their respective standardized running variable. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC* are shown in parenthesis. Romano-Wolf multiple hypotheses adjusted standard p-values are marked in bold. *** p<0.01, ** p<0.05, * p<0.10. Figures A13 and A14 show the same analysis in graphical form for each program. Table A5 reports results for those items that were excluded from the main scale.

Table 5:
RDD Impact Estimates of AFAM-PE and TUS on Additional Perceptions and Opinions

	AFAM-PE				TUS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: *Other outcomes*								
Perceived position in income distribution	0.328	0.174	0.085	-0.047	-0.698*	-0.723*	-0.025	-0.021
	(0.280)	(0.294)	(0.673)	(0.413)	(0.404)	(0.397)	(0.973)	(0.636)
Observations	917	917	917	917	560	560	560	560
Recipients should feel ashamed of themselves	-0.007	0.008	-0.042	-0.044	-0.027	-0.034	-0.009	-0.023
	(0.043)	(0.044)	(0.062)	(0.065)	(0.032)	(0.033)	(0.054)	(0.051)
Observations	905	905	905	905	555	555	555	555
Grade of support towards the program	0.016	-0.123	0.039	-0.084	-0.264	-0.275	-0.451	-0.493
	(0.160)	(0.171)	(0.213)	(0.227)	(0.274)	(0.276)	(0.457)	(0.430)
Observations	878	878	878	878	554	554	554	554
Ashamed of appearance	0.105	0.113	0.228**	0.248**	0.055	0.056	-0.029	-0.008
	(0.078)	(0.075)	(0.109)	(0.114)	(0.109)	(0.109)	(0.178)	(0.165)
Observations	915	915	915	915	556	556	556	556
Panel B: *Life satisfaction outomes*								
Life satisfaction	-0.188	-0.334	-0.310	-0.392	-0.504	-0.659	-0.998	-0.425
	(0.439)	(0.454)	(0.640)	(0.667)	(0.596)	(0.558)	(0.999)	(0.947)
Life satisfaction (1 = above median)	-0.058	-0.097	-0.176	-0.197	-0.253**	-0.290**	-0.335**	-0.359*
	(0.098)	(0.102)	(0.143)	(0.150)	(0.120)	(0.119)	(0.205)	(0.193)
Observations	916	916	916	916	560	560	560	560
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial degree	1	1	2	2	1	1	2	2

Notes: This table shows parametric RDD impact estimates for other subjective outcomes regarding perceptions and opinions of individuals. Perceived position in income distribution reports results for the following question: "*Imagine a scale from 1 to 10, where in 1 are the poorest and in 10 the richest people, where do you place yourself?*". Recipients should feel ashamed of themselves refers to answers in response to: "*Do you agree with the statement that people who receive AFAM-PE (or TUS) should be ashamed of themselves?*". Possible answers are 0 (No) or 1 (Yes). Meanwhile, Grade of support towards the program references two different questions. For AFAM-PE: "*Do you think that AFAM-PE benefits should be provided less in cash and a part should be given through a food card? (It is always the same money)*" The answer scale goes from 1 (strongly disagree) to 5 (strongly agree). For TUS: '*Do you think that TUS is a...?*". The possible responses range from 1 (very bad benefit) to 5 (very good benefit). In addition, Ashamed of appearance refers to the question of "*Have you thought about not attending or have you not attended a work, family or social event during the last month because you felt you did not have the clothes or appearance required for that venue?*". Possible answers are 0 (No) or 1 (Yes). Lastly, Life satisfaction shows results for this question: "*On a scale of 1 to 10, where 1 is very dissatisfied and 10 is very satisfied, how satisfied are you in relation to to your life in general*". Columns (1) to (4) show impact estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while columns (5) to (8) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC* are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.10.

Table 6:
Robustness Test: Estimates for Alternative Bandwidth Selection

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Panel A:	*Cut 0.01 from tails*		*Cut 0.1 from tails*	
Shame z-score	0.494**	0.573***	0.671**	0.716**
	(0.217)	(0.222)	(0.290)	(0.295)
Humiliation z-score	0.223	0.192	0.609**	0.641**
	(0.203)	(0.206)	(0.264)	(0.260)
Bandwidth	[−0.03, 0.06]	[−0.03, 0.06]	[−0.3, 0.3]	[−0.3, 0.3]
Observations	811	811	515	515
Panel B:	*Cut 0.015 from tails*		*Cut 0.15 from tails*	
Shame z-score	0.539**	0.600***	0.761**	0.808**
	(0.229)	(0.231)	(0.315)	(0.321)
Humiliation z-score	0.289	0.270	0.663**	0.659**
	(0.212)	(0.216)	(0.281)	(0.274)
Bandwidth	[−0.025, 0.055]	[−0.025, 0.055]	[−0.25, 0.25]	[−0.25, 0.25]
Observations	738	738	480	480
Panel C:	*Cut 0.02 from tails*		*Cut 0.2 from tails*	
Shame z-score	0.493**	0.554**	0.882**	0.929**
	(0.248)	(0.252)	(0.365)	(0.371)
Humiliation z-score	0.311	0.288	0.907***	0.873***
	(0.232)	(0.233)	(0.338)	(0.304)
Bandwidth	[−0.02, 0.05]	[−0.02, 0.05]	[−0.2, 0.2]	[−0.2, 0.2]
Observations	567	567	405	405
Control variables	No	Yes	No	Yes
Polynomial degree	1	1	1	1

Notes: This table presents alternative estimates for the main outcome variables of shame and humiliation scales for different bandwidth selection over the running variable. Panel A shows results for a trimmed running variable with 0.01 (0.1) cuts from each tail for the AFAM-PE (TUS) sample. Meanwhile, Panel B shows results for 0.015 (0.15) cuts and Panel C shows results for 0.02 (0.2) for the AFAM-PE (TUS) sample. Estimates are obtained through regressions following the preferred specification: a linear fuzzy RDD, as described in Equation 8. All coefficients correspond to the parameter of interest β, which represent the effect of each program on each considered outcome. Columns (1) and (2) show impact estimates for the AFAM-PE, while columns (3) and (4) show estimates for the TUS program. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC* are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.10. Figure A15 shows graphical the same results in graphical form for additional cut-off trimmings.

Table 7:
Robustness Test: Donut RDD Estimates

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Panel A:	$n \notin [-0.001, 0.001]$		$n \notin [-0.01, 0.01]$	
Shame z-score	0.434**	0.566***	0.568**	0.620**
	(0.195)	(0.218)	(0.257)	(0.264)
Humiliation z-score	0.104	0.090	0.638***	0.680***
	(0.188)	(0.200)	(0.230)	(0.232)
Observations	904	904	541	541
Panel B:	$n \notin [-0.005, 0.005]$		$n \notin [-0.05, 0.05]$	
Shame z-score	0.457**	0.580**	0.767**	0.830**
	(0.214)	(0.228)	(0.327)	(0.334)
Humiliation z-score	0.079	0.043	0.643**	0.707**
	(0.217)	(0.224)	(0.299)	(0.304)
Observations	840	840	472	472
Control variables	No	Yes	No	Yes
Polynomial degree	1	1	1	1

Notes: This table presents donut RDD estimates for the main outcome variables of shame and humiliation. Panel A shows results for a trimmed running variable with 0.001 (0.01) cuts right next to the threshold of AFAM-PE (TUS). Meanwhile, Panel B shows results for 0.005 (0.05) cuts for the AFAM-PE (TUS) sample. Estimates are obtained through regressions following the preferred specification: a linear fuzzy RDD, as described in Equation 8. All coefficients correspond to the parameter of interest β, which represent the effect of each program on each considered outcome. Columns (1) and (2) show impact estimates for the AFAM-PE, while columns (3) and (4) show estimates for the TUS program. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC* are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.10. Figure A16 shows the same analysis in graphical form for the 0.005 and 0.05 cuts.

Table 8:
Robustness Test: Non-Parametric RDD Estimates

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Panel A: *Uniform kernel*				
Shame z-score	0.641**	0.686**	0.732*	0.757**
	(0.307)	(0.308)	(0.379)	(0.374)
Humiliation z-score	0.379	0.340	0.692**	0.667**
	(0.278)	(0.277)	(0.342)	(0.336)
Panel B: *Triangular kernel*				
Shame z-score	0.634**	0.679**	0.785*	0.816**
	(0.308)	(0.308)	(0.411)	(0.399)
Humiliation z-score	0.377	0.340	0.705*	0.687*
	(0.277)	(0.278)	(0.377)	(0.362)
Panel C: *Epanechnikov kernel*				
Shame z-score	0.641**	0.686**	0.799**	0.829**
	(0.307)	(0.308)	(0.400)	(0.391)
Humiliation z-score	0.379	0.340	0.729**	0.705**
	(0.277)	(0.278)	(0.364)	(0.352)
Control variables	No	Yes	No	Yes
Polynomial degree	1	1	1	1
Observations	917	917	560	560

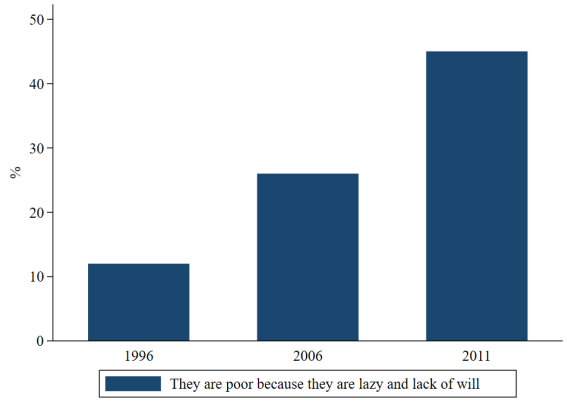
Notes: This table presents non-parametric estimates proposed by Cattaneo et al. (2018) for the main outcome variables of shame and humiliation scales. Estimates are obtained through regressions following the preferred fuzzy RDD specification, local polynomials of first order, as described in Equation 9. Estimates of β are bias-corrected, following Calonico et al. (2014). Panel A shows results using a uniform kernel specification, while Panel B shows results with a triangular kernel, and Panel C with a epanechnikov kernel. Columns (1) and (2) report impact estimates for the AFAM-PE, while columns (3) and (4) report estimates for the TUS program. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Robust standard errors (clustered by ICC*) following Calonico et al. (2014). *** p<0.01, ** p<0.05, * p<0.10.

Appendix A Appendix Figures and Tables

Appendix Figures

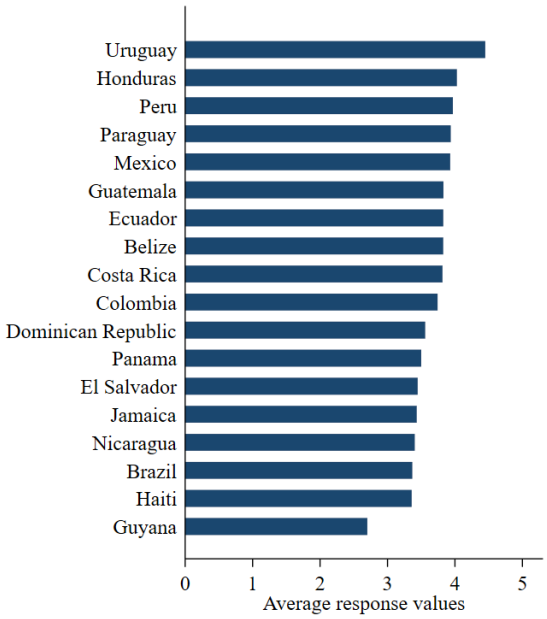
Figure A1:
Uruguayans' Beliefs Towards the Poor and Those on Welfare

(a) Why are there people in need in Uruguay?



Source: World Values Surveys (1996, 2006, 2011)

(b) Do you agree that those who are on welfare are lazy?

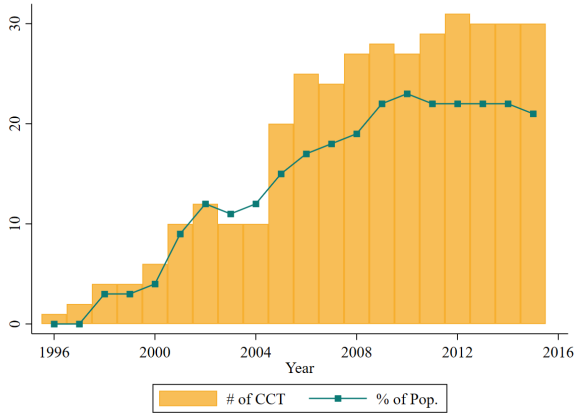


Source: Americas Barometer (2012)

Notes: This figure shows survey data about recent trends and international comparisons of beliefs of Uruguayans towards the poor and those who are on welfare. Panel (a) shows data for Uruguay from the World Value Survey (WVS) for three years between 1996 and 2011 about people´s beliefs of why there are still people in need in the country. It presents the % of people in Uruguay on the World Value Survey samples for the selected years that answer the question "*Why are there people in need in Uruguay*" with the following response: "*Because they are lazy and lack of will*". Such percentage has almost quadrupled over the 15-year period, being 12%, 26% and 45% for 1996, 2006 and 2011, respectively. Panel (b) shows data for Latin America and the Caribbean from the Americas Barometer for 2012. It displays the average answer values from each available country to this question: "*Some people say that those who receive social assistance from government programs are lazy. How much do you agree or disagree?*". Uruguay reports the highest average response (4.4) among the available countries in Latin American and the Caribbean.

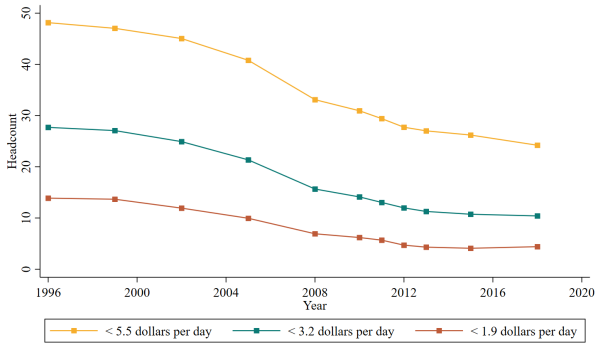
Figure A2:
Poverty and cash transfers trends in Latin American and the Caribbean

(a) Cash transfer programs and population covered by year



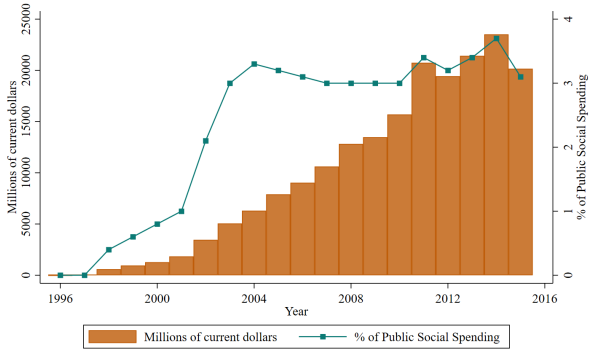
Source: Atuesta and Cecchini (2017)

(b) 1.9, 3.2 y 5.5 (2011 PPP values) poverty lines by year



Source: World Bank estimates

(c) Public expenditure in cash transfers by year



Source: Atuesta and Cecchini (2017)

Figure A3:
Electronic Card Comparison Between Both Programs

(a) AFAM–PE electronic debit card

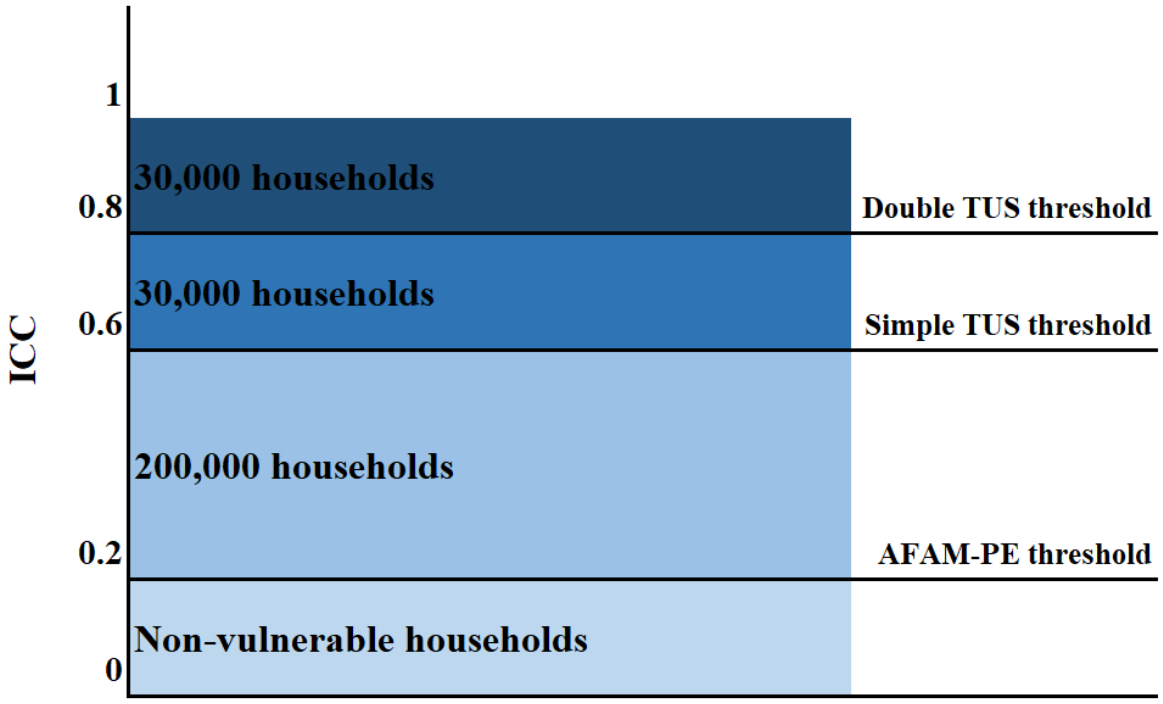
Source: Banco de la República Oriental del Uruguay

(b) TUS electronic food card

Source: Banco de la República Oriental del Uruguay

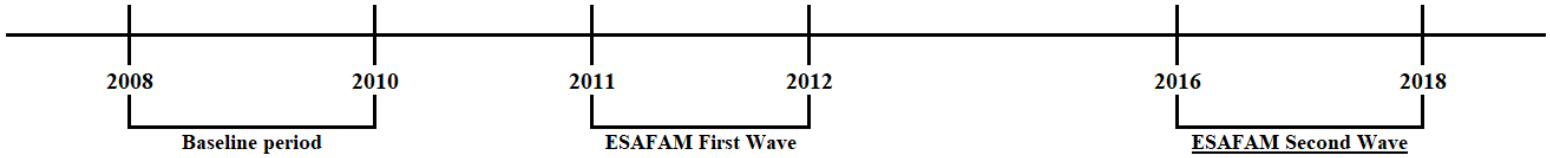
Notes: This figure shows the aspect of the electronic cards through which each cash transfer is delivered to households. Panel (a) exhibits the debit card for AFAM-PE. Panel (b) exhibits the food card for TUS.

Figure A4:
ICC Thresholds Across MIDES Cash Transfers



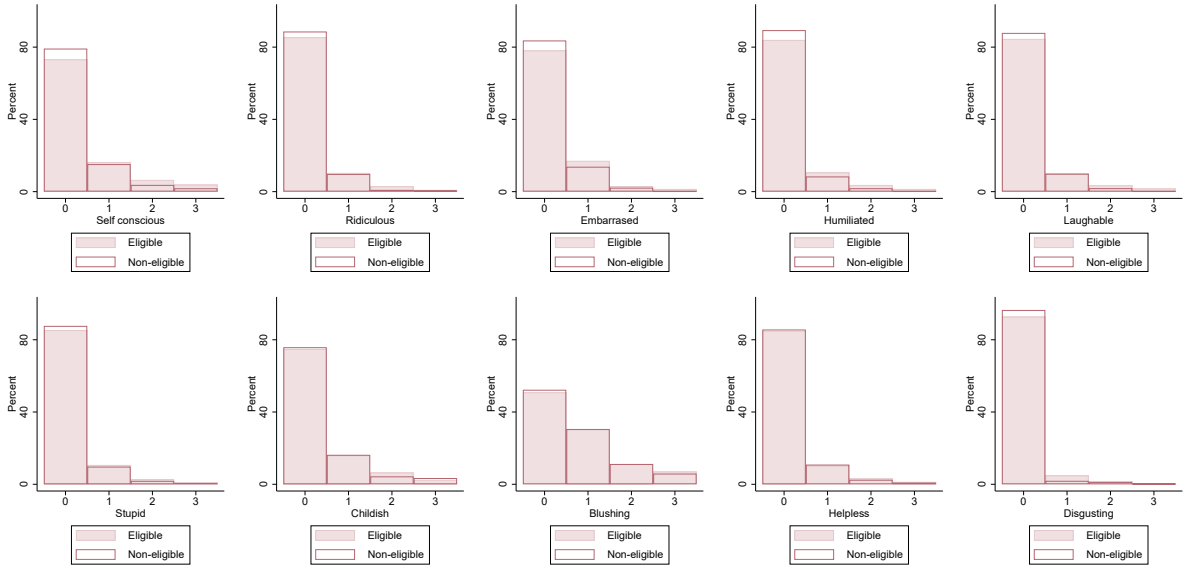
Notes: This figure exhibits the eligibility thresholds and intended amount of beneficiary households over the *Índice de Carencias Críticas* (ICC) for each of the MIDES cash transfers. Double TUS is just twice the money corresponding to a simple TUS. **Source:** Author's own elaboration based on MIDES (2013).

Figure A5:
Time-Frame of the Study



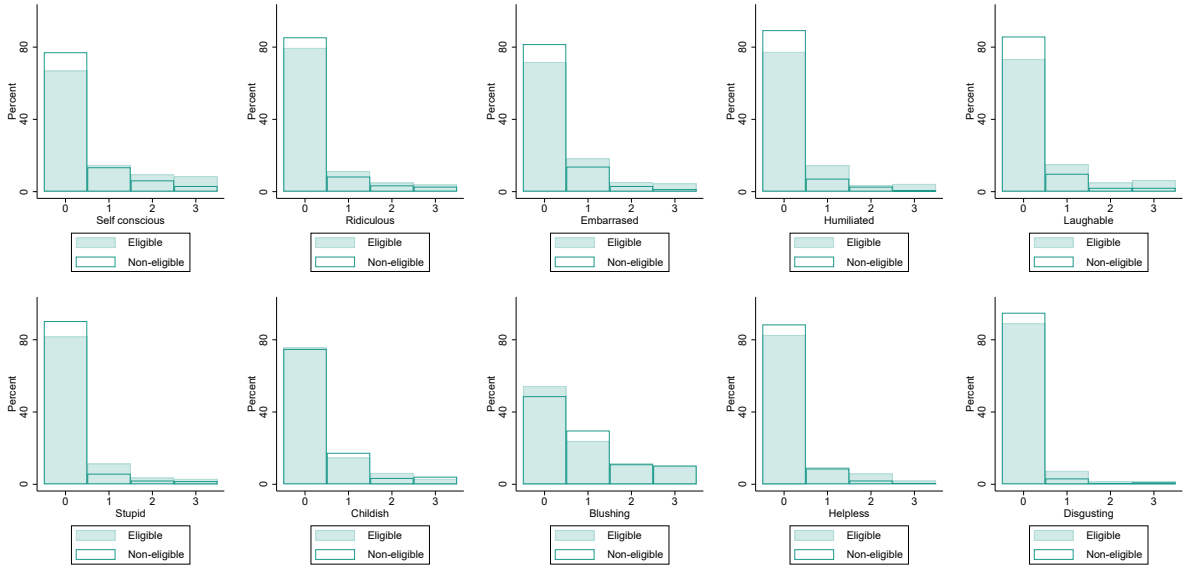
Notes: This figure exhibits the time-frame of both programs and points out the timing of the conducted surveys. The baseline period over which the empirical strategy is designed and from where the baseline administrative data is obtained lies between 2008 and 2010. Subsequently, the ESAFAM survey collected data between 2011 and 2012 during its first wave, and 2016 and 2018 during its second wave. The outcome variables used in this study are only available for the second wave, which occurs between 6 to 8 years after the baseline period. **Source:** Author's own elaboration.

Figure A6:
Histograms of responses to the items of the internal shame proneness scale:
AFAM-PE sample



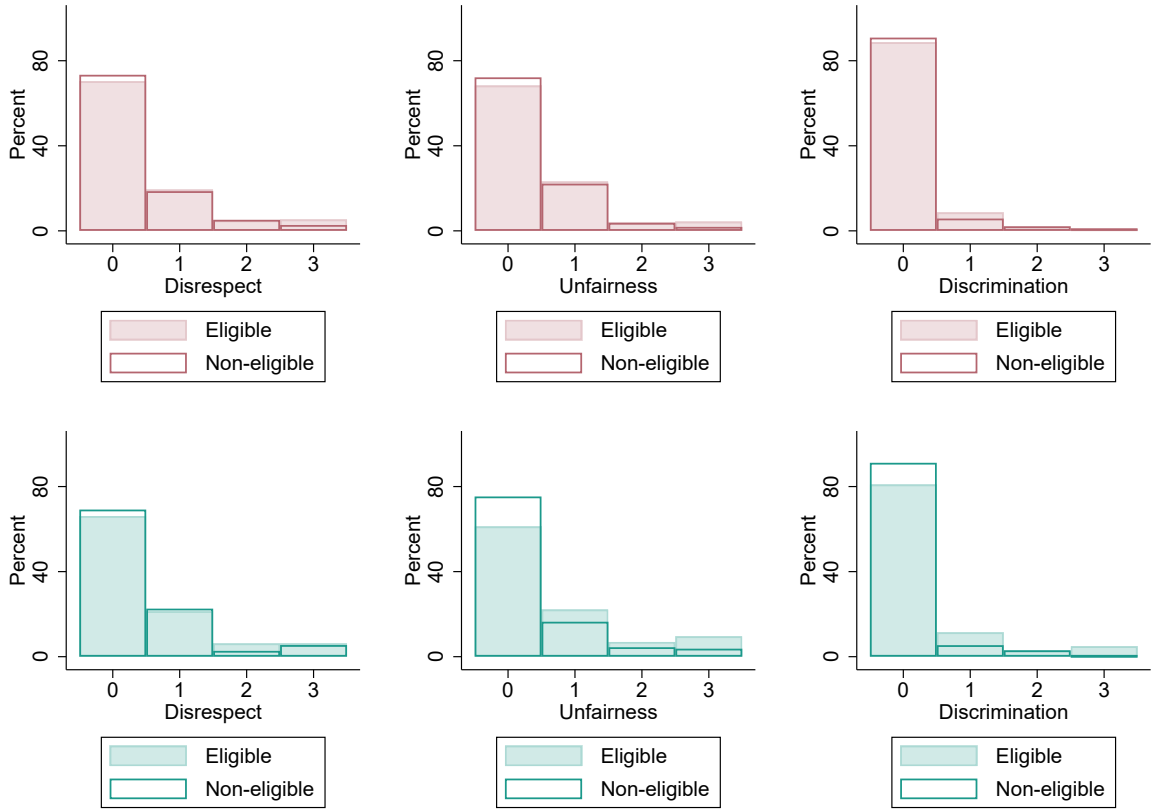
Notes: This figure shows histograms of the proportion of responses to the individual items of the internal shame proneness scale for the AFAM-PE sample, distinguishing between eligible and non-eligible groups. 0 = Rarely or Never; 1 = Occasionally; 2 = Often, 3 = Always or almost always.

Figure A7:
Histograms of responses to the items of the internal shame proneness scale:
TUS sample



Notes: This figure shows histograms of the proportion of responses to the individual items of the internal shame proneness scale for the TUS sample, distinguishing between eligible and non-eligible groups. 0 = Rarely or Never; 1 = Occasionally; 2 = Often, 3 = Always or almost always.

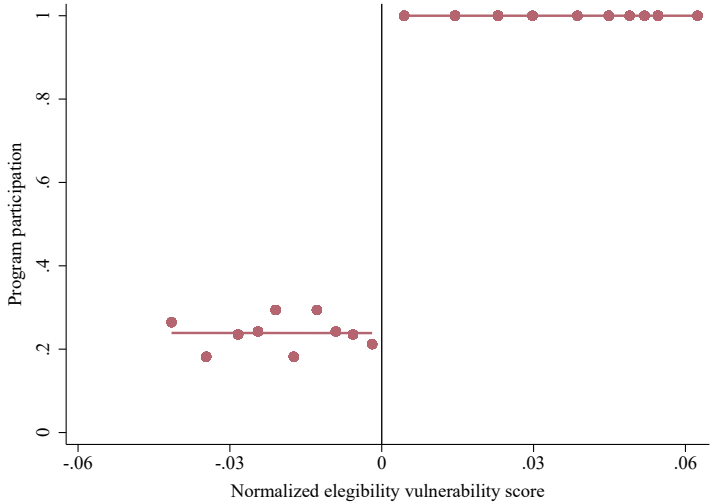
Figure A8:
Histograms of responses to the items of the external humiliation scale:
AFAM-PE and TUS sample



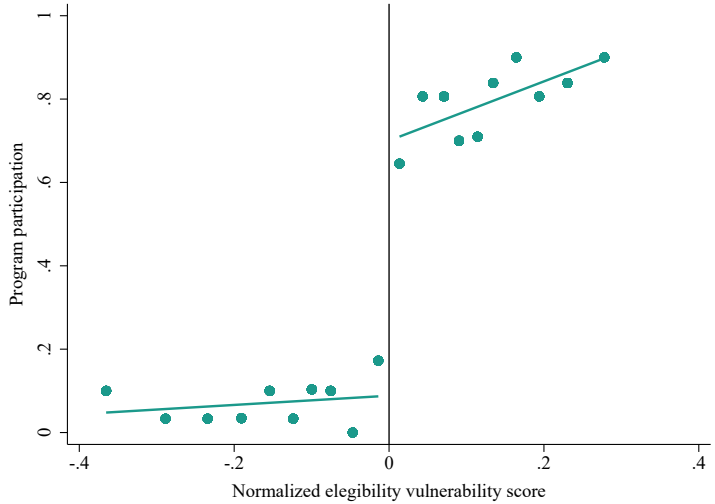
Notes: This figure shows histograms of the proportion of responses to the individual items of the external humiliation scale for both the AFAM-PE (red) and TUS sample (green), distinguishing between eligible and non-eligible groups in each program. 0 = Rarely or Never; 1 = Occasionally; 2 = Often, 3 = Always or almost always.

Figure A9:
First Stage Estimates: Participation and Eligibility According to the Vulnerability Score
Sub-sample analyses

(a) AFAM–PE [Women sub-sample]

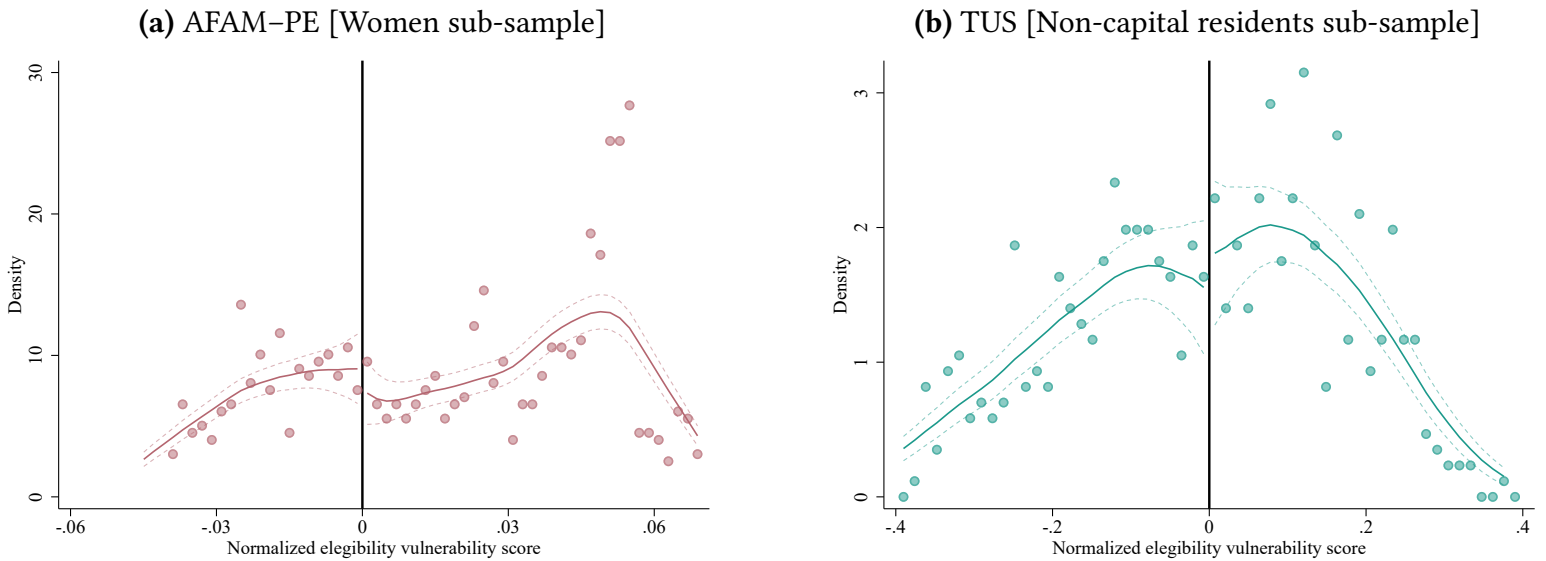


(b) TUS [Non-capital residents sub-sample]



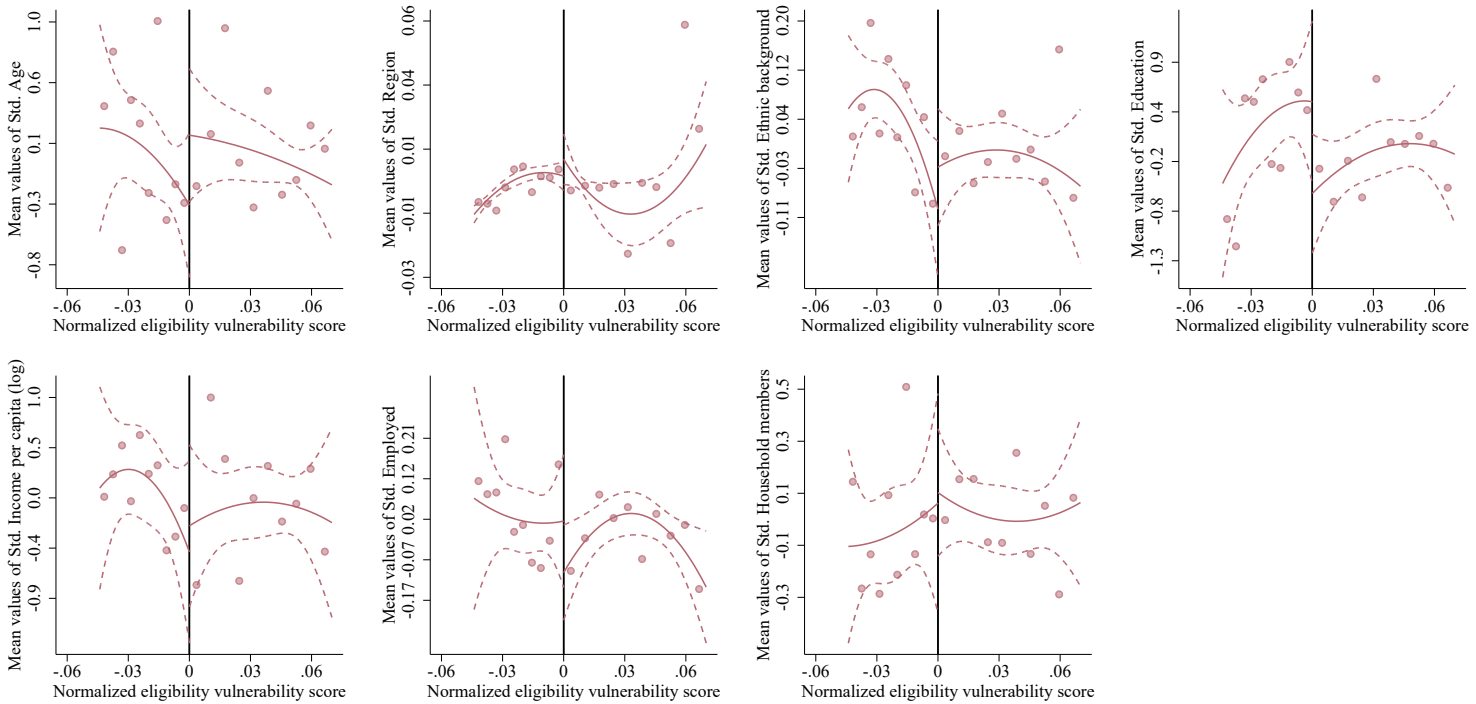
Notes: These figures show first stage estimates of each program's effective participation according to the standardized eligibility vulnerability score, *Índice de Carencias Críticas* (ICC*), for sub-samples of the main analysis. The dots represent the share of participants in the program over ten bin intervals at each side of the ICC* cut-offs at zero, represented by the vertical black lines. The lines correspond to linear estimated fits. Panel (a) displays results for the AFAM-PE women sub-sample, while Panel (b) displays results for the non-capital residents TUS sub-sample for [-0.047; 0.073] and [-0.400 ; 0.400] bandwidths of each standardized running variable, respectively. The value of the jump at the cut-off which reflects the probability of effectively participating in the program according its eligibility is around 0.74 and 0.70, for AFAM-PE and TUS respectively. The F-Test values for these first stages are 296.28 and 83.68

Figure A10:
McCrary Density Test: Bunching Around the Eligibility Threshold
Sub-Sample Analyses



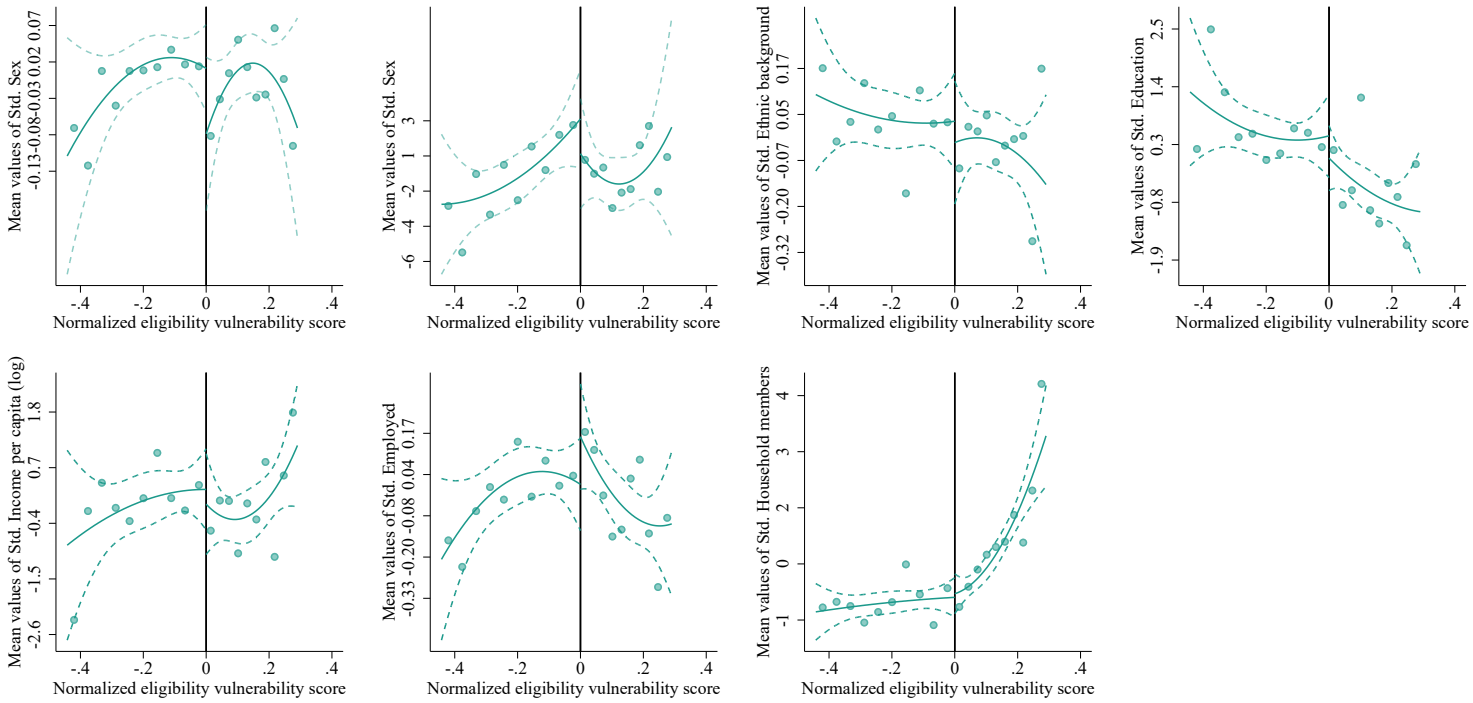
Notes: These figures show the McCrary (2008) density test for the sub-sample anal to check for the presence of systematic bunching at the right side of the cut-off of each program, represented by the vertical black lines at zero. The solid thick line indicates the density estimate, computed from a local linear regression with separate trends for each side of the threshold. The dashed lines represent 95% confidence intervals. Panel (a) displays results for the AFAM-PE sample, while Panel (b) displays results for the TUS sample for [-0.047; 0.073] and [-0.400 ; 0.400] bandwidths of each standardized running variable, respectively. The figures look really similar to those from Figure 3. The tests reject the null hypothesis that there is any bunching next to the cut-offs, and thus provide evidence of non-manipulation of the running variable. The same conclusion is reached when considering other newer manipulation tests: the p-values of the tests for each program are 0.493 and 0.835 for the one proposed by Cattaneo et al. (2018) and 0.576 and 0.999 for the one proposed by Bugni and Canay (2021), for AFAM-PE and TUS respectively.

Figure A11:
Balance Test Around the Eligibility Threshold
AFAM–PE Women Sub-Sample



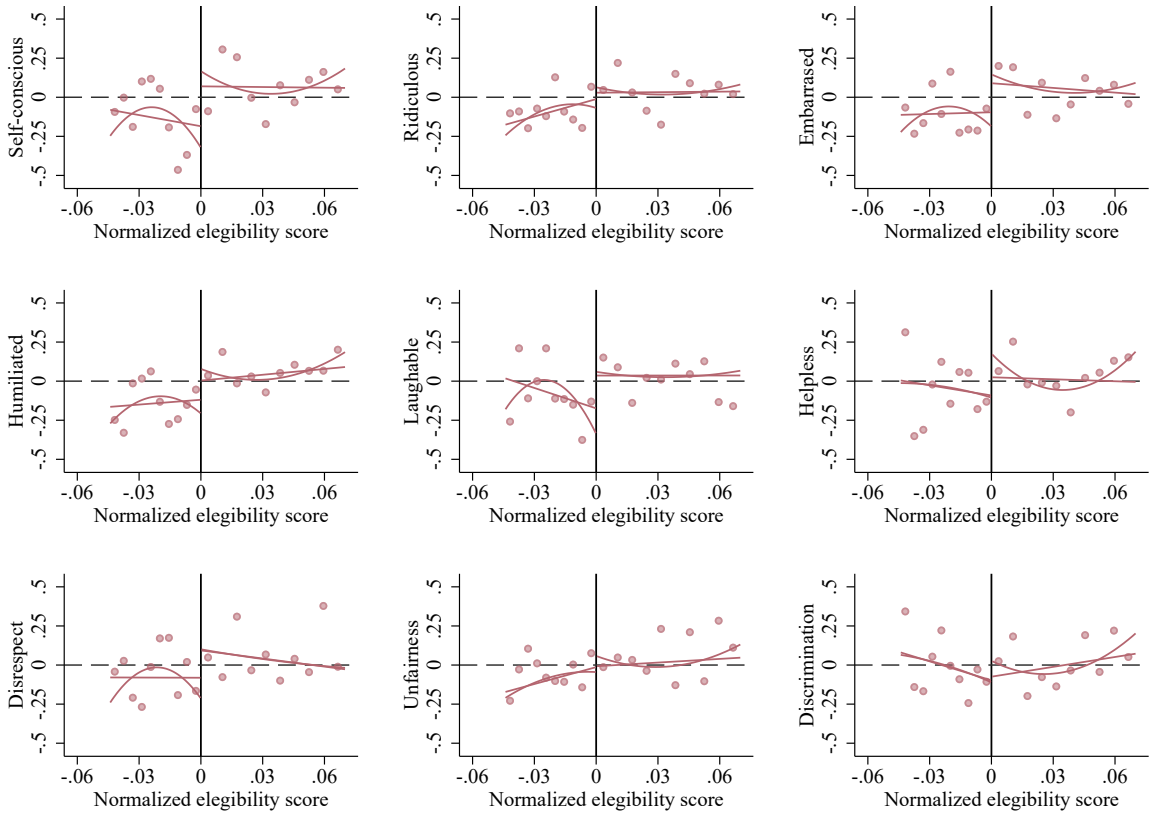
Notes: This figure exhibits graphically the balance test displayed in Table 2 for the AFAM-PE program, which checks for systematic imbalances of covariates at each side of the cut-off. It shows RDD estimates using baseline covariates as dependent variables for a [-0.047; 0.073] bandwidth of its standardized running variable. The dots represent their mean standardized value over ten bin intervals at each side of the ICC* cut-off, which is outlined with the vertical bar at zero. The thick line represents a quadratic fit. The dashed lines represent 90% confidence intervals. Control variables include: region (Montevideo = 1) and ethnicity (white = 1), each being taken out when in the presence of collinearity, e.g region is taken out when it's being evaluated as an outcome. Standard errors clustered by ICC*.

Figure A12:
Balance Test Around the Eligibility Threshold
TUS Non-Capital Residents Sub-Sample



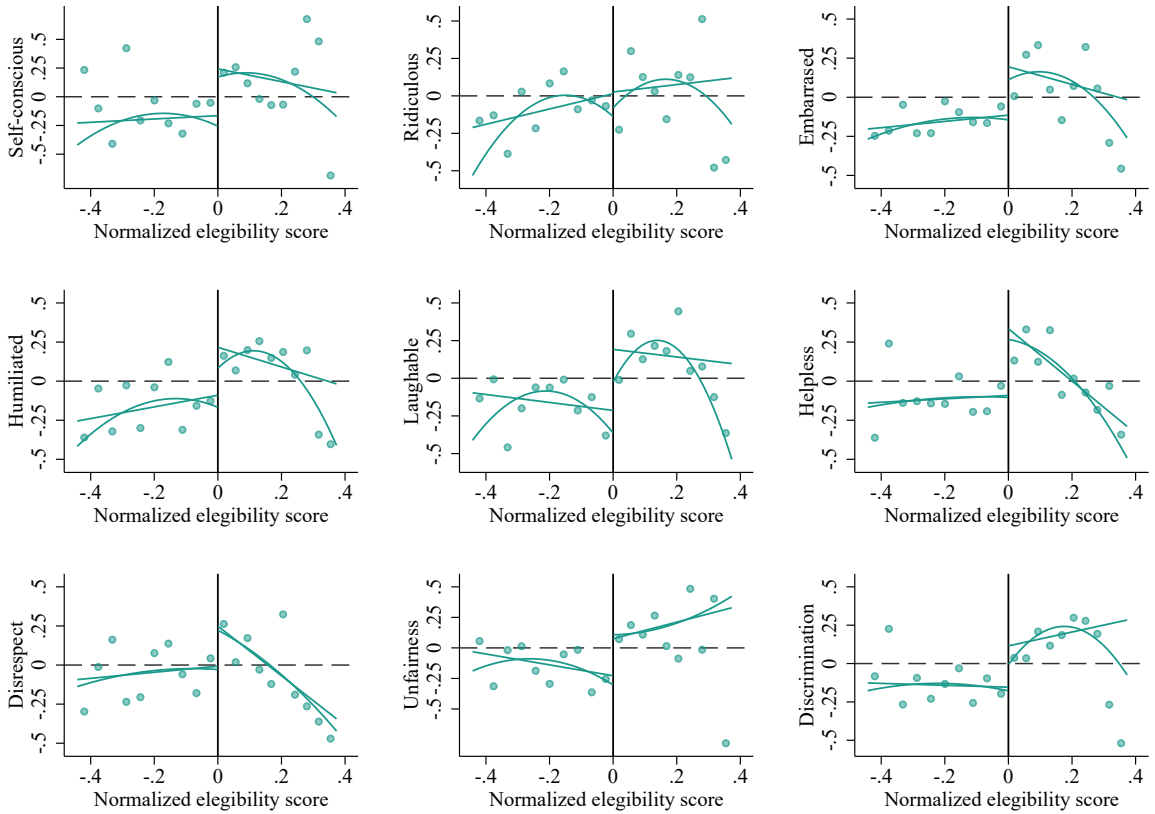
Notes: This figure exhibits graphically the balance test displayed in Table 2 for the TUS program, which checks for systematic imbalances of covariates at each side of the cut-off. It shows RDD estimates using baseline covariates as dependent variables for a [-0.040; 0.040] bandwidth of its standardized running variable. The dots represent their mean standardized value over ten bin intervals at each side of the ICC* cut-off, which is outlined with the vertical bar at zero. The thick line represents a quadratic fit. The dashed lines represent 90% confidence intervals. Control variables include: sex (woman = 1), age and ethnicity (white = 1), each being taken out when in the presence of collinearity, e.g sex is taken out when it's being evaluated as an outcome. Standard errors clustered by ICC*.

Figure A13:
RDD Impact Estimates of AFAM-PE on Individual Item Outcomes



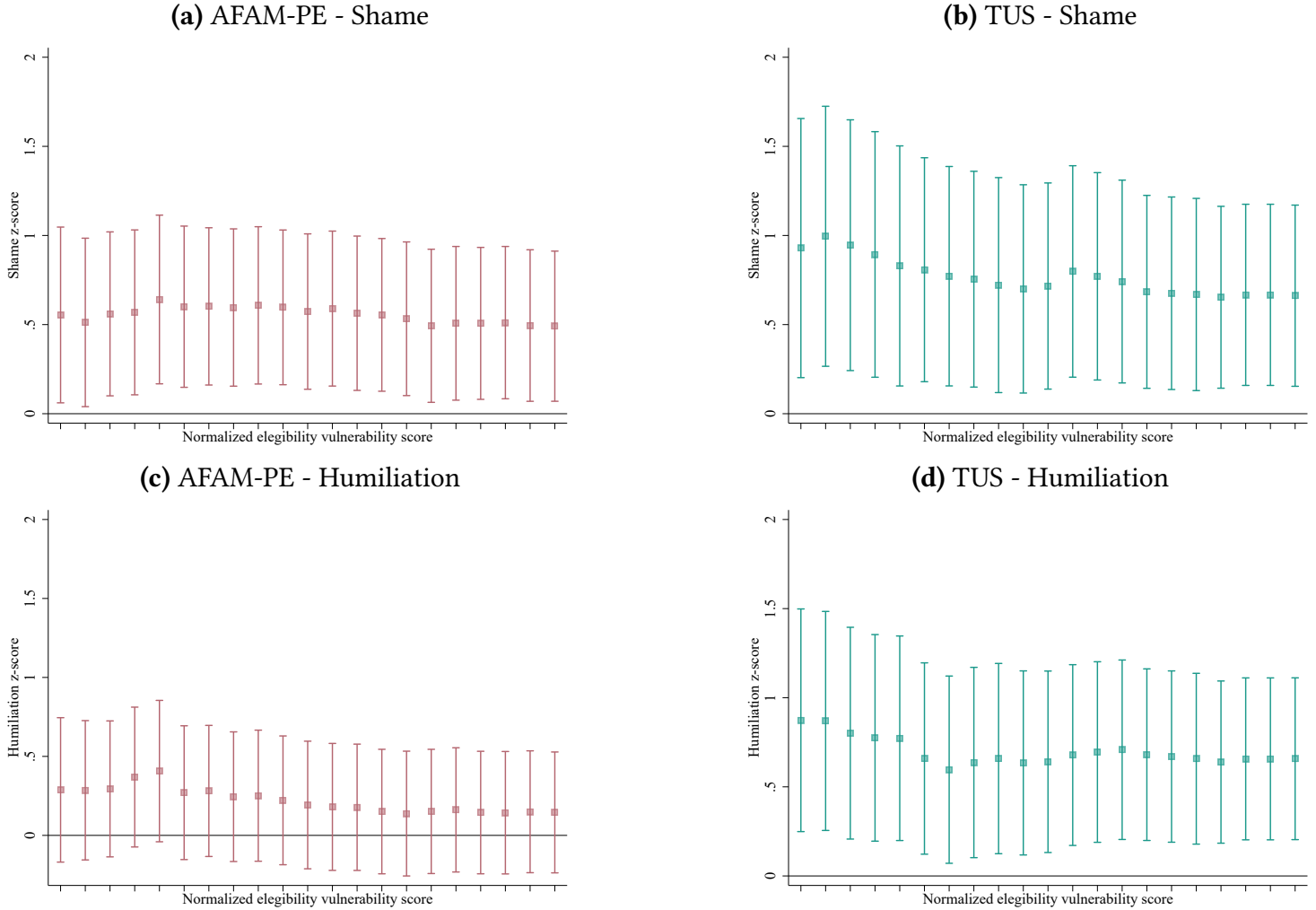
Notes: These figures show parametric Sharp RDD impact estimates for the individual item z-scores of the aggregated shame and humiliation scales. The dots represent the mean value of the items over ten bin intervals at each side of the ICC* cut-off, which is outlined with the vertical bar at zero. The lines represent linear and quadratic fits, respectively. Control variables include: sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC*. Table 4 shows the same analysis in regression form

Figure A14:
RDD Impact Estimates of TUS on Individual Item Outcomes



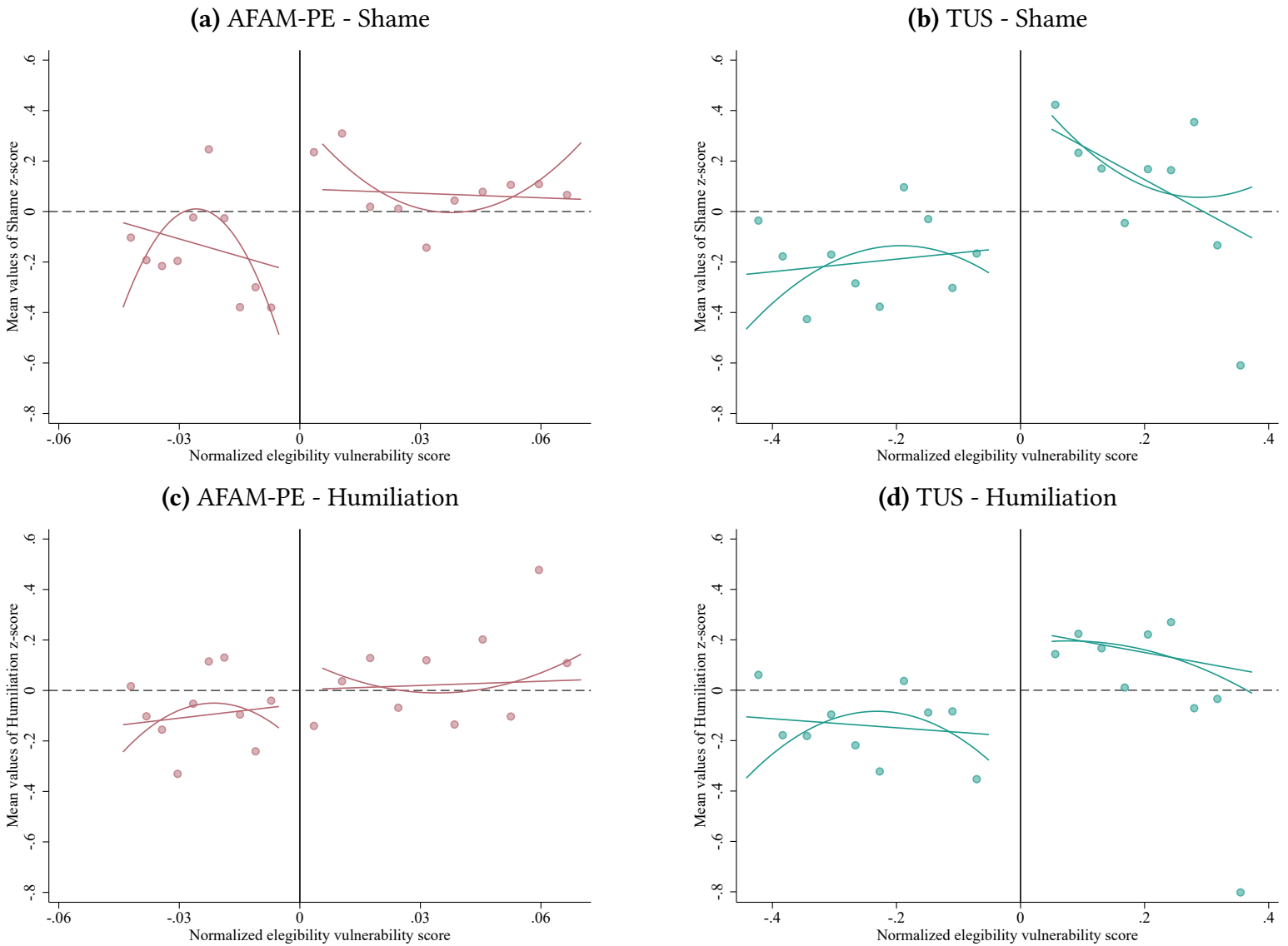
Notes: These figures show parametric Sharp RDD impact estimates for the individual item z-scores of the aggregated shame and humiliation scales. The dots represent the mean value of the items over ten bin intervals at each side of the ICC* cut-off, which is outlined with the vertical bar at zero. The lines represent linear and quadratic fits, respectively. Control variables include: sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC*. Table 4 shows the same analysis in regression form

Figure A15:
Robustness Test: Estimates for Alternative Bandwidth Selection



Notes: These figures present alternative estimates for the main outcome variables of shame and humiliation for different bandwidth selection over the normalized running variable of each program. They plot the coefficients of interest β, which represent the effect of each program on each outcome. Each estimate is computed by trimming the tails of the running variable by making progressive 0.001 (0.01) cuts from each side for the AFAM-PE (TUS) sample. The graph reads from right to left. In each panel, the first coefficient is the baseline estimate without any trimming. Then, from right to left each estimate trims an additional 0.001 (0.01) cut until reaching a 0.02 (0.2) cut at each tail of the normalized running variable of each program. Coefficients are obtained through regressions following the preferred specification: a linear Fuzzy RDD with controls, as described in Equation 8. Panels (a) and (c) show estimates for the AFAM-PE program while Panels (b) and (d) show estimates for the TUS program. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC*. 95% confidence intervals are displayed.

Figure A16:
Robustness Test: Donut RDD Estimates



Notes: These figures show donut RDD estimates for the main outcome variables of shame and humiliation with a 0.005 (0.05) cut around the AFAM-PE (TUS) threshold. The dots represent the mean value of the shame (humiliation) z-score index over ten bin intervals at each side of the ICC* cut-off, which is marked with a vertical bar at zero. The lines represent linear and quadratic estimated fits, respectively. Panels (a) and (c) show estimates for the AFAM-PE program, while Panels (b) and (d) show estimates for the TUS program. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC*. Table 7 shows the same analysis in regression form.

Appendix Tables

Table A1:
Monthly TUS amounts as of January 2023

Number of Children	Uruguayan pesos ($)		US dollars (US$)	
	Simple TUS	Double TUS	Simple TUS	Double TUS
0 to 1 children	955	1910	33	66
2 children	1448	2896	51	102
3 children	1843	3686	65	130
4 or more children	2567	5134	90	180

Notes: This table shows the monthly values as of January 2017 for the amount of TUS transfers, which depend on the number of children in the beneficiary households. **Source:** MIDES (2017)

Table A2:
Main Survey Outcomes Regarding Stigma, Shame and Humiliation

	For each of the following listed feelings, please place a number from 1 to 4 reflecting how frequent the feeling is for you	AFAM-PE Obs. (1)	AFAM-PE Mean (2)	AFAM-PE Std. Dev. (3)	TUS Obs. (4)	TUS Mean (5)	TUS Std. Dev (6)
Panel A: *Shame indicators*							
Self-conscious	Feeling self-conscious	965	0.368	0.733	592	0.478	0.874
Ridiculous	Feeling ridiculous	952	0.176	0.499	587	0.286	0.703
Embarrassed	Feeling embarrassed	966	0.251	0.559	593	0.337	0.701
Humiliated	Feeling humiliated	966	0.196	0.537	593	0.250	0.635
Laughable	Feeling laughable	944	0.198	0.543	579	0.328	0.744
Stupid	Feeling stupid	951	0.183	0.507	587	0.216	0.576
Childish	Feeling childish	961	0.362	0.719	585	0.368	0.740
Blushing	Feeling blushed	969	0.733	0.908	589	0.801	0.999
Helpless	Feeling helpless, paralyzed	961	0.203	0.545	587	0.213	0.576
Disgusting	Feeling disgusting to others	937	0.085	0.374	572	0.121	0.464
	Have you felt that you have been [...] during the last three months?						
Panel B: *Humiliation indicators*							
Disrespect	Treated without respect	982	0.426	0.783	596	0.487	0.833
Unfairness	Treated unfairly	967	0.415	0.731	594	0.51	0.877
Discrimination	Treated with discrimination	985	0.142	0.459	600	0.222	0.624

Notes: This table exhibits the questions of interest developed by Zavaleta (2007) regarding stigma, shame and humiliation feelings in contexts of poverty. 0 = Rarely or Never; 1 = Occasionally; 2 = Often, 3 = Always or almost always. It also presents descriptive statistics for the AFAM-PE and TUS samples, separately. Figures A6, A7, and A8 display the proportion of responses for each individual item, differentiating between the AFAM-PE and TUS samples, and eligible and non-eligible groups within each program.

Table A3: Descriptive Statistics for Main Aggregated Outcomes

	Observations	Mean	Std. Deviation	Minimum	Maximum	Non-Eligible	Eligible	Mean Diff. Test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: AFAM-PE								
A.1: Summarized								
Shame sum	917	1.373	2.304	0.0	17.0	1.062	1.528	0.466***
Humiliation sum	917	0.965	1.312	0.0	8.0	0.850	1.023	0.173*
A.2: Standardized								
Shame z-score	917	-0.003	1.0	-0.599	6.786	-0.138	0.064	0.202***
Humiliation z-score	917	-0.015	0.981	-0.737	5.245	-0.102	0.028	0.130*
Panel B: TUS								
B.1: Summarized								
Shame sum	560	1.877	2.973	0.0	18.0	1.333	2.394	1.061***
Humiliation sum	560	1.239	1.700	0.0	9.0	0.960	1.510	0.551***
B.2: Standardized								
Shame z-score	560	-0.005	1.001	-0.637	5.426	-0.188	0.170	0.357***
Humiliation z-score	560	0.010	1.016	-0.730	4.645	-0.157	0.169	0.326***

Notes: This table shows summary statistics for the main aggregated outcomes of interest regarding shame and humiliation. For each index, statistics for either the standardized and non-standardized version are displayed. Panel A shows statistics for the AFAM-PE program, while Panel B shows descriptive data for the TUS program. The last three columns show the mean values for eligible and non-eligible individuals in each program, and a difference in means test between both averages. *** p<0.01, ** p<0.05, * p<0.10. Figure 1 displays kernel distributions of these indexes. Table A2 exhibits the questions from the individual items these indexes contain.

Table A4:
Other Subjective Outcomes: individual perceptions and opinions

	AFAM-PE			TUS		
	Obs. (1)	Mean (2)	Std. Dev. (3)	Obs. (4)	Mean (5)	Std. Dev (6)
Perceived position in income distribution	917	4.230	1.436	560	3.705	1.617
Recipients should feel ashamed of themselves	905	0.049	0.215	555	0.032	0.177
Grade of support towards the program	889	2.708	1.126	552	4.062	0.728
Ashamed of appearance	915	0.149	0.357	556	0.288	0.453
Life satisfaction	916	6.509	2.293	560	6.259	2.422
Life satisfaction (1 = above mode)	916	0.511	0.500	560	0.548	0.498

Notes: This table shows summary statistics for the other subjective variables regarding perceptions and opinions of individuals that are considered as additional outcomes.

Table A5:
RDD Impact Estimates of AFAM-PE and TUS on Excluded Items

	AFAM-PE		TUS	
	(1)	(2)	(3)	(4)
Excluded shame items				
Stupid	-0.094	0.066	0.134	0.125
	(0.184)	(0.182)	(0.227)	(0.231)
	(0.184)	**(0.182)**	**(0.227)**	**(0.231)**
Childish	-0.266	-0.113	0.240	0.297
	(0.203)	(0.196)	(0.240)	(0.238)
	(0.203)	**(0.196)**	**(0.240)**	**(0.238)**
Blushing	-0.045	-0.021	0.032	0.017
	(0.182)	(0.187)	(0.252)	(0.249)
	(0.182)	**(0.187)**	**(0.252)**	**(0.249)**
Disgusting	0.220	0.282	0.499	0.534
	(0.186)	(0.192)	(0.202)**	(0.207)***
	(0.186)	**(0.192)**	**(0.202)***	**(0.207)****
Control variables	No	Yes	No	Yes
Polynomial degree	1	1	1	1
Observations	917	917	560	560

Notes: This table shows parametric RDD impact estimates for the individuals items that were excluded from the main shame aggregated index in the PCA analysis. Items are standardized with mean zero and standard deviation of one. Estimates follow the linear fuzzy specification of Equation 8. All coefficients correspond to the parameter of interest β, which represent the effect of each program on each considered outcome. Columns (1) and (2) show impact estimates for the AFAM-PE program for a [-0.047; 0.073] bandwidth of its standardized running variable, while columns (3) and (4) show estimates for the TUS program for a [-0.400 ; 0.400] bandwidth of its standardized running variable. Control variables include sex (woman = 1), age, region (Montevideo = 1) and ethnicity (white = 1). Standard errors clustered by ICC* are shown in parenthesis. Romano-Wolf multiple hypotheses adjusted standard p-values are marked in bold. *** p<0.01, ** p<0.05, * p<0.10.

Appendix B Principal Component Analysis (PCA)

This appendix details the construction of the aggregate index of shame proneness based on the survey module proposed by Zavaleta (2007). Following Filmer and Pritchett (2001) and Osborne (2008), the Principal Component Analysis (PCA) method is carried out for purposes of data reduction of related items regarding latent constructs.

The PCA is based on the estimation of the following system of equations and its main objective is to determine which items to consider for the latent constructs to be constructed and which weighting to give to each of these items:

$$y_{ij} = \lambda_{i1}F_{1j} + \lambda_{i2}F_{2j} + \ldots + \lambda_{iN}F_{Nj} \qquad j = 1, \ldots, N \qquad i = 1, \ldots, N$$

Where y_{ij} are the items, $_{ij}$ the latent constructs, λ_{ij} the regression coefficients or as they are known in this literature 'loadings', and μ_{ij} the error terms that are assumed to be independent of the latent constructs.

In first place, the analysis begins with all the items belonging to the shame scale shown in Table A2. Then their polychoric correlation matrix is first calculated and the PCA is performed on it. The results of this first phase of the PCA are presented in Tables B1 and B2. The first table shows that two factors are formed with eigenvalues greater than the unit value, of which the first one explains 51.4% of the total variance of the set of items. Furthermore, compute goodness-of-fit indicators are computed in order to assess the strength of correlation between the items and thus the internal validity of the construct. The indicator are the Kaiser-Meyer-Olkin (KMO) and Cronbach Alpha tests. Both range from 0 to 1 and the higher their score the better the adequacy of the items to form a common scale. In this case, these indicators take values of 0.87 and 0.77, respectively, which are considered fairly high in the literature (Yin and Etilé, 2019).

The second table shows which items load in Component 1 and the magnitude of this "loading". Component 1, which refers to the shame proneness construct, gathers its corresponding items, with the exception of the items stupid, childish and blushing, which do not have loadings higher than 0.3, the threshold usually used in the literature, and therefore are not considered to load on either component, following Yin and Etilé (2019). In addition, the item regarding the emotion of feeling stupid is dropped since it does not directly correlate to the context of poverty and welfare recipiency in which the questionnaire is being used. This is reflected by the fact that this variable's inter-item correlation is never higher than 0.14 with respect to the rest of the items. Moreover, after dropping this variable, both the average inter-item correlation and the proportion of the variance explained by the principal component (62%) increase.

Table B1: Principal components: initial eigenvalues and explained variance

Component	Eigenvalue	Difference	Proportion	Cumulative
Component 1	5.144	4.092	0.514	0.514
Component 2	1.052	0.320	0.105	0.620
Component 3	0.732	0.076	0.073	0.693

Notes: This table shows the results of a principal component analysis performed on the polychoric correlation matrix of the items, restricting the number of components to two. The first four components with eigenvalues higher than the unite value are shown. The first component explains 54.4% of the total variance.

Table B2: Component loadings of initial items

Variable	Component 1	Uniqueness
Self-conscious	0.322	0.466
Ridiculous	0.370	0.294
Embarrassed	0.339	0.408
Humiliated	0.345	0.375
Laughable	0.315	0.491
Stupid	0.361	0.329
Childish		0.768
Blushing		0.682
Helpless	0.330	0.439
Disgusting		0.604

Notes: This table shows the components loadings and uniqueness obtained by means of a principal component analysis performed on the polychoric correlation matrix of the items, restricting the number of components to two. The component loadings represent the relationship between each of the items and the latent constructs. Following Yin and Etilé (2019) rule, only those loadings with values greater than 0.3 are shown. As for the uniqueness column, it represents the proportion of variance that is specific to an item and is not shared by the latent constructs.

Given the previous results, the PCA is now repeated but discarding the items that have low loadings. This new analysis increases the variance explained by the first component up to 62.0%, as shown in Table B3. Also, this new set keep fairly high KMO and Cronbach Alpha values of 0.83 and 0.77 respectively. As for the loadings of the items, Table B4 shows that all items loads in Component 1 with loadings higher than 0.3.

Table B3: Principal components: final eigenvalues and explained variances

Component	Eigenvalue	Difference	Proportion	Cumulative
Component 1	3.712	3.088	0.620	0.620
Component 2	0.632	0.073	0.105	0.725

Notes: This table shows the results of a principal component analysis performed on the polychoric correlation matrix of the items, restricting the number of components to two. The first four components with eigenvalues higher than the unite value are shown. The first component explains 62.0% of the total variance.

Table B4: Component loadings of final items

Variable	Component 1	Uniqueness
Self-conscious	0.393	0.426
Ridiculous	0.447	0.256
Embarrassed	0.393	0.425
Humiliated	0.428	0.425
Laughable	0.379	0.318
Helpless	0.406	0.467

Notes: This table shows the components loadings and uniqueness obtained by means of a principal component analysis performed on the polychoric correlation matrix of the items, restricting the number of components to two. The component loadings represent the relationship between each of the items and the latent constructs. Following Yin and Etilé (2019) rule, only those loadings with values greater than 0.3 are shown. As for the uniqueness column, it represents the proportion of variance that is specific to an item and is not shared by the latent constructs.

Based on the above, the shame proneness index is constructed. Although the index can be built using different aggregation methods (addition, average, PCA, FA, PCA with rotated weights, FA with rotated weights), the simpler one, the sum (standardized), is finally chosen for the main analysis. Nevertheless, results found are robust to other alternatives.