Welfare Analysis of Changing Notches: Evidence from Bolsa Família*

Katy Bergstrom†, William Dodds‡, and Juan Rios§

February 8, 2022

Abstract

This paper develops a framework to bound the welfare impacts of reforms that change notches in transfer schedules. We show that we can construct bounds from two empirical objects: (1) the number of households bunching at the old notch who move toward the new notch, and (2) the number of households who "jump" down to the new notch as a result of the reform. Importantly, our bounds hold in a wide class of models; thus, we refer to our two empirical objects as "sufficient statistics". We estimate these statistics using a generalized difference-in-difference strategy in the context of a reform to the Brazilian anti-poverty program, Bolsa Família. We find that the reform's marginal value of public funds is bounded between 0.90 and 1.12; thus, the reform was welfare improving (decreasing) if the government values giving \$0.90 (\$1.12) to eligible households more (less) than spending one dollar on their best alternative.

Keywords: bunching, jumping, notches, MVPF, bounds

*Some of the ideas in this paper were part of a chapter of Juan's dissertation at Stanford University, titled "Welfare Analysis of Transfer Programs with Jumps in Reported Income: Evidence from the Brazilian Bolsa Família". We are grateful to Douglas Bernheim, Raj Chetty, and Florian Scheuer for their advice and guidance. We also want to thank Juan Camilo Castillo, Dave Donaldson, Alex Gelber, Matthew Gentzkow, Han Hong, Caroline Hoxby, Melanie Morten, Muriel Niederle, Luigi Pistaferri, Petra Persson, Emmanuel Saez, Joel Slemrod, and seminar participants at Stanford, University of Michigan, Pompeu Fabra, CEMFI, PUC Rio, INSPER, University of Auckland, The World Bank, and FGV, and three anonymous referees for comments and suggestions. Financial support from the B.F. Haley and E.S. Shaw Fellowship Fund, the Shultz Fellowship, and the Stanford EDGE Initiative are gratefully acknowledged. This research was supported in part using high performance computing (HPC) resources and services provided by Technology Services at Tulane University, New Orleans, LA. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

†World Bank. Email: kbergstrom@worldbank.org.

‡Tulane University. Email: wdodds@tulane.edu.

§181 Sanchez Street, 612, San Francisco CA 94114, E-mail: juanriosriver@gmail.com

1 Introduction

Throughout developed and developing countries, tax and transfer schedules often feature notches wherein incremental changes in behavior lead to discrete changes in benefits or tax liabilities. For example, in Pakistan, Kleven and Waseem (2013) document notches of up to 4 percentage points in *average* personal income tax rates; in the United States, individuals are eligible for Medicaid only if their income is below a certain threshold (Yelowitz, 1995); in Costa Rica, the average corporate tax rate jumps by 10 percentage points at discrete revenue thresholds (Bachas and Soto, 2021). Given this prevalence, it is important for economists and policymakers to understand the welfare impacts of policy reforms that change the structure of notches in tax and transfer systems.

Typically the behavioral and welfare impacts of notches are analyzed using the standard bunching approach whereby one non-parametrically estimates the mass of agents bunching around a notch and then translates this bunching mass into a structural parameter(s) (Kleven, 2016). One can then use these structural parameters to gauge the welfare impacts of previous policy changes or to predict the impacts of future policy reforms. However, the chief limitation of this approach is that translating a bunching mass into a structural parameter typically relies heavily on the modeling assumptions of the agent problem. For example, assumptions on preferences, the types of behavioral responses available to agents, and the degree of optimization frictions can all greatly impact the mapping between a bunching mass and a structural parameter.[1] Consequently, Kleven (2016) states "although bunching provides compelling non-parametric evidence of a behavioral response, moving from observed bunching to a structural parameter that can be used to predict the effects of policy changes is difficult."

This paper develops a technique that uses changes in bunching to assess the welfare impacts of changes to notches in transfer schedules whilst imposing minimal structure on the agent problem. In particular, we show how to bound the welfare impact of such a reform using two sufficient statistics: (1) the number of agents bunching at the original threshold who move toward the new threshold as a result of the reform, and (2) the number of agents who "jump" down to bunch at the new threshold as a result of the reform.[2] Importantly, our bounds require only minimal structure on the agent problem: agents can have arbitrary preference heterogeneity, agents can have any number of choice variables, and agents may face optimization frictions such as limited choice sets and/or adjustment costs. Thus, our approach overcomes a key limitation of the standard bunching approach by using reduced-form bunching evidence to inform welfare analysis without making strong assumptions about the agent's optimization problem. However, the benefits of our approach have corresponding costs relative to standard bunching methods: our approach only provides bounds on the welfare impacts of a particular reform and cannot be used to infer the welfare impact of proposed future reforms. Thus, the strengths and weaknesses of our approach are complementary to those of the standard bunching approach.

We then use our method to bound the welfare impacts a reform that expanded a notch in one of the world's largest cash transfer programs, Bolsa Família (BF). Our bounds suggest that the marginal value of public funds for this reform lies between 0.90 and 1.12, which implies that the reform was welfare improving if the government values giving R$0.90 to BF recipients (in a non-distortionary way) more than spending

1. This is, of course, not always the case: in some applications, structural parameters appear robust to model augmentations, e.g., Best et al. (2020).

2. We define jumping behavior as a discontinuous change in an agent's choice variable when a policy changes by an infinitesimal amount.

one dollar on their next best alternative and welfare decreasing if the government values giving R$1.12 to BF recipients (in a non-distortionary way) less than spending one dollar on their next best alternative. Because recipients of Bolsa Família are typically very poor (Bastagli (2008), Lindert et al. (2007)), we argue that it is very likely to be the case that the government is willing to spend R$1 to get R$0.90 to BF recipients. While the welfare impact of this reform is important in its own right, our application also highlights more generally how our approach can provide meaningful information for welfare analysis.

The paper begins by building intuition in a simple, static model in which a government provides a constant transfer only to households who report an income below a certain threshold; hence, the transfer schedule features a notch. Households are endowed with an income and choose how much income to report to the government subject to some cost of misreporting. We derive bounds on the marginal value of public funds (MVPF) of changing the notch, comprising of both an increase in the benefit level and eligibility threshold. The MVPF is the ratio of households' willingness-to-pay (WTP) for the reform relative to the government's budgetary cost of the reform (Hendren and Sprung-Keyser, 2020). We show that aggregate WTP for the reform (and, in turn, the MVPF) can be bounded using two empirical objects: (1) the number of "bunching households" who bunch at the old notch and move to (or towards) the new notch as a result of the reform, and (2) the number of "jumping households" who jump down to bunch at the new notch as a result of the reform. To calculate WTP for the reform, we first note that for households who do not jump or bunch, their WTP is simply equal to the amount of additional money they receive as a result of the reform. However, for the jumping and bunching households, their WTP for the reform differs from their increase in benefits. In particular, by jumping down to the new notch, jumping households now receive the benefit but also incur a misreporting cost. For infinitesimal reforms, their WTP for the reform is exactly zero because the only households that jump are those that are initially indifferent between misreporting at the notch to get the benefit and reporting truthfully. However, for larger reforms, jumping households have a positive WTP which is bounded above by the size of the new benefit given that misreporting costs are weakly positive. Thus, each jumping household has a WTP for the reform between 0 and the size of the new benefit. On the other hand, bunching households see an increase in the size of their benefit as well as a reduction in their misreporting costs as they move toward the new, higher notch.[3] Thus, bunching households' WTP for the reform is at least as large as their increase in benefits; moreover, their WTP cannot exceed the size of the new benefit level because if it did, it would not have been optimal for these households to bunch at the original notch in the first place. We can use these bounds on household WTP to bound aggregate WTP for the reform (and, in turn, the MVPF) as long as we can observe the number of bunching and jumping households. Because the welfare impact of a reform is equal to the MVPF multiplied by a normative welfare weight all minus one, our bounds on the MVPF enable us to make welfare statements about the impacts of changing a notch using estimable, reduced-form objects.

While our results are initially stated in the context of a simple, static misreporting model to build intuition, we show that our welfare bounds are highly robust to the agent problem. In particular, our bounds hold in models with any sort of behavioral response margin (e.g., labor supply responses instead of misreporting responses), arbitrary preference heterogeneity, adjustment costs, limited choice sets (e.g., labor supply frictions), and (non-extreme) misperceptions of the benefit schedule. Moreover, we discuss how our

3. Note, even for infinitessimal reforms, the utility gain from bunching households re-optimizing in response to a change in the location of the notch is first order; this is because utility is not continuous in the location of the notch, so one cannot apply the typical envelope theorem logic to argue that behavioral responses have only second order impacts on utility for infinitesimal reforms.

bounds can be augmented to allow for dynamic decision making with uncertainty and for more complex policy environments such as in-kind transfers and the presence of other tax and transfer schedules. We view this generality as a key contribution: the number of jumping and bunching households can be used to bound the MVPF of a notch change in a wide class of models. We therefore refer to the number of jumping and bunching households as the sufficient statistics to bound the MVPF, and in turn, the welfare impact of a notch change.

The second part of this paper is concerned with bounding the welfare impacts of a reform to a notch in one of the world's largest cash transfer programs, Bolsa Família (BF). BF is a Brazilian anti-poverty program that started in 2003 and gave benefits to around 14 million households as of 2014 (Gazola Hellmann, 2015). Importantly, the BF schedule features a pronounced notch that was expanded in June 2014. Prior to June 2014, households reporting a per-capita income below the extreme-poverty threshold of R$70 per-month were eligible for an unconditional monthly benefit of R$70 per-month, whereas households reporting an income above this threshold were not.[4] However, from June 2014 onward, both the eligibility threshold and benefit were raised by 10%: the eligibility threshold increased from R$70 per-capita, per-month to R$77 per-capita, per-month, while the unconditional benefit increased from R$70 per-month to R$77 per-month.

We have access to administrative data spanning December 2011 to September 2016 from Cadastro Único, which is the Brazilian government's national registry used to determine eligibility for all federal social welfare programs. Using this data, we seek to estimate the number of original bunching households who move towards the new notch and the number of households jumping to the new notch as a result of the reform. The number of bunching households is simply equal to the reduction in the bunching mass at the old notch (as a result of reform), while the number of jumping households is equal to the increase in the bunching mass at the new notch less the reduction in the bunching mass at the old notch (as a result of the reform). However, when looking at the raw data, there are clear time trends in the number of households locating at the old and new notch in the pre-reform period. Thus, our identification challenge is to understand how the number of households reporting incomes at the old and new notch would have evolved in the post-reform period had the reform not occurred.

At a high level, our identification strategy is based on the following insight: the incentives to report an income at and above the original notch are affected by the reform, but the incentives to report an income below the original notch are unaffected by the reform. Hence, we can use portions of the reported income distribution below the original notch as control groups for underlying time trends in the reported income distribution at and above the original notch. To do so, we use a generalized difference-in-difference strategy as employed in Wolfers (2006) and Mora and Reggio (2013). A standard difference-in-difference estimator requires that the control groups and treatment groups have "parallel trends" pre-reform that would have persisted in the post-reform period had the reform not occurred. Similarly, the generalized difference-in-difference estimator requires that there exist stable relationships, which are well approximated by relatively low order polynomials, between treatment and control groups which existed pre-reform that would have persisted in the post-reform period had the reform not occurred. In our setting, we discretize the reported income distribution into bins, using bins below the original notch as control bins while using bins that include the original and new notches as treatment bins. We then use the estimated relationships between treatment

4. BF also includes a conditional variable benefit for households with children. To receive this variable benefit, households must make specified health and education investments in their children and have a reported per-capita income below the "poverty threshold" (which is higher than the extreme-poverty threshold). We restrict our attention to the *unconditional* component of BF in this paper.

and control bins in the pre-reform period to create counterfactuals for how the treatment bins would have evolved absent the reform. Fundamentally, there are complex structural relationships between the number of people reporting in each income bin; these relationships are determined by how the true income distribution is evolving over time, how the true income distribution is mapped to the reported income distribution, as well as growth in applicants to the BF program over time. Our generalized difference-in-difference strategy essentially assumes that the differences over time between the number of people reporting in each income bin are well approximated by polynomials which would have persisted in absence of the reform; this assumption appears to hold in the pre-reform period, lending credence to our identification assumptions.

Using this generalized difference-in-difference strategy, we find that the number of households reporting incomes at the new notch increased by approximately 50,000 and the number of households reporting incomes at the old notch decreased by approximately 27,000 as a result of the reform. Hence, the reform induced 27,000 bunching households to move to the new threshold and an additional 23,000 households to jump down to the new threshold. This translates into a lower bound for the MVPF of the reform of approximately 0.9 and an upper bound for the MVPF of approximately 1.12. In terms of welfare implications, as long as the government values giving R$0.90 (in a non-distortionary manner) to the BF households more than spending R$1 on their next best alternative, the reform was welfare improving. We argue that this is likely the case given that BF has high coverage of households in extreme poverty and given that the households misreporting to get the benefit typically fall in the poorest half of the population (Bastagli (2008) and Lindert et al. (2007)). A back-of-the-envelope calculation suggests that the welfare effect of spending R$1 on the BF reform is at least as high as the welfare effect of spending R$1.50 on a non-distortionary universal transfer, suggesting that the BF reform was welfare improving.

We believe it is helpful to further compare and contrast our paper with the growing literature that explores bunching at notches and kinks (starting with Saez (2010) and Kleven and Waseem (2013)). As mentioned, our approach overcomes a key limitation of standard bunching analyses: we can use reduced-form bunching evidence to make welfare statements without making strong assumptions about the agent's optimization problem. However, this advantage comes with limitations: our bounds are only relevant for the reform in question and cannot be used to infer the welfare impacts of alternative reforms. Moreover, our approach is more data intensive: to estimate changes in bunching masses without structural assumptions, one requires cross-sectional data in both the pre- and post-reform period. Our paper also differs from standard bunching analyses with respect to our empirical strategy. We use a difference-in-difference strategy to estimate *changes* in bunching as a result of a reform.[5] In contrast, papers analyzing notches (or kinks) typically estimate bunching by assuming that portions of the distribution far from the notch are unimpacted by the presence of the notch and then make smoothness assumptions on the underlying distribution to non-parametrically identify the excess bunching induced by the notch. However, as discussed further in Section 4.3, our distribution of reported incomes is extremely non-smooth (we have extreme bunching at numbers equal to 0 mod 50, substantial bunching at numbers which were notches many years earlier, and less extreme bunching at numbers equal to 0 mod 10); while bunching estimation methods have been augmented to deal with "round-number" bunching and bunching at "reference-points" (e.g., Kleven and Waseem (2013) and Best and Kleven (2017)), we believe that the pervasiveness and the variability of round-number and reference-point

5. To the best of our knowledge, few papers attempt to estimate changes in bunching as a result of a change in a kink/notch while controlling for underlying time dynamics. One exception is Carril (2021) who uses standard bunching techniques to estimate the counterfactual distributions that would ensue in absence of a notch in both the pre- and post-reform periods and then uses changes in these counterfactual distributions to control for underlying dynamics.

bunching in our setting will make it too difficult to precisely identify the counterfactual distributions around our notches (in particular, around the original notch of R$70 which is equal to 0 mod 10). Hence, using a difference-in-difference strategy to estimate behavioral responses to changes in notches is a contribution of this paper and will hopefully be helpful in other settings where estimating bunching directly is difficult.

Finally, perhaps the closest work to this paper on the theoretical side is a contemporaneous paper by Lockwood (2020). Lockwood observes that in the presence of notches, the sufficient statistic approach to welfare analysis of tax systems needs to be augmented to include an additional term which captures the change in bunching at a notch in response to a tax reform. Similarly, our approach also requires the change in bunching at a notch to evaluate welfare impacts of a tax reform. Technically, we allow for changes in the location of the notch (which Lockwood (2020) does not) which means we need to observe the change in bunching at the new notch as well as the old notch. Our key theoretical contribution relative to Lockwood (2020) is the generality of our results. Whereas Lockwood (2020) characterizes the welfare impacts of infinitesimal reforms to notches in a model with quasi-linear utility, one dimension of agent heterogeneity, and one choice variable, our approach allows us to bound the welfare impacts of reforms to programs with notches of any size while placing very little structure on the agent problem.

The remainder of the paper is organized as follows. Section 2 introduces the theoretical framework for welfare analysis of changing notches, Section 3 discusses the structure of the Bolsa Família program, Section 4 discusses our strategy to empirically identify the number of jumping and bunching households for the June 2014 reform, Section 5 presents our results and robustness analysis, while Section 6 concludes.

2 Theoretical Framework for Welfare Analysis of Changing Notches

This section derives bounds for the welfare impact of changing a notch in a transfer program. To build intuition, we begin by deriving bounds in a simple, static misreporting model, showing that our bounds can be expressed in terms of two empirical objects: (1) the number of households bunching at the old notch who move towards the new notch as a result of the reform, and (2) the number of households who jump down to the new notch as a result of the reform. We then show that our bounds still hold in a much more general model with minimal structure on the household problem, thereby arguing that our two empirical objects are "sufficient statistics" to bound the change in welfare.

2.1 Baseline Model Set-Up

To begin, we consider a static world in which households choose how much income to report, $\hat{y}$, subject to a policy $\mathbf{p} = \{b, \tau\}$ where b denotes the level of the benefit and τ denotes the cut-off level of reported income s.t. those reporting $\hat{y} \leq \tau$ receive b and those reporting $\hat{y} > \tau$ receive nothing.[6] Households have two

6. For the sake of parsimony, our theoretical framework ignores other aspects of the tax and transfer system. Thus, we assume that behavioral responses to changes in the notch do not have budgetary impacts on other programs. For instance, we assume changing reported income does not impact income tax payments; this is innocuous in our setting as households receiving BF are almost certainly too poor to be liable for income taxation (e.g., in 2015, individuals earning less that R$1903.98 per-month are exempt from income taxation). We also assume that receipt of other welfare programs is not impacted by changes in reported income. This again is reasonable in our setting because while there are numerous welfare programs in addition to BF in Brazil, these programs typically have thresholds well above the BF thresholds (see Appendix B.2 for details on other programs tied to the Cadastro Único registry). Despite this, we discuss how our framework can account for other aspects of the tax and transfer

dimensions of heterogeneity: (1) endowed income y distributed according to CDF $F(y)$, and (2) aversion to misreporting governed by $\mu \in \{1, 2\}$ with probability mass function $\pi(\mu)$. Type $\mu = 1$ households are "truth-telling" households whereas type $\mu = 2$ households are willing to misreport their income. For simplicity, we assume there is no fixed cost of reporting an income and that all households know the policy $\mathbf{p}$. We also assume households have quasi-linear utility in consumption. Type $\mu = 1$ households therefore have utility under policy $\mathbf{p}$ given by:

$$U^*(y, \mu = 1; \mathbf{p}) \equiv y + b\mathbb{1}(y \leq \tau) \tag{1}$$

Utility for type $\mu = 2$ households under policy $\mathbf{p}$ is equal to:

$$\begin{aligned} U^*(y, \mu = 2; \mathbf{p}) \equiv \max_{\hat{y}} \; & c - v(y - \hat{y})\,\mathbb{1}(y > \hat{y}) \\ \text{s.t. } & c = y + b\mathbb{1}(\hat{y} \leq \tau) \end{aligned} \tag{2}$$

where $v(y - \hat{y})\,\mathbb{1}(y > \hat{y})$ captures the disutility of reporting $\hat{y}$ when true income is y. We assume $v' > 0$ and $v'' \geq 0$ so that the cost of misreporting is increasing and convex in the discrepancy between true and reported income.[7] If a household is indifferent between reporting truthfully and over-reporting, we break their indifference by assuming they report truthfully.

Optimal reported incomes, $\hat{y}^*$, for type $\mu = 2$ households are characterized as follows (a formal derivation of optimal household reported income is given in Appendix A.1):

$$\hat{y}^*(y, \mu = 2; \mathbf{p}) = \begin{cases} y & \text{if } y \leq \tau \\ \tau & \text{if } y \in (\tau, y^c(\mathbf{p})] \\ y & y > y^c(\mathbf{p}) \end{cases}$$

where $y^c(\mathbf{p})$ is the income level for which households are indifferent between misreporting at τ and reporting truthfully, implicitly defined by $y^c(\mathbf{p}) + b - v(y^c(\mathbf{p}) - \tau) = y^c(\mathbf{p})$. In words, type $\mu = 2$ individuals with $y < \tau$ report truthfully and get the benefit, those with $y \in (\tau, y^c(\mathbf{p})]$ get the benefit by misreporting and bunching at the notch, and those with $y > y^c(\mathbf{p})$ report their income truthfully and do not get the benefit.

Next, we define $G(x; \mathbf{p})$ to capture the share of the population reporting an income less than or equal to x under policy $\mathbf{p}$:

$$G(x; \mathbf{p}) = \sum_{\mu \in \{1,2\}} \int_Y \mathbb{1}(\hat{y}^*(y, \mu; \mathbf{p}) \leq x) dF(y|\mu)\pi(\mu)$$

Hence, $G(\tau; \mathbf{p})$ captures the total number of individuals locating under the eligibility threshold τ. Finally, we assume that total social welfare under policy $\mathbf{p}$, $\mathcal{W}(\mathbf{p})$, is given by a weighted sum of household utilities evaluated at optimal choices under policy $\mathbf{p}$ less the budgetary cost of the policy multiplied by the shadow value of public funds λ, which captures the welfare gain of spending a dollar on the government's next best alternative. In other words, we assume that spending \$1 on the program comes at the cost of spending \$1 less on some other program which decreases welfare by a constant value λ. Welfare under policy $\mathbf{p}$ is therefore given by:

$$\mathcal{W}(\mathbf{p}) = \sum_{\mu \in \{1,2\}} \int_Y \phi(y, \mu) U^*(y, \mu; \mathbf{p}) dF(y|\mu)\pi(\mu) - \lambda b G(\tau; \mathbf{p}) \tag{3}$$

schedule in Section 2.3 and how potential fiscal externalities impact our findings in Section 5.2

7. Modeling evasion using a convex and deterministic cost function $v(\cdot)$ dates back to Mayshar (1991) and Slemrod (2001).

where $\phi(y, \mu)$ denotes the government's welfare weight on a household with income y and type μ.

2.2 Welfare Effect of the Reform in Baseline Model

Our goal is to evaluate the welfare impact of a reform from policy $\mathbf{p} = \{b, \tau\}$ to $\mathbf{p}' = \{b', \tau'\}$, defining $\{\Delta b, \Delta\tau\} = \mathbf{p}' - \mathbf{p}$. For ease of exposition, let us assume $\Delta b > 0$ and $\Delta\tau > 0$ (i.e., we increase the level and location of the notch). Our goal is to derive bounds for $\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p})$ in terms of empirically observable objects (we discuss the advantages of focusing on bounding the welfare impact as opposed to characterizing the exact impact in Section 2.3). Hendren and Sprung-Keyser (2020) show that the welfare impacts of any policy reform can be expressed in terms of a *normative* welfare weight along with a *positive* sufficient statistic, denoted the MVPF, which captures households' willingness-to-pay (WTP) for the reform relative to the total budgetary cost of the reform. Our goal then is to derive bounds for the marginal value of public funds (MVPF).

Because the utility function is quasi-linear, a household with income y and misreporting type μ has a WTP for the reform implicitly defined by:

$$U^*(y, \mu; \mathbf{p}) = U^*(y, \mu; \mathbf{p}') - WTP$$

Equivalently, WTP is simply equal to the compensating variation. We now heuristically derive bounds on the WTP for all households impacted by the reform (we formally derive these bounds in the proof to Lemma 1). To do so, we will split the set of households impacted by the reform into four groups.

The first group of households we consider is the *mechanical households*. These are the households who report at or below τ under policy $\mathbf{p}$ and who do not change their behavior as a result of the reform. The number of mechanical households is given by $M = G(\tau; \mathbf{p}')$. Notably, M is not equal to the mass reporting at or below τ under policy $\mathbf{p}$ as this mass includes bunching households who do update their behavior as a result of the reform (see discussion below). Conversely, those reporting at or below τ under $\mathbf{p}'$ will also report at or below τ under $\mathbf{p}$. Mechanical households receive an increase in benefits equal to Δb. Hence, their WTP for the reform is exactly equal to Δb.

The second group of households is the *bunching households*. These are the households who misreport and bunch at the original threshold τ under policy $\mathbf{p}$ and move with the threshold as it is increased (i.e., who report between $(\tau, \tau']$ under policy $\mathbf{p}'$).[8] Thus, the number of bunching households is equal to reduction in households locating at or below τ *as a result of the reform*: $B = G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')$. The bunching households receive an increase in benefits equal to Δb. Moreover, they experience a reduction in their misreporting costs as the move from τ to τ'. Hence, the bunching households have a WTP of at least Δb.[9] Moreover, by revealed preference arguments, the most the bunching households can value this reduction in misreporting costs is b dollars. If they value this reduction by more than b dollars, it would not have been optimal for these households to bunch at τ under policy $\mathbf{p}$.[10] Thus, the bunching households' WTP for the reform is in

8. Not all bunching households will move all the way to τ'. In particular, bunchers with $y \in (\tau, \tau')$ will report $\hat{y} = y$ under $\mathbf{p}'$.

9. Note that even for an infinitesimal reform, the utility gain bunching households experience from moving with the notch has a first-order impact on welfare. This is because the envelope theorem does not apply for these households as utility changes discontinuously as we perturb $\mathbf{p}' \to \mathbf{p}$ (holding decisions fixed at optimal decisions under policy $\mathbf{p}'$).

10. Suppose bunching households have a WTP for the relaxation in their misreporting costs greater than b: $v(y-\tau) - v(y-\tau') > b$. This implies $v(y - \tau) > b$. However if $v(y - \tau) > b$, bunching households would have preferred to report truthfully above τ over misreporting at τ under policy $\mathbf{p}$.

$[\Delta b, \Delta b + b] = [\Delta b, b']$.

The third group of households is the *threshold households*. These are the households who report in $(\tau, \tau']$ under policy $\mathbf{p}$. Thus, the number of threshold households is given by $T = G(\tau'; \mathbf{p}) - G(\tau; \mathbf{p})$. Notably, these households will not update their behavior in response to the reform. This is because all households reporting above τ given policy $\mathbf{p}$ are reporting truthfully regardless of their type μ. Under $\mathbf{p}'$ the optimal choice for these households is to remain reporting truthfully as doing so allows them to receive b' and not incur any misreporting costs. These households go from receiving no benefits to receiving b' in benefits. Hence their WTP is equal to b'.

Finally, the fourth group of households is the *jumping households*. These are the households who report above τ' given policy $\mathbf{p}$ but who report at τ' given policy $\mathbf{p}'$. In particular, households who were previously close-to-indifferent between misreporting at the threshold and truthfully reporting above the threshold but opted for the latter (i.e., type $\mu = 2$ households with $y \in (y^c(\mathbf{p}), y^c(\mathbf{p}')]$) will now jump and misreport to the new threshold τ'.[11] Notably, we describe the behavioral response of these households as a "jump" because these households experience a discontinuous change in their optimal reported income as we move from $\mathbf{p}$ to $\mathbf{p}'$. The number of jumping households, J, is given by $J = G(\tau'; \mathbf{p}') - G(\tau'; \mathbf{p})$, which is equivalent to the increase in households reporting at or below the new threshold *as a result of the reform* (note, the number of jumping households is not equal to the mass of households locating at τ' given policy $\mathbf{p}'$ as some of the households locating at τ' may be original bunching households who moved with the notch). By revealed preference, jumping households' utility is improved by changing their behavior; hence, their WTP is weakly positive. Moreover, after the reform, jumping households get b' more dollars but incur misreporting costs. Therefore jumping households' WTP cannot exceed b' as misreporting costs are weakly positive. Hence the WTP of jumping households is in $[0, b']$.[12]

Table 1 summarizes the bounds on the WTP for each household in each of our four groups along with the number of households falling into each group and the cost to the government of each household in each group. Moreover, Figure 1 illustrates the change in the reported income density as a result of the reform as well as the set of households making up each group.

Table 1: WTP and Cost to the Government of the Reform

Group	Number of households	WTP (per-household)	Cost to Govt. (per-household)
Mechanical Households	$M = G(\tau; \mathbf{p}')$	Δb	Δb
Bunching Households	$B = G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')$	$\in [\Delta b, b']$	Δb
Threshold Households	$T = G(\tau'; \mathbf{p}) - G(\tau; \mathbf{p})$	b'	b'
Jumping Households	$J = G(\tau'; \mathbf{p}') - G(\tau'; \mathbf{p})$	$\in [0, b']$	b'

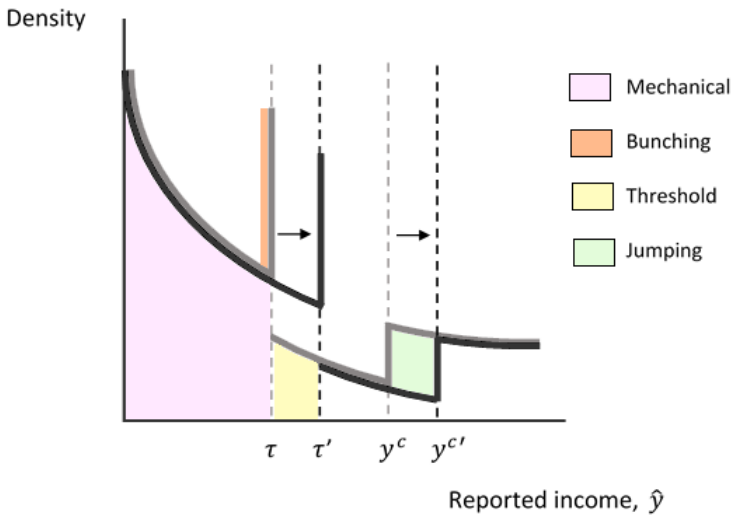
Note: This table shows the willingness-to-pay (WTP) and cost to the government for all the households impacted by the reform of moving from policy $\mathbf{p}$ to $\mathbf{p}'$. We split the households into four groups: mechanical, bunching, threshold, and jumping households.

This brings us to Lemma 1 which bounds the total WTP for the reform:

Lemma 1. *If individuals solve Problem (2), the total WTP of the reform from* $\mathbf{p}$ *to* $\mathbf{p}'$ *with* $\mathbf{p}' - \mathbf{p} =$

11. For large changes in τ s.t. $y^c(\mathbf{p}) < \tau'$, only $\mu = 2$ households with $y \in (\tau', y^c(\mathbf{p}')]$ jump and misreport to the new notch.

12. For an infinitesimal reform, jumping households have a WTP of 0 as the only households who jump are those that are indifferent between locating at the notch and reporting truthfully. Thus, our lower bound for the WTP for jumping households is exact for an infinitesimal reform. Note, however, that the jumping households still have a first order impact on social welfare through their effect on the government's budget as each jumping household costs the government b' dollars (despite the fact that the mass of jumpers is measure 0 for an infinitesimal reform; see Bergstrom and Dodds (2021) for further discussion).



Note: This figure shows a hypothetical density of reported incomes under the initial policy $\mathbf{p} = \{b, \tau\}$ (in grey) and how this density changes as a result of a reform that increases the policy to $\mathbf{p}' = \{b', \tau'\}$ (in black). Note, the vertical grey line at τ and the vertical black line at τ' represent the bunching households under the initial policy and new policy, respectively. This figure also depicts which households are classified as mechanical households, bunching households, threshold households, and jumping households.

FIGURE 1: DENSITY OF REPORTED INCOMES UNDER $\mathbf{p}$ AND $\mathbf{p}'$

$\{\Delta b, \Delta\tau\} > 0$ *can be bounded as follows:*

$$\Delta b(M+B) + b'T \leq \textit{Total WTP} \leq \Delta bM + b'(B+T+J)$$

where M, B, T *and* J *denote the mass of mechanical, bunching, threshold and jumping households mathematically defined in Table 1.*

Proof. See Appendix A.2. □

Moreover, using the cost per-household to the government given in Table 1, we can express the total budgetary cost of the reform as follows:

$$\text{Total Cost} \equiv b'G(\tau'; \mathbf{p}') - bG(\tau; \mathbf{p}) = \Delta b(M+B) + b'(T+J) \tag{4}$$

Hence, we can construct bounds for the MVPF of the reform as follows:

Proposition 1. *If individuals solve Problem (2), the marginal value of public funds of the reform from* $\mathbf{p}$ *to* $\mathbf{p}'$ *with* $\mathbf{p}' - \mathbf{p} = \{\Delta b, \Delta\tau\} > 0$ *can be bounded as follows:*

$$MVPF \geq MVPF_L \equiv \frac{\Delta b(M+B) + b'T}{\Delta b(M+B) + b'(T+J)} = 1 - b'\frac{J}{\textit{Total Cost}}$$

$$MVPF \leq MVPF_U \equiv \frac{\Delta bM + b'(B+T+J)}{\Delta b(M+B) + b'(T+J)} = 1 + b\frac{B}{\textit{Total Cost}}$$

Proof. This follows directly from Lemma 1 and Equation (4). □

We can use these bounds to construct bounds on the money metric welfare gain relative to the budgetary cost of the reform using Proposition 2:[13]

Proposition 2. *If individuals solve Problem (2) and social welfare is given by Equation (3), then the money metric welfare gain relative to the budgetary cost of the reform from* $\mathbf{p}$ *to* $\mathbf{p}'$ *with* $\mathbf{p}' - \mathbf{p} = \{\Delta b, \Delta\tau\} > 0$ *can be bounded as follows:*

$$\omega_L MVPF_L - 1 \leq \frac{\frac{1}{\lambda}\left[\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p})\right]}{b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p})} \leq \omega_U MVPF_U - 1$$

where ω_L *(*ω_U*) captures the weighted average money-metric welfare gain from giving a dollar to mechanical, bunching, threshold, and jumping households, where the weights are determined by the relative size of each group's lower bound (upper bound) for WTP.*

Proof. We do not provide a separate proof for Proposition 2; we only provide a proof for Proposition 3, which nests Proposition 2 (see Appendix A.3). □

Proposition 2 bounds the increase in total welfare from spending \$1 on the reform. In particular, $\omega_L MVPF_L$ captures a lower bound for the increase in total welfare, measured in dollars, of spending \$1 on the reform ignoring the budgetary cost of this \$1: $MVPF_L$ captures our lower bound on the total WTP of the mechanical, bunching, threshold, and jumping households when we spend \$1 on the reform, while ω_L denotes the welfare gain, measured in dollars, of splitting \$1 among the mechanical, bunching, threshold, and jumping households (where the split is determined by the lower bounds on each group's WTP for the reform). Thus, subtracting the budgetary cost of \$1 from $\omega_L MVPF_L$ gives a lower bound for the total welfare gain of spending \$1 on the reform. Symmetric logic explains why $\omega_U MVPF_U - 1$ is an upper bound for the increase in total welfare, measured in dollars, of spending \$1 on the reform.

Finally, in light of Proposition 1, we can express $MVPF_L$ and $MVPF_U$ in terms of two positive objects: (1) the number of bunching households, and (2) the number of jumping households.[14] Thus, the number of bunchers and jumpers are the positive objects to construct bounds for the welfare impact of the reform.

2.3 Robustness to Model Specification

To highlight the robustness of our bounds for the welfare impact, we now show that we actually require very little structure on preferences or behavioral responses of agents. Suppose households have several decisions variables denoted by the vector x within a choice set X. Household decisions are made conditional on primitives denoted by the vector $\theta \in \Theta$ and the policy $\mathbf{p}$. Households get the benefit b if their reported income $\hat{y}$, which can be a decision variable or a function of decision variables and primitives, is below τ. Household income, denoted y, is also potentially a function of decisions x.[15] Households solve:

$$\begin{aligned} U^*(\theta;\mathbf{p}) = &\max_{x\in X} \; u\left(c, x; \theta\right) \\ &\text{s.t. } c = y(x,\theta) + b\mathbb{1}(\hat{y}(x,\theta) \leq \tau) \end{aligned} \tag{5}$$

13. Note, that the welfare change is expressed in dollar units (as opposed to welfare units) as we divide through by the shadow value of public funds, λ.

14. Technically, we also need to know the total cost of the reform; however, Equation (4) shows that the total cost is a function solely of $G(\tau';\mathbf{p}')$ and $G(\tau;\mathbf{p})$, which are necessary objects to construct J and B.

15. For example, in a labor supply model, x could equal household labor supply l, θ could include productivity n, and $\hat{y} = y = nl$.

where c denotes consumption. We assume total welfare is given by a weighted sum of utilities, with welfare weights given by $\phi(\theta)$:

$$\mathcal{W}(\mathbf{p}) = \int_{\Theta} \phi(\theta) U^*(\theta; \mathbf{p}) dF(\theta) - \lambda b G(\tau; \mathbf{p}) \tag{6}$$

where λ represents the shadow value of public funds and $G(\tau; \mathbf{p}) = \int_{\theta:\hat{y}(\theta,\mathbf{p})\leq\tau} dF(\theta)$ represents the number of households receiving the benefit under policy $\mathbf{p}$. More generally, we define $G(z; \mathbf{p}) = \int_{\theta:\hat{y}(\theta,\mathbf{p})\leq z} dF(\theta)$. This setup allows us to substantially generalize Proposition 1 and Proposition 2:

Proposition 3. *Suppose households solve Problem (5), welfare is given by Equation (6), and $\tau' > \tau$. Defining:*

$$MVPF_L \equiv 1 - \frac{b' \left[G(\tau'; \mathbf{p}') - G(\tau'; \mathbf{p})\right]}{b' G(\tau'; \mathbf{p}') - b G(\tau; \mathbf{p})} = 1 - b' \frac{J}{\textit{Total Cost}} \tag{7}$$

$$MVPF_U \equiv 1 + \frac{b[G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')]}{b' G(\tau'; \mathbf{p}') - b G(\tau; \mathbf{p})} = 1 + b \frac{B}{\textit{Total Cost}} \tag{8}$$

Then as long as $b' G(\tau'; \mathbf{p}') - b G(\tau; \mathbf{p}) > 0$ have:

$$\omega_L MVPF_L - 1 \leq \frac{\frac{1}{\lambda}\left[\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p})\right]}{b' G(\tau'; \mathbf{p}') - b G(\tau; \mathbf{p})} \leq \omega_U MVPF_U - 1 \tag{9}$$

where ω_L (ω_U) captures the weighted average money-metric welfare gain from giving a dollar to mechanical, bunching, threshold, and jumping households, where the weights are determined by the relative size of each group's lower bound (upper bound) for WTP.

Proof. See Appendix A.3. □

Proposition 3 highlights that we can bound the MVPF and, consequently, bound the welfare impacts of the reform using empirically observable objects while only putting limited structure on the household problem.[16] [17] In particular, to bound the MVPF we simply need to estimate how the number of people locating below the new notch changes as a result of the reform, $J = G(\tau'; \mathbf{p}') - G(\tau'; \mathbf{p})$, as well as how the number of people locating below the old notch changes as a result of the reform, $B = G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')$.[18] In the context of our simple baseline model, B equals the number of households bunching at the original notch and J equals the number of households jumping down to bunch at the new notch. However, if households solve the more general household problem (5), the interpretation of these terms may be changed. For instance, consider a labor supply model in which households can only work full-time, half-time, or not at all. In this case, there will be no "bunching households" so it is a misnomer to say $G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')$ captures the number of bunching households as in Proposition 1. Nonetheless, for ease of exposition, we will continue to refer to

16. We are slightly abusing notation here. Our upper bound for the MVPF, $MVPF_U$, is actually the upper bound on the MVPF from moving from policy $\mathbf{p}'$ to $\mathbf{p}$. This is because our upper bound on the WTP is actually an upper bound on the negative WTP of moving from policy $\mathbf{p}'$ to $\mathbf{p}$, or equivalently, it is an upper bound on the willingness-to-accept (WTA) to move from policy $\mathbf{p}$ to $\mathbf{p}'$. We obviate this distinction in Proposition 1 by making the assumption that utility is quasi-linear in consumption, so that WTP=WTA.

17. Note that in Proposition 3 we have assumed $\tau' > \tau$ and that $b' G(\tau'; \mathbf{p}') - b G(\tau; \mathbf{p}) > 0$. Assuming $\tau' > \tau$ is WLOG because one can also use Equation (9) to get bounds on the welfare gain from moving from $\mathbf{p}'$ to $\mathbf{p}$. On the other hand, if $b' G(\tau'; \mathbf{p}') - b G(\tau; \mathbf{p}) < 0$, both inequalities in Equation (9) are simply reversed.

18. Again, we also need to know the total cost of the reform; but the total cost can be constructed from $G(\tau'; \mathbf{p}')$ and $G(\tau; \mathbf{p})$, which are necessary objects to construct J and B.

$B = G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')$ as the number of bunching households and $J = G(\tau'; \mathbf{p}') - G(\tau'; \mathbf{p})$ as the number of jumping households.

The core intuition for Proposition 3 is that we can again use revealed preference arguments to bound WTP for a reform from policy $\mathbf{p}$ to policy $\mathbf{p}'$. And these revealed preference arguments require very little structure on household utility, household heterogeneity, the choice variables, or choice sets available to households. Problem (5) can encompass a variety of important realisms: (1) households may respond by changing labor supply instead of misreporting their income (or respond on a variety of dimensions),[19] (2) households may face limited choice sets (e.g., restrictions on labor supply), (3) households may face reporting costs (e.g., hassle or time costs) and thereby face a decision of whether to report/update their income on the registry[20], (4) households may have a wide range of heterogeneity in their utility functions. We view this robustness as perhaps the most important aspect of our theory: we can construct bounds for the MVPF of changing a notch in a manner which is, in large part, model-free.

But of course there are *some* implicit restrictions encoded in the assumed household problem (Problem (5)) used to prove Proposition 3. Perhaps most importantly, Proposition 3 requires that households correctly perceive the benefit schedule and the reform. The proof to Proposition 3 uses the fact that household re-optimization *improves* utility; if misperceptions are extreme for many households this may not be the case. However, it is straight-forward to extend Proposition 3 when households misperceive the schedule if we are willing to assume that, on average, behavioral responses to the reform from $\mathbf{p}$ to $\mathbf{p}'$ improve welfare (i.e., perceptions are not so extreme that households, on average, harm themselves by responding to the reform; see Appendix A.5). Problem (5) is also a static problem, so that Proposition 3 does not allow for dynamic decision making or uncertainty. We augment Proposition 3 in Appendix A.6 to show that we can bound the discounted welfare impact of the policy over time relative to the discounted total budgetary cost in a more general dynamic model with income dynamics, savings, and stochastic shocks. In this case, the relevant bounds for the MVPF are constructed using the discounted sum of the expected number of jumping households over time and the discounted sum of the expected number of bunching households over time. Proposition 3 also requires that there are no externalities from household decisions so that decisions of one household do not directly impact the utility of any other household. Our bounds can be generalized to allow for externalities by augmenting the upper and lower bounds for the MVPF with an additional term measuring the WTP for these externalities relative to the total cost; however, measuring WTP for externalities is likely difficult in practice.

Additionally, Proposition 3 can be augmented to allow for more complex policy environments. Proposition 3 assumes that there is no underlying tax and transfer system beyond the benefit b given to those with a reported income $\hat{y} \leq \tau$. But we can easily extend Proposition 3 to account for more complex underlying tax and transfer schedules; in this case, the total budgetary cost of the reform must include the impact that behavioral responses have on other programs that impact the government's budget (i.e., we need to calculate the fiscal externalities associated with the reform).[21] Proposition 3 also assumes that the transfer is given in cash; if the transfer is paid in-kind, one would require an estimate of the average WTP for \$1 worth of the

19. This is consistent with Feldstein (1999), who argues that the efficiency costs of taxation do not depend on whether behavioral responses occur on the labor supply margin or the misreporting margin.

20. To capture adjustment/updating costs, one could suppose $x = \hat{y}_t$ and θ consists of current income y_t, aversion to misreporting μ, and prior reported income $\hat{y}_{t-1}$; households incur an adjustment cost k if their reported income today differs from prior reported income.

21. In particular, the total cost of the reform would equal $b'G(\tau'; \mathbf{p}') + R(\mathbf{p}') - bG(\tau; \mathbf{p}) - R(\mathbf{p})$ where $R(\mathbf{p})$ denotes net government spending under policy $\mathbf{p}$ (exclusive of spending on BF), and $\Delta R = R(\mathbf{p}') - R(\mathbf{p})$ captures the fiscal externalities of

in-kind good to apply an analogue to Proposition 3.

Lastly, we discuss the advantages of bounding the welfare impacts (as opposed to exactly characterizing the welfare impacts) of a notch reform. First, bounding welfare impacts requires very little structure on the household optimization problem whereas exactly characterizing welfare impacts requires one to take a stance on which margins households are responding, the functional form of their utility function, the sorts of frictions they face, etc. In fact, we prove in Appendix A.4 that the bounds in Proposition 3 are as tight as possible without making additional assumptions on the household problem. Second, our bounds on the welfare impact of a notch change are expressed in terms of estimable reduced-form objects, whereas exactly characterizing welfare impacts would require one to estimate a potentially large number of structural parameters.

We now turn to our empirical application: estimating the number of bunching and jumping households for the June 2014 Bolsa Família reform, and, in turn, calculating bounds for the welfare impact of this reform. We begin by describing the structure of the Bolsa Família program, our data, and the June 2014 reform.

3 The Bolsa Família Program

The Brazilian anti-poverty program Bolsa Família (BF) provides cash transfers to poor households based on their reported, monthly per-capita income. BF was implemented in October 2003 and is administered by the social development ministry (Ministério do Desenvolvimento Social, or MDS). BF is one of the world's largest cash transfer programs with around 14 million households receiving benefits in 2014 (Gazola Hellmann, 2015). Applicants report their information, including information on household income, expenditures, assets, socioeconomic characteristics, and demographic characteristics, to interviewers at Cadastro Único agencies, which are program offices spread across Brazil's 5,570 municipalities. Information is entered by interviewers into the Cadastro Único registry. Beneficiaries of the BF program are required to update their information once every 2 years to maintain their benefits.

Eligibility for the BF program is based on reported household per-capita income. Household per-capita income is calculated in five steps. First, the applicant is asked to report labor income for the last month as well as average monthly labor income over the past year for each member in the household (the applicant must present a government-issued ID for herself and for each family member thus making it difficult to register fictitious family members). Second, the applicant is asked to report average monthly income received from five additional sources for each household member (see Appendix B.1, which shows the questionnaire used to calculate monthly income). Third, for each individual, the computer calculates the minimum between the average monthly labor income over the past year and last month's labor income. Fourth, the computer sums this minimum monthly labor income along with income from the five additional sources to get a measure of total monthly income for each individual. Finally, the computer sums this individual total monthly income across all household members and divides it by the number of household members.

The government then transfers a monthly, unconditional benefit (referred to as the "basic benefit") to all

the reform. In this case, the lower and upper bounds for the MVPF are given by:

$$MVPF_L \equiv 1 - \frac{b'\left[G(\tau';\mathbf{p}') - G(\tau';\mathbf{p})\right] - \Delta R}{b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p}) + \Delta R} = 1 - b'\frac{J}{\text{Total Cost}} - \frac{\Delta R}{\text{Total Cost}}$$

$$MVPF_U \equiv 1 + \frac{b[G(\tau;\mathbf{p}) - G(\tau;\mathbf{p}')] + \Delta R}{b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p}) - \Delta R} = 1 + b\frac{B}{\text{Total Cost}} - \frac{\Delta R}{\text{Total Cost}}$$

Thus, in addition to B and J, we need to observe ΔR to assess the welfare impact of the reform.

households reporting a per-capita income below the "extreme-poverty threshold" and transfers an additional benefit (referred to as the "variable benefit") to households with children who have a reported per-capita income below the higher "poverty threshold" conditional on these households making health and education investments in their children.[22] [23] Between July 2009 and June 2014, the extreme-poverty threshold was R$70 per-capita, per-month and the poverty threshold was R$140 per-capita, per-month.

Finally, the MDS has several enforcement mechanisms to prevent income misreporting. First, during the interview, the income questions come at the end of the questionnaire so that questions on expenditures and assets can help the interviewer asses the veracity of the reported income (Bastagli, 2008). Second, during the interview, the applicant is reminded of her responsibility to provide true statements under penalty of losing the right to be eligible for government programs (Gazola Hellmann, 2015). Third, the ministry conducts audits, which can be triggered by citizens' complaints and cross-checks of registry data with other datasets such as administrative data on formal employment, deaths, or automobile purchases (Gazola Hellmann, 2015). However, despite these attempts, the large informal sector in the Brazilian economy leaves substantial scope for misreporting, which we conjecture is an important response margin to the schedule reform.

3.1 Data Sources and Sample Description

We have access to the Cadastro Único household registry, which is used to determine the eligibility of households for BF as well as all other targeted federal social programs (Veras Soares, 2011). Many of these other programs have eligibility criteria above the BF thresholds which explains the large number of ineligible applicants in the registry.[24] These other programs do not change concurrently with the BF reform that we analyze and are discussed in more detail in Appendix B.2.

The final data set is constructed by appending eight extractions of the registry: one in December of each year from 2011 until 2015, one in April 2015, one in August 2015, and one in September 2016.[25] Each extraction contains the latest information for all households on the registry at the time of the extraction date. For instance, if a household updated its information in August of 2011 and September 2013, its information will appear as of August of 2011 in the 2011 and 2012 extractions and as of September 2013 in the 2013, 2014, 2015, and 2016 extractions. The reform to BF that we study occurred in June 2014. Summary statistics on household per-capita income and number of family members as of June 2014 are displayed in Table 2. Note, Table 2 also presents separate statistics for single individual households (households with one adult and no children) as our main analysis will focus solely on these households (we discuss why below).

22. The size of this conditional transfer varies based on the number of children you have and their age(s); see Appendix B.4.

23. Technically, eligible families do not automatically become BF beneficiaries. There is a quota (cap) on the number of beneficiaries per municipality. Prior to 2009, these quotas were based on the predicted number of households below the poverty threshold in each municipality. Post 2009, these quotas were based on the predicted number of households below the poverty threshold scaled by 1.18 (Gerard, Naritomi and Silva, 2021). If the number of eligible families exceeds the quota in a municipality, priority is assigned first to quilombola families, second to families with lower per-capita income, and third to families with a larger number of children (Gazola Hellmann, 2015). Given our focus will be on households reporting below the extreme-poverty threshold post 2009, we believe that it is reasonable to assume that the vast majority of these households do in fact receive the basic benefit.

24. Moreover, the government aims to register all families with per-capita incomes below half the minimum wage (or total incomes below three times the minimum wage) (Veras Soares, 2011). The minimum wage was R$724 per-month in 2014; thus half the minimum wage was R$362, which is substantially higher than the highest BF threshold.

25. Every time the MDS analyzes the Cadastro Único data, it creates one of these extractions. Therefore, the frequency of the extractions are a result of previous data analyses by the ministry. Appendix B.3 contains a figure depicting the timeline of the data extractions.

Table 2: Summary Statistics as of June, 2014

Variables	Mean	Median
Per Capita Income (single individual households)	330.51	200.00
	(324.30)	
Per Capita Income (all households)	151.17	76.66
	(184.97)	
Number of Members Per Household	3.04	3.00
	(1.46)	
Observations (single individual households)	4,232,528	
Observations (all households)	28,932,001	

Note: This table shows summary statistics for households in the Cadastro Único database as of June, 2014. Per-capita income denotes household, monthly, per-capita income and is measured in Brazilian reais. The PPP conversion from US dollars to Brazilian reais was 1.813 in 2014 (OECD, `https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm`).

3.2 The Transfer Schedule and the June 2014 Reform

At the beginning of our data (December 2011), the extreme-poverty threshold was equal to R$70 per-capita, per-month and the poverty threshold was equal to R$140 per-capita, per-month.[26] The basic benefit was equal to R$70 per-month, while the variable benefits were based on the number and ages of the children in the household (see Appendix B.4 for more information on the variable benefits). In June 2014, the government increased both the benefits and thresholds by 10%. Thus, the extreme-poverty threshold was raised from R$70 to R$77 per-capita, per-month, the poverty threshold was raised from R$140 to R$154 per-capita, per-month, the basic benefit was raised from R$70 to R$77 per-month, and the variable benefits were also increased by 10%. This reform was announced on national television by the president in April 2014.

As mentioned, our main analysis will focus on single individual households. These households are not eligible for the variable benefits as they do not have children. Thus, prior to June 2014, these households received R$70 per-month if their reported income was less than or equal to R$70 per-month and 0 otherwise. After June 2014, these households received R$77 per-month if their reported income was less than or equal to R$77 per-month and 0 otherwise. Thus, the benefit schedule for these households has a single notch which increased both in level and location as a result of the June 2014 reform; see Figure 2.

26. In 2003, the extreme-poverty threshold and the poverty threshold were set to equal one-fourth and one-half of the monthly minimum wage of R$200, respectively. These thresholds have since been periodically adjusted for inflation; the time betwen adjustments is *ad hoc* and not linked to the minimum wage. Prior to June 2014, the last readjustment was in July 2009 (Gazola Hellmann, 2015).

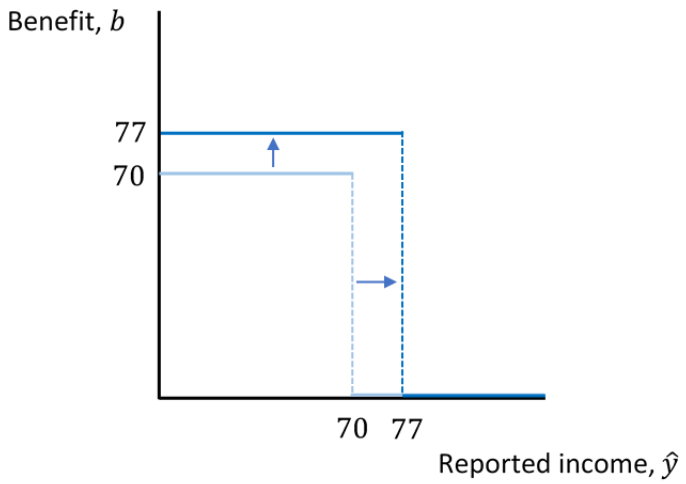


FIGURE 2: JUNE 2014 REFORM TO THE BENEFIT SCHEDULE FOR SINGLE INDIVIDUAL HOUSEHOLDS

We focus on single individual households because in February 2013, the government instituted a guaranteed minimum income of R$70 per-capita for all households which was subsequently raised to R$77 in the June 2014 reform.[27] However, because the basic benefit is equal to the guaranteed minimum income, the benefit schedule for single individual households is not impacted by this minimum. In contrast, for all other households, this minimum creates a kink in the benefit schedule below the extreme poverty threshold (the location of this kink will vary based on household composition); moreover, the location of this kink changes with the 2014 reform. For example, prior to the reform, households with two adults and no children had a kink at the reported per-capita income level of R$35 which increased to R$38.5 post June 2014.[28] This is problematic because, as will be discussed in Section 4, one of our identification assumptions is that the reported income distribution below the extreme-poverty threshold is unaffected by the June 2014 reform; clearly this is not necessarily true for households with more than one member because the kink created by the guaranteed minimum income changes concurrently with the notch. Nonetheless, we will discuss the impacts of the June 2014 reform on households with more than one individual in Section 5.5.

Finally, in June 2016, there was another reform to the BF program where both the benefit and the threshold were further increased. This reform, like the June 2014 reform, affected households of all compositions. Thus, in our empirical analysis, we will restrict our attention to the 4 year window around the June 2014 reform (i.e., June 2012 - June 2016).[29]

4 EMPIRICAL STRATEGY TO BOUND THE MVPF OF THE JUNE 2014 REFORM

Next, we apply our theoretical framework from Section 2 to empirically bound the MVPF of the June 2014 Bolsa Família reform for single individual households. Using the notation from Section 2, the June 2014

27. This guaranteed minimum income was instituted earlier for households with children. In particular, this guarantee was instituted in June 2012 for households with children below the age of 6, in November 2012 for households with children below the age of 15, and in February 2013 for all remaining households.

28. In particular, prior to June 2014, a two member household with a reported per-capita income less than R$35, $\hat{y} < 35$, will receive an additional monthly benefit equal to $2(70 - \hat{y}) - 70$. E.g., a two member household reporting a per-capita income of R$20 will receive R$70 in the basic benefit and an additional benefit of R$30.

29. We do not have enough data beyond June 2016 to analyze the effects of this later reform because our final data extraction is in September 2016.

reform changed the policy from $\mathbf{p} = \{b, \tau\} = \{70, 70\}$ to $\mathbf{p}' = \{b', \tau'\} = \{77, 77\}$. Thus, we seek to calculate:

$$MVPF_{L,\bar{t}} \equiv 1 - \frac{77 \times [G_{\bar{t}}(77; \mathbf{p}') - G_{\bar{t}}(77; \mathbf{p})]}{77 \times G_{\bar{t}}(77; \mathbf{p}') - 70 \times G_{\bar{t}}(70; \mathbf{p})} = 1 - 77 \times \frac{J_{\bar{t}}}{\text{Total Cost}_{\bar{t}}} \tag{10}$$

$$MVPF_{U,\bar{t}} \equiv 1 + \frac{70 \times [G_{\bar{t}}(70; \mathbf{p}) - G_{\bar{t}}(70; \mathbf{p}')]}{77 \times G_{\bar{t}}(77; \mathbf{p}') - 70 \times G_{\bar{t}}(70; \mathbf{p})} = 1 + 70 \times \frac{B_{\bar{t}}}{\text{Total Cost}_{\bar{t}}} \tag{11}$$

where $\bar{t}$ denotes the end of our four-year-window centered around the reform, i.e., $\bar{t}$ =June 2016. For simplicity, we focus on bounding the MVPF at the end of our four year window, which can be used to bound the welfare impact of the reform in June 2016.[30] Thus, we need to estimate the number of jumping and bunching households at the end of our four year window, which, in turn, requires us to estimate the number of households locating below the old and new notch under both $\mathbf{p}$ and $\mathbf{p}'$ at $\bar{t}$: $G_{\bar{t}}(70; \mathbf{p})$, $G_{\bar{t}}(70; \mathbf{p}')$, $G_{\bar{t}}(77; \mathbf{p})$, and $G_{\bar{t}}(77; \mathbf{p}')$. However, because this was a national reform, only $G_{\bar{t}}(77; \mathbf{p}')$ and $G_{\bar{t}}(70; \mathbf{p}')$ are observed directly because only policy $\mathbf{p}'$ was offered in period $\bar{t}$. Thus, our goal is to estimate the number of households reporting less than or equal to R\$70 and R\$77 in period $\bar{t}$ had the reform not happened: $G_{\bar{t}}(77; \mathbf{p})$ and $G_{\bar{t}}(70; \mathbf{p})$.

Ideally, we would estimate $G_{\bar{t}}(77; \mathbf{p})$ and $G_{\bar{t}}(70; \mathbf{p})$ using control groups from random experimental variation (e.g., from staggering the implementation of the reform randomly across geographies); however, as mentioned above, the BF reform was a national reform implemented everywhere starting June 2014. Consequently, our identification strategy will rely on using regions of the reported income distribution that *were not* impacted by the reform to control for underlying time trends in the portions of the reported income distribution that *were* impacted by the reform. This brings us to our two identification assumptions.

4.1 Identification Assumption 1

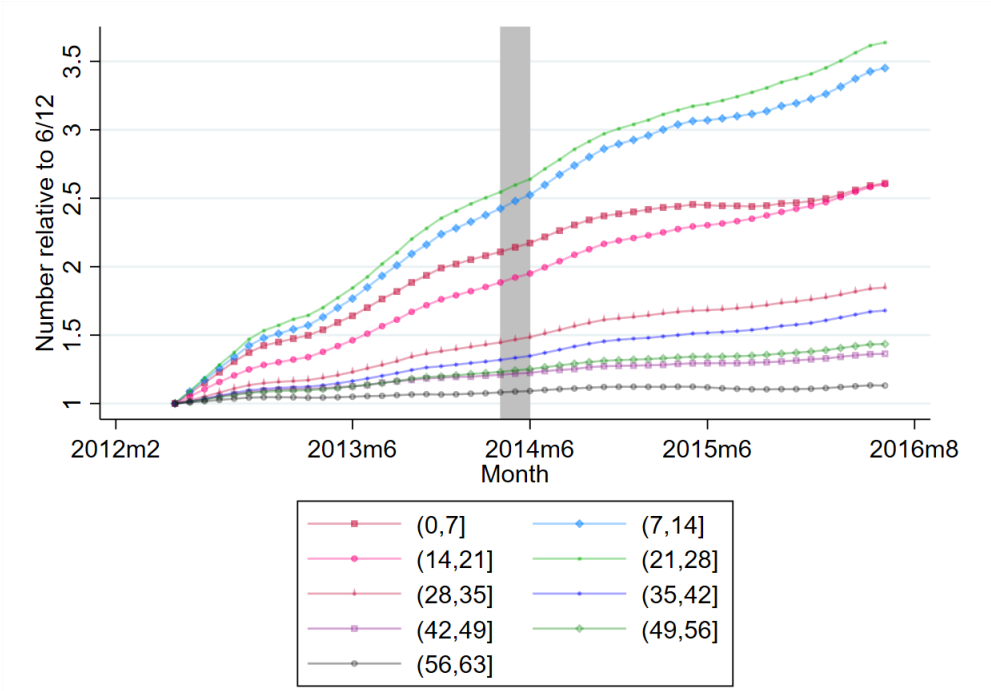
Our identification strategy relies on first finding a region of the reported income distribution which was not impacted by the reform. In the context of the baseline model in Section 2.1, the number of single individual households reporting an income strictly below R\$70 should be unchanged by the reform. The intuition for this is that anyone who has a true income below R\$70 always reports truthfully both pre- and post-reform and anyone who has a true income above R\$70 prefers to misreport at the threshold rather than misreport to an income level below the threshold (because misreporting costs are increasing in the distance between true and reported incomes). While in theory bunching should occur precisely at R\$70, in reality bunching is typically more diffuse due to small optimization errors and/or frictions (Kleven, 2016). Thus, our first identification assumption is that the number of people reporting incomes at or below R\$63 is unaffected by the reform (i.e., we assume that R\$63 is sufficiently far below R\$70 such that there are no "bunchers" at or below R\$63):

Identification Assumption 1. *The distribution of reported incomes below R\$63 is unaffected by the reform:* $G_t(x; \mathbf{p}) = G_t(x; \mathbf{p}') \ \forall \ x \leq 63, \ t.$

To provide suggestive evidence that Assumption 1 is reasonable, Figure 3 plots the number of single individual households reporting in income bins of size 7 from R\$0 to R\$63 (with the numbers in each bin

30. Implicitly, we assume that agents are not making dynamic decisions (e.g., agents are myopic) so that we can apply Proposition 3 and calculate the welfare impact of the reform in a given time period. Note, there is another reform to the schedule in June 2016 so we cannot estimate the welfare impact of the reform beyond June 2016.

normalized to 1 in June 2012). Figure 3 shows that it is not obvious any of the bins below R$63 were impacted by the reform.[31]



Note: This figure shows the number of single individual households with reported incomes in the various bins. The number in each bin is normalized to 1 in June, 2012. For example, this means that there are 3.5 times as many single individual households with reported incomes in R$(7, 14] in June, 2016 as in June, 2012. The timing of the reform is indicated by the gray, shaded region.

FIGURE 3: NUMBER OF SINGLE INDIVIDUAL HOUSEHOLDS REPORTING INCOMES IN VARIOUS BINS, NORMALIZED TO 1 IN JUNE, 2012

Thus, Identification Assumption 1 implies that $G_{\bar{t}}(63;\mathbf{p}) = G_{\bar{t}}(63;\mathbf{p}')$. This in turn implies that $J_{\bar{t}}$ and $B_{\bar{t}}$ can be expressed solely in terms of how the number of households reporting incomes in bins R$(63, 70] and R$(70, 77] changed as a result of the reform:[32]

$$\begin{aligned} B_{\bar{t}} =& - \left[G_{\bar{t}}(70;\mathbf{p}') - G_{\bar{t}}(70;\mathbf{p})\right] \\ =& - \left(\left[G_{\bar{t}}(70;\mathbf{p}') - G_{\bar{t}}(63;\mathbf{p}')\right] - \left[G_{\bar{t}}(70;\mathbf{p}) - G_{\bar{t}}(63;\mathbf{p})\right]\right) \end{aligned} \tag{12}$$

$$\begin{aligned} J_{\bar{t}} =& G_{\bar{t}}(77;\mathbf{p}') - G_{\bar{t}}(77;\mathbf{p}) \\ =& \left[G_{\bar{t}}(77;\mathbf{p}') - G_{\bar{t}}(70;\mathbf{p}')\right] - \left[G_{\bar{t}}(77;\mathbf{p}) - G_{\bar{t}}(70;\mathbf{p})\right] + \\ & \left[G_{\bar{t}}(70;\mathbf{p}') - G_{\bar{t}}(63;\mathbf{p}')\right] - \left[G_{\bar{t}}(70;\mathbf{p}) - G_{\bar{t}}(63;\mathbf{p})\right] \end{aligned} \tag{13}$$

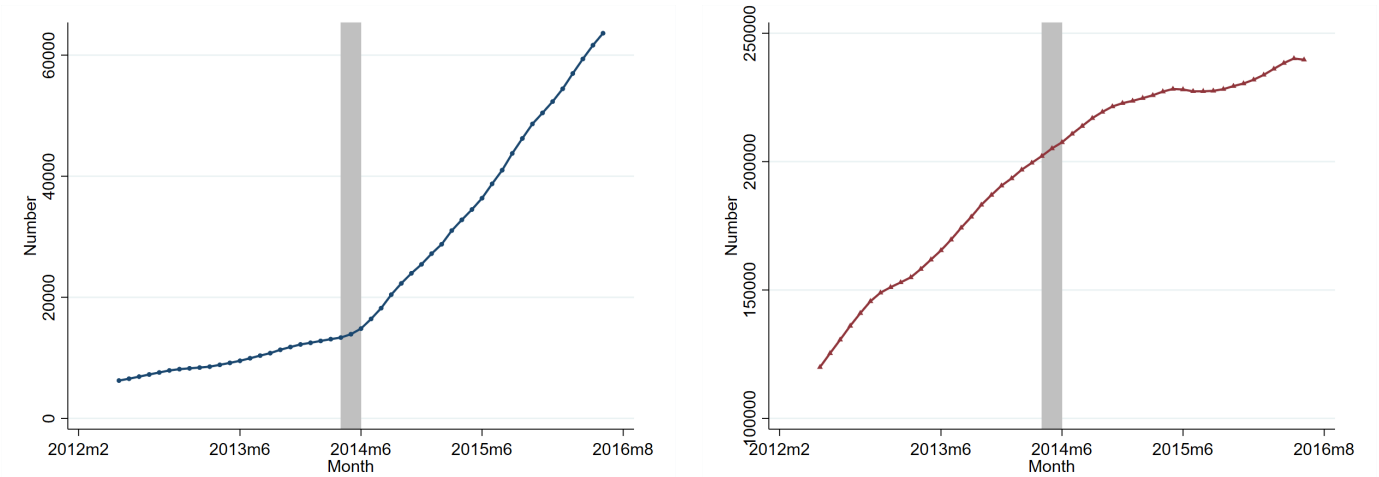
In particular, $B_{\bar{t}}$ is equal to the reduction in households locating in R$(63,70] as a result of the reform (i.e., the reduction in households bunching at the old notch), while $J_{\bar{t}}$ is equal to the increase in households locating in R$(70,77] less the reduction in households locating in R$(63,70] as a result of the reform (i.e., the increase in households bunching at the new notch less the reduction in households bunching at the old notch).

31. We discuss reasons why Identification Assumption 1 may fail to hold and explore robustness to Identification Assumption 1 in Section 5.3.

32. The total cost of the reform can be estimated from the components necessary to estimate $B_{\bar{t}}$ and $J_{\bar{t}}$ as Total Cost $= 77G_{\bar{t}}(77;\mathbf{p}') - 70G_{\bar{t}}(70;\mathbf{p}) = 77[G_{\bar{t}}(77;\mathbf{p}') - G_{\bar{t}}(70;\mathbf{p}') + G_{\bar{t}}(70;\mathbf{p}') - G_{\bar{t}}(63;\mathbf{p}')] + 70[G_{\bar{t}}(70;\mathbf{p}) - G_{\bar{t}}(63;\mathbf{p})] + 7G_{\bar{t}}(63;\mathbf{p}')$.

Figure 4 shows how the number of households reporting incomes in bins R$(63, 70] and R$(70, 77] evolved in a four year window around the reform from June 2012 to June 2016. In contrast to Figure 3, Figure 4a depicts a clear trend departure commensurate with the reform. Thus, Figure 4a provides highly suggestive evidence that the BF reform induced a substantial behavioral response that increased the number of households bunching at the new threshold. Note that the number of individuals reporting incomes in R$(70, 77] is increasing over time prior to the reform - this is likely due to a combination of factors including population growth and income dynamics.[33] In particular, Brazil was experiencing slowing and then negative per-capita GDP growth from 2012-2016 with a large drop in 2015 (World Bank Data). However, the slope of the line in Figure 4a clearly changes at the time of the reform, leading to a sharp increase in the number of individuals locating in R$(70, 77]. The fact that this reform induced a slope shift rather than a level shift is likely due to households learning about the threshold shift over time.[34]

In contrast to Figure 4a, there is less convincing *prima facie* that the reform induced a substantial change in the number of households locating in R$(63,70] - see Figure 4b. While Figure 4b may indicate that the number of households in R$(63,70] declined as a result of the reform, underlying time trends make it difficult to assess the extent to which this is the case. Hence, we need a way to control for underlying time trends in the numbers reporting in R$(63, 70] and R$(70, 77] to estimate how these quantities would have evolved over time *in absence of the reform*. This brings us to our second identification assumption.



(a) Number Reporting in R$(70,77]

(b) Number Reportig in R$(63,70]

Note: This figure shows the number of single individual households that report incomes in the intervals R$(63, 70] and R$(70, 77] for each month between June 2012 to June 2016. The timing of the reform (from the announcement in April 2014 to the enactment in June 2014) is indicated by the gray, shaded region.

Figure 4: Number of Single Individual Households Reporting an Income in R$(63,70] and R$(70,77]

33. Similarly, the number of households on the Cadastro Único registry is growing over this time period - see Figure 12 in Appendix C.1. This too is likely due to a variety of factors including population growth, a struggling Brazilian economy, and increased awareness/understanding of the BF program over time.

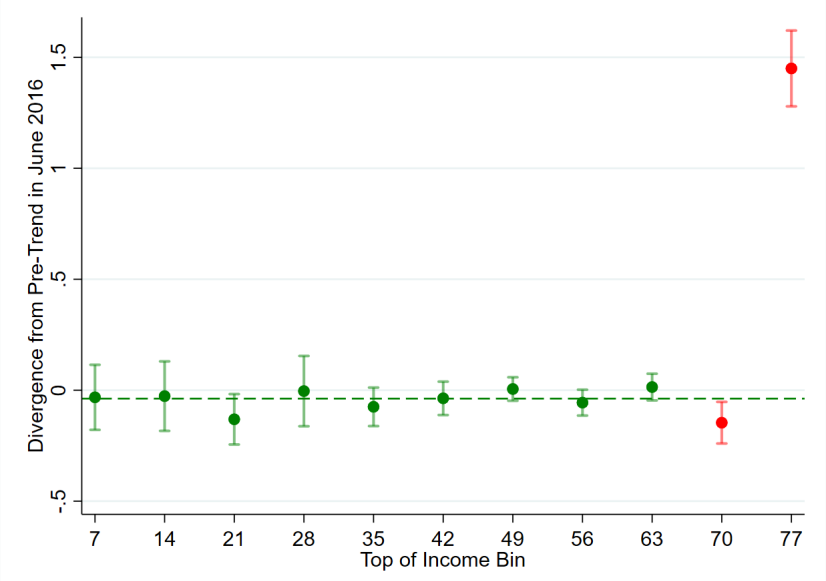
34. Alternatively, households may wait to update simply because they are only required to update every two years and/or to avoid suspicion of misreporting that could result from updating immediately after the reform.

4.2 Identification Assumption 2

At a high level, our second identification assumption is going to relate how the numbers reporting in bins below R\$63, R\$(0,7],....,R\$(56,63], evolve over time to how the numbers reporting in our two bins of interest, R\$(63, 70] and R\$(70, 77], would have evolved over time in absence of the reform. In this sense, bins below R\$63 can be viewed as our "control bins" while R\$(70,77] and R\$(63,70] can be viewed as our "treatment bins". Building towards our second identification assumption and our main empirical specification, consider the following regression, where $N_{(x-7,x],t}$ denotes the number of individuals reporting in income bin R\$$(x-7, x]$ in month t:

$$\underbrace{\log\left(N_{(x-7,x],t}\right)}_{\log \# \text{ in } (x-7,x]} = \underbrace{\sum_{k=0}^{K} \alpha_{k,x} t^k}_{\text{bin-specific polynomial}} + \underbrace{\beta_{1,x} post_t + \beta_{2,x} post_t \times t}_{\text{post-reform deviation from polynomial}} + \epsilon_{xt} \tag{14}$$

where $post_t$ takes value 1 if month t is after the reform and 0 otherwise (i.e., $post_t = 1$ if $t \geq$June 2014 and 0 otherwise). Regression (14) simply estimates a break from the (bin-specific) polynomial time trend in the post-reform period for the (log) number of people reporting in R\$$(x-7, x]$. Running Regression (14) separately for each bin (using cubic bin-specific polynomials), Figure 5 plots $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}\bar{t}$ for each $x \in \{7, 14, ..., 77\}$ where $\bar{t}$ represents the final month in our analysis period (June, 2016). In other words, Figure 5 plots how the log number of people reporting in each seven increment bin below R\$77 deviated from its bin-specific cubic time trend in the post-reform period:



Note: This figure plots how the log number of people reporting in each seven increment bin deviated in the post-reform period from a bin-specific, cubic time trend estimated via Regression (14) along with 95% confidence intervals using robust standard errors. For each $x \in \{7, 14, ..., 77\}$, we estimate the deviation from the bin-specific time trend in June, 2016 as $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}\bar{t}$, where $\bar{t}$ represents the final month in our analysis period (June, 2016). The green horizontal line plots the average deviation (-0.038) from trend for bins with $x \leq R\$63$.

FIGURE 5: DEVIATION FROM CUBIC TIME TREND IN POST-REFORM PERIOD BY INCOME BIN

Figure 5 shows that the number of people reporting in income bins below R\$63 saw very minor (and mostly statistically insignificant) trend breaks in the post-reform period. The dashed green line in Figure 5

indicates an average trend break of -0.038 across these bins, indicating that the number of people reporting incomes in bins $\{R\$(0,7], R\$(7,14], ..., R\$(56,63]\}$ was, on average, approximately 3.8% lower in June, 2016 relative to polynomial trend. However, under Identification Assumption 1, any post-reform trend break for these bins must be due to underlying time variation unrelated to the reform. Loosely speaking, our second identification assumption is going to be that our two treatment bins, R\$(63, 70] and R\$(70, 77], would have seen the same deviation post-reform as the control bins had the reform not happened, i.e., these two bins also would have seen a 3.8% reduction in the number of people in June, 2016 (relative to their bin-specific polynomial time trends) had the reform not happened. However, as can be seen in Figure 5, the number of people reporting incomes in R\$(63, 70] saw a reduction of around 13.6% (a 0.146 log-point decrease) while the number of people reporting incomes in R\$(70, 77] saw an increase of 326.3% (a 1.45 log-point increase) as of June 2016. Formally, our second identification assumption is:

Identification Assumption 2. *In absence of the reform, the (log) number of people reporting in each bin evolves according to:*

$$\log\left(N_{(x-7,x],t}\right) = h(t) + \sum_{k=0}^{K} \alpha_{k,x} t^k + \nu_{x,t} \quad \textit{for } x \in \{7, 14, ..., 77\}$$

Under Identification Assumption 2, the difference between the (log) number of households in any two 7-increment bins below R\$77 is, *in absence of the reform*, governed by a stable polynomial plus a random error term. Combining Assumptions 1 and 2, we can use our control bins (i.e., bins ≤R\$63) to identify $h(t)$ in the post-reform period.[35] In other words, we can use our control bins to identify the expected deviation from bin-specific polynomials in the post-reform period if the reform did not occur. Any deviation observed above and beyond $h(t)$ in our two treatment bins is then attributed to the reform. This brings us to our main empirical specification: a generalized difference-in-difference specification where we allow for flexible pre-treatment dynamics between treatment and control bins:

$$\log\left(N_{(x-7,x],t}\right) = \delta_t + \sum_{k=0}^{K} \alpha_{k,x} t^k + [\beta_{1,x} post_t + \beta_{2,x} post_t \times t] \times treat_x + \epsilon_{xt} \text{ for } x \in \{7, 14, ..., 77\} \qquad (15)$$

where δ_t represents a set of month fixed-effects; $treat_x$ takes value 1 if $x \in \{70, 77\}$ and 0 otherwise; and $\sum_{k=0}^{K} \alpha_{k,x} t^k$ captures polynomial time trends for each bin (indexed by x) which predate the reform and are assumed to persist into the post-reform period had the reform not occurred. Under our two identification assumptions, the causal impact of the reform on the (log) number of households locating in $R\$(x-7, x]$ in month t, denoted $\Delta \log\left(N_{(x-7,x],t}\right)$, is equal to $\hat{\beta}_{1,x} + \hat{\beta}_{2,x} t$ for $x \in \{70, 77\}$.

Before we present the results from Equation (15) in Section 5 (using various polynomial degrees K), let

35. We use the distribution below R\$63 rather than the distribution above R\$77 to control for underlying time trends because the distribution above R\$77 is presumably impacted by the reform given that the jumping households would have located above R\$77 in absence of the reform. However, reported incomes *sufficiently far* above R\$77 should not be impacted by the reform. The question is, how far is sufficiently far? Kleven and Waseem (2013) use an iterative procedure to determine how far above a notch is sufficiently far so that the distribution beyond this point is unaffected by the notch. Their procedure relies on equating the excess mass at and below the notch with the missing mass above the notch. However, such an exercise is not feasible in our setting due to extensive margin responses. In particular, while some of the "excess" households locating below R\$77 would have located above R\$77 in absence of the reform, some of them may have simply decided to not report an income and, thus, not be on the registry. Thus, we cannot equate the excess mass below R\$77 with the missing mass above R\$77. Consistent with our selection of control bins, Kleven (2016) notes that when extensive margin responses are strong, one should consider only using data below the notch to estimate the bunching mass.

us discuss the statistical and economic interpretations of Identification Assumption 2. From a statistical perspective, Identification Assumption 2 is simply a generalization of the standard difference-in-difference assumption (this generalization is employed in, for example, Wolfers (2006) and is discussed more generally in Mora and Reggio (2013)).[36] Setting $K = 0$ results in a standard difference-in-difference estimator in which the control groups and treatment groups are assumed to have "parallel trends" pre-reform that would have persisted in the post-reform period had the reform not occurred. The standard difference-in-difference strategy therefore assumes that the difference between the treatment and control groups follow a 0^{th} order polynomial in absence of the reform (i.e., that they differ by a constant). Larger values of K require progressively weaker parallel assumptions. For instance, setting $K = 1$ requires the "parallel growths" assumption which asserts that the growth rates in the treatment and control groups over time are the same (i.e., that second and higher differences between treatment and control groups are constant over time).[37] Setting $K = 2$ leads to the "parallel accelerations" assumption which asserts that the acceleration rates in the treatment and control groups over time are the same (i.e., third and higher differences between treatment and control groups are constant over time). Setting $K > 2$ results in what Mora and Reggio (2013) refer to as the "parallel-K" assumption which asserts that the $K + 1$ and higher differences between the treatment and control groups are constant over time.

But what is the economic meaning of Identification Assumption 2? Fundamentally, there are complex structural relationships between the number of people reporting in each income bin; these relationships are determined by how the true income distribution is evolving over time, how the true income distribution is mapped to the reported income distribution, as well as growth of the BF registry over time. By making Identification Assumption 2, we are essentially assuming that the differences over time between the number of people reporting in each income bin (which are determined by unknown structural relationships) are well approximated by polynomials which would have persisted in absence of the reform. Identification Assumption 2 is not fully testable in the same way that the standard differences-in-differences assumption (that parallel pre-trends persist post treatment) is not fully testable. However, as in the case of the standard "parallel trends" assumption, we can gauge whether Identification Assumption 2 seems sensible by looking at pre-reform data. Identification Assumption 2 implies that the difference between the log number of households in any two 7-increment bins below R$77 is governed by a stable polynomial in absence of the reform. Appendix C.4 shows how the differences $\log\left(N_{(70,77]}\right) - \log\left(N_{(x-7,x]}\right)$ and $\log\left(N_{(63,70]}\right) - \log\left(N_{(x-7,x]}\right)$ evolved in the pre-reform period for $x \in \{7, 14, ..., 63\}$. These differences follow stable, low order polynomial relationships in the pre-period. While this of course does not imply that these relationships would have persisted into the post-reform period (just as parallel pre-trends do not imply that the trends would continue on a parallel path post treatment), it is at least suggestive that stable relationships exists between the various income bins.

However, because we have several control bins, we can *partially* test the validity of our two identification assumptions. In particular, under Identification Assumption 1, our control bins should be unaffected by the reform. Thus, in the post-reform period, we can test whether the number of households locating in each of our control bins can be accurately predicted by our other control bins using Identification Assumption 2. Thus, at the end of Section 5, we run a number of placebo tests, suggesting that our two identification assumptions appear to be reasonable.

36. Equation (2) in Wolfers (2006) uses an analogous generalized difference-in-difference accounting for differential quadratic pre-trends. Our Equation (15) is a combination of Equation (17) and Equation (20) from Mora and Reggio (2013).

37. A number of papers discuss the extended version of differences-in-differences under the "parallel growths" assumption (e.g., Mora and Reggio (2013), Rambachan and Roth (2020), and Bilinski and Hatfield (2019)).

4.3 Could We Use A Bunching Estimator to Bound the MVPF?

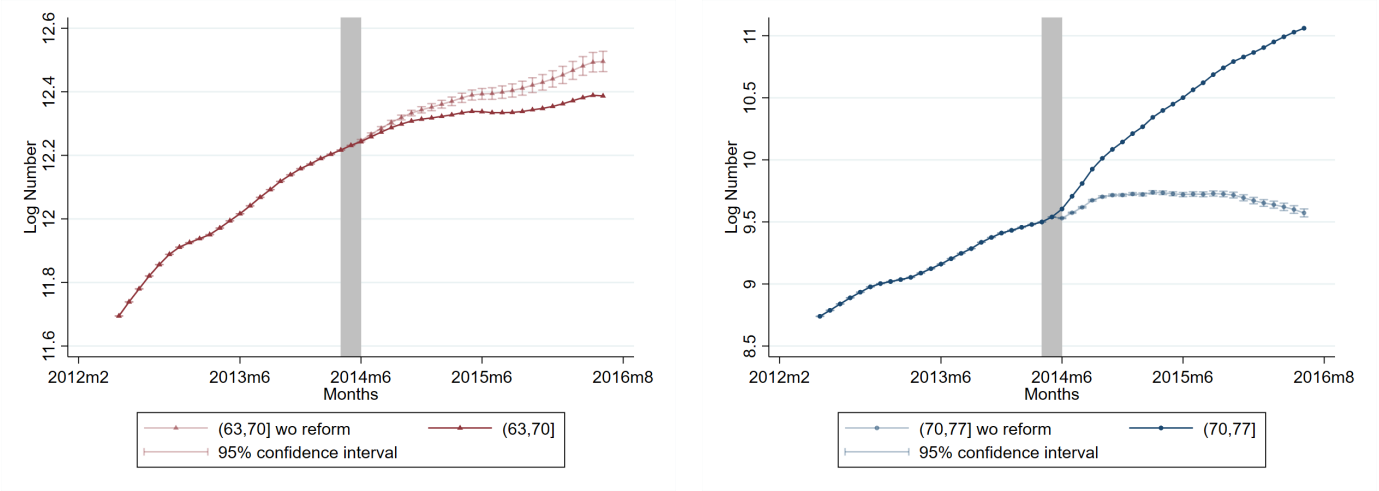
Our empirical strategy amounts to estimating the reduction in the mass of households bunching at the old notch and the increase in the mass of households bunching at the new notch as a result of the reform. Thus, one may naturally wonder whether we can estimate the required empirical objects using standard bunching techniques (as developed in Kleven and Waseem (2013) and Saez (2010)). Of course, even using standard bunching techniques, one must account for time variation, i.e., one must estimate how many people would have bunched at the old notch and at the new notch in the post-reform period had the reform not occurred. As an example of how one might use standard bunching techniques while also controlling for time variation, one could first use these techniques to back out the counterfactual reported income distributions that would result if there were no notches in both the pre- and post-reform period (e.g., in June 2014 and June 2016). To do so, one would use the observed reported income distributions in the pre- and post-reform periods and make the following two assumptions (which are standard in the bunching literature): (a) the reported income distribution would be smooth in absence of the notch, and (b) the reported income distribution sufficiently far away from the notch is unaffected by the notch. One could then assume that the relationship that exists between the counterfactual and the observed reported income distributions in the pre-reform period would have stayed the same in the post-reform period in absence of the reform. This assumption combined with the estimated counterfactual reported income distribution in the post-reform period would then allow one to calculate the reported income distribution in the post-reform period had the reform not happened. A similar strategy is employed in Carril (2021). However, this strategy is unlikely to be feasible in our setting. As seen in Figure 17 in Appendix C.3, the reported income distribution in Brazil is *extremely* non-smooth. For instance, in June 2014, in addition to substantial bunching at the notch of R$70, there is extreme bunching at numbers equal to 0 mod 50, less severe bunching at numbers equal to 0 mod 10, and substantial bunching at R$60. R$60 was a previous BF eligibility threshold (the notch of R$60 was implemented in 2006 while the notch of R$70 was implemented in 2009; see Gazola Hellmann (2015)). While bunching estimators have been augmented to deal with "round-number" bunching and other "reference-point" bunching (e.g., Kleven and Waseem (2013) and Best and Kleven (2017)), we believe that the pervasiveness and the variability of round-number and reference-point bunching in our setting will make it too difficult to precisely identify the counterfactual distributions around our notches (in particular, around the original notch of R$70 which is equal to 0 mod 10). Thus, we view estimating changes in bunching that result from a notch reform using a difference-in-difference strategy as a contribution as of this paper and hope our approach will be useful in other settings where estimating bunching directly is difficult.

5 Results

In this Section, we present results from our generalized difference-in-difference specification, Equation (15), and, in turn, calculate bounds on the MVPF. Based on these bounds, we then discuss the welfare implications of the reform. Next, we show robustness of our results to a variety of factors: we show robustness to Identification Assumption 1 by allowing the number of households between R$(56, 63]$ to also be impacted by the reform; we show robustness to a more general version of Identification Assumption 2; and, we show robustness of our results to household composition. Finally, we conduct a placebo exercise which partially tests the validity of our two identification assumptions.

5.1 *Difference-in-Difference Results*

First, we present results from estimating Equation (15) seting $K = 3$, which assumes the difference in the number of households reporting in any two income bins below R$77 approximately follows a cubic polynomial over time (in absence of the reform). Once we have estimated Equation (15), we can recover the causal impact of the reform on the log number of households in R$$(x-7, x]$ for $x \in \{70, 77\}$ in any given month t, denoted $\Delta \log\left(N_{(x-7,x],t}\right) = \hat{\beta}_{1,x} + \hat{\beta}_{2,x} t$. Figure 6 plots the evolution of $\log\left(N_{(63,70],t}\right)$ and $\log\left(N_{(70,77],t}\right)$ over time along with the estimated counterfactual path if the reform did not occur: $\log\left(N_{(63,70],t}\right) - \Delta \log\left(N_{(63,70],t}\right)$ and $\log\left(N_{(70,77],t}\right) - \Delta \log\left(N_{(70,77],t}\right)$. The difference between the actual path and the counterfactual path yields the causal impact of the reform under our two identification assumptions.



(a) $N_{(63,70]}$: Actual & Counterfactual

(b) $N_{(70,77]}$: Actual & Counterfactual

Note: This figure shows the log number of single individual households reporting incomes in R$$(63, 70]$ and R$$(70, 77]$ over time along with the counterfactual paths had the reform not happened. The counterfactual paths are equal to the actual log number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x} t$, estimated using Equation (15) where we set $treat_x = 1$ if $x \in \{70, 77\}$ and $K = 3$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

Figure 6: Causal Impact of the Reform Estimated From Equation (15)

Consistent with the theoretical model from Section 2.1, Figure 6 shows that the reform led to a large, significant increase of about 1.5 log points in the number of households reporting in bin R$$(70, 77]$ and a significant decrease of about 0.1 log points in the number of households reporting in bin R$$(63, 70]$.[38] Translating these log changes into levels, the reform led to an increase of approximately 49,000 households locating in R$$(70, 77]$ and a decrease of approximately 27,000 households locating in R$$(63, 70]$. Plugging these numbers into Equations (12) and (13)), we get that $B_{\bar{t}} \approx 27,000$ and $J_{\bar{t}} \approx 22,000$. Thus, of the additional 49,000 households bunching at the new notch, 27,000 would have bunched at the old notch had

38. All standard errors for our analysis are clustered at the bin level (using STATA's default small-number-of-clusters bias adjustment) as this is the level of "treatment assignment", see Abadie et al. (2017). Thus, we have 9 control clusters and 2 treatment clusters. However, Imbens and Kolesár (2016) shows that these standard errors are modestly under-estimated when the number of clusters is small (they find that for samples with 10 clusters, 95% confidence intervals only have a 91% coverage rate). Hence, our standard errors may be modestly overstating the true statistical confidence we have in our estimates. An alternative approach is to use non-clustered wild bootstrapped p-values (note, wild *cluster* bootstrapped p-values lead to severe bias in difference-in-difference settings with a small number of treated clusters, see Roodman et al. (2019)). In our setting, these p-values are smaller than those generated from our cluster-robust standard errors. Finally, Ferman and Pinto (2019) suggest another alternative for difference-in-difference settings with a small number of treated clusters; unfortunately, their results rely on asymptotic theory in the number of control clusters, which we believe is inappropriate for our setting given that we only have 9 control clusters.

the reform not happened, while 22,000 would have located above the new notch had the reform had not happened (i.e., 22,000 households jump into the program as a result of the reform).[39]

We repeat this exercise under the alternative assumptions that the differences in the (log) numbers in any two bins follow quadratic, quartic, or quintic polynomials over time, i.e., we re-estimate Equation (15) for $K = 2, 4, 5$.[40] Figures of the counterfactual paths for our two treatment bins under these alternative assumptions are shown in Appendix C.2. The predicted changes in the numbers reporting in the two treatment bins for $K = 2, 3, 4, 5$ are shown in Columns (1) and (2) of Table 3. Table 3 presents the estimated impact of the reform on the numbers locating in R$(63,70] and R$(70,77], $\Delta N_{(63,70],t}$ and $\Delta N_{(70,77],t}$, along with the estimated number of bunching households $B_{\bar{t}}$, estimated number of jumping households $J_{\bar{t}}$, and lower and upper bounds on the MVPF in June 2016. All of these quantities are stable across different values of K. Under our preferred specification ($K = 3$), we find a lower bound for the MVPF of the reform around 0.9 and an upper bound of around 1.12.

Table 3: Impacts of Reform and MVPF Bounds in June 2016 Estimated from Equation (15)

	(1)	(2)	(3)	(4)	(5)	(6)
Polynomial Degree, K	$\Delta N_{(63,70],\bar{t}}$	$\Delta N_{(70,77],\bar{t}}$	$B_{\bar{t}}$	$J_{\bar{t}}$	$MVPF_{L,\bar{t}}$	$MVPF_{U,\bar{t}}$
Quadratic, $K = 2$	-26,279	51,759	26,279	25,480	0.88	1.11
	(6, 163)	(2, 160)	(6, 163)	(6, 659)	(0.03)	(0.03)
Cubic, $K = 3$	-27,452	49,247	27,452	21,794	0.90	1.12
	(4,357)	(234)	(4, 357)	(4, 592)	(0.02)	(0.02)
Quartic, $K = 4$	-29,338	50,873	29,338	21,535	0.90	1.13
	(6,257)	(1, 345)	(6, 257)	(6, 503)	(0.03)	(0.03)
Quintic, $K = 5$	-29,240	50,559	29,240	21,318	0.90	1.13
	(5,912)	(1,184)	(5,912)	(6,141)	(0.03)	(0.03)

Note: Columns (1) and (2) show the estimated impacts of the reform on the number of single individual households reporting incomes in bins R$(63,70] and R$(70,77] for June 2016: $\Delta N_{(63,70],\bar{t}}$ and $\Delta N_{(70,77],\bar{t}}$. Estimates are calculated using Equation (15) with various polynomial degrees $K \in \{2, 3, 4, 5\}$. Columns (3) and (4) show the estimated number of bunching and jumping households for June 2016, $B_{\bar{t}}$ and $J_{\bar{t}}$, calculated using Equations (12) and (13). Columns (5) and (6) show the estimated upper and lower bounds for the MVPF for June 2016, calculated using Equations (10) and (11). Standard errors are presented in parentheses and are computed from the delta method from the clustered standard errors estimated in Equation (15).

5.2 Discussion

Our lower bound on the MVPF of the June 2014 BF reform is approximately 0.90 whereas our upper bound on the MVPF is approximately 1.12. This implies that the WTP for the reform relative to the budgetary cost is between 0.90 and 1.12. What can we say about the welfare implications of an MVPF between 0.90 and 1.12? We can be sure that the reform was welfare improving as long as the government values giving R$0.90 to BF households in a non-distortionary manner more than spending R$1 on their next best alternative. On the other hand, the reform was welfare decreasing if the government values spending R$1

39. One may then wonder whether we estimate a reduction of 22,000 households locating above R$77 as a result of the reform. To investigate this, we can include income bins above R$77 in Equation (15) as additional *treatment* bins. Cumulatively, we estimate that income bins between R$77 and R$147 saw a decline of around 14,300 households when setting $K = 3$ (with a standard error of around 6800). However, as mentioned in footnote 35, the set of jumping households consists of both households who would have located above R$77 in absence of the reform *and* households who would not have reported an income to the registry at all in absence of the reform (i.e., new entrants). Because of these new entrants, it is not surprising that we estimate the number of jumping households as exceeding the reduction in households reporting above R$77.

40. Our results are also robust to estimating Equation (15) with more granular income bins, see Appendix C.5.

on their next best alternative more than giving R$1.12 to BF households in a non-distortionary manner.[41]

But what is the government's best alternative use of funds? In general, this is a difficult question to answer. Lindert et al. (2007) show that BF performs better (in terms of targeting performance) than other anti-poverty programs in both Brazil and other countries in Latin America; this suggests that there is not a better alternative anti-poverty program that the government could finance. Hence, for expositional simplicity we assume that the next best use of funds is a non-distortionary UBI. Under this assumption, we can perform a conservative back-of-the-envelope calculation to argue that the 2014 BF reform was almost certainly welfare improving. Suppose the government is utilitarian and that the true income distribution is log-normal with parameters taken from the World Bank's PovcalNet database for Brazil in 2016.[42] Conservatively, we assume that the BF household income distribution matches the distribution for the bottom 50% of income earners in Brazil.[43] Moreover, we assume utility of consumption is CRRA: $u(c) = c^{1-\gamma}/(1-\gamma)$. Based on these assumptions, the government prefers giving R$0.90 (in a non-distortionary manner) to BF households more than spending R$1 on a non-distortionary UBI as long as $\gamma > 0.14$. While estimates for γ vary widely, typically estimates fall between 1 and 10 (Outreville, 2014). Alternatively, if $\gamma = 1$ (i.e., $u(c) = \log(c)$), giving R$0.90 (in a non-distortionary manner) to BF households yields equivalent welfare gains to spending R$1.50 on a non-distortionary UBI. Thus, if the next best alternative policy is a non-distortionary UBI and the government is utilitarian, the BF reform was almost certainly welfare improving.

Finally, when calculating our bounds for the MVPF, we have ignored any potential fiscal externalities when determining the total cost of the reform (i.e., we have ignored the impact that behavioral responses have on other components of the government budget). There is evidence to suggest that some beneficiaries partially substitute formal sector employment for informal sector employment to become eligible for the BF benefit (as informal income is less easily verified making misreporting less costly; see De Brauw et al. (2015)). While it is easier to evade taxes in the informal sector, it is unlikely that this behavioral response will affect tax revenues as BF beneficiaries are almost certainly sufficiently poor so as to be exempt from income taxation.[44] However, there is evidence showing that increased BF transfers lead to an increase in formal sector employment of non-beneficiaries, which leads to higher GDP and tax revenue (Gerard, Naritomi and Silva, 2021). If this positive fiscal externality exists for the June 2014 reform, this would lower our estimate for the total cost of the reform, thus raising our estimates for both the lower and upper bound of the MVPF (as the denominator of both bounds would be reduced). Thus, incorporating this positive fiscal externality would only reinforce our finding that the 2014 BF reform was welfare improving relative to a non-distortionary UBI.

41. Technically, the reform is welfare improving as long as the government values spending R$1 on their next best alternative less than splitting R$0.90 (in a non-distortionary manner) among the mechanical, threshold, bunching, and jumping households, where the split is determined by the lower bounds on each group's WTP for the reform. Similarly, the reform is welfare decreasing if the government values spending R$1 on their next best alternative more than splitting R$1.12 (in a non-distortionary manner) among the mechanical, threshold, bunching, and jumping households, where the split is determined by the upper bounds on each group's WTP for the reform.

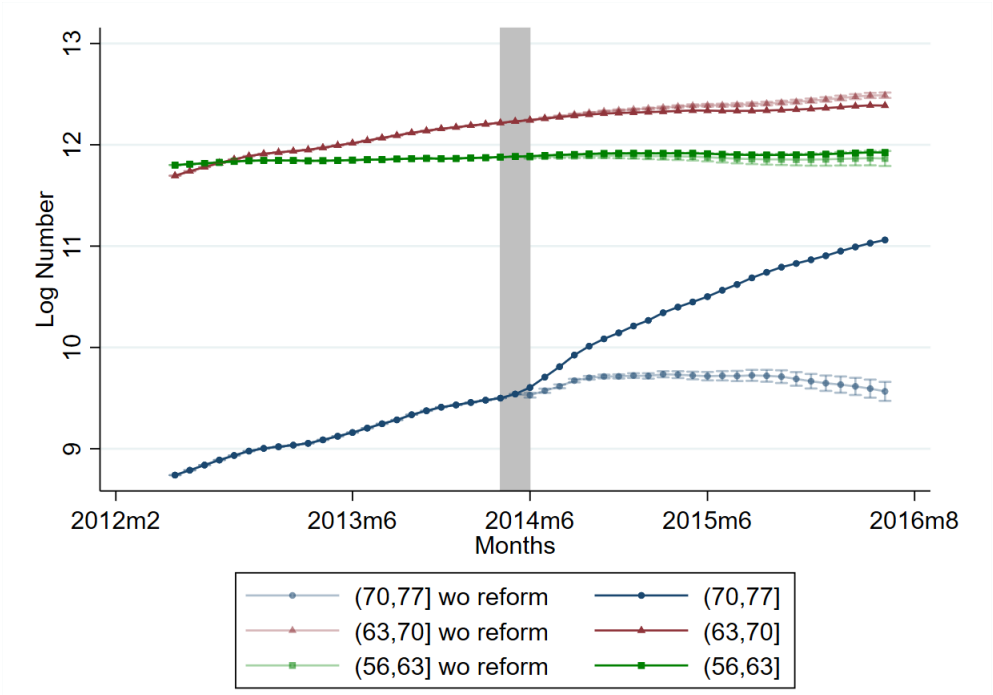
42. In 2016, the mean per-capita income in Brazil was R$615.32 and the Gini coefficient was 0.533; these parameters fully pin down the income distribution if we assume log-normality.

43. Bastagli (2008) finds that of all transfers paid in 2004, 91% went to households in the bottom 50% of the income distribution. Alternatively, Lindert et al. (2007) suggests even better targeting performance, finding that 94% of all benefits go to the bottom 40%. Notably, transfers are not evenly distributed across the bottom 50% of households: the poorest households within this group receive more (e.g., Bastagli (2008) finds that the bottom 20% of the income distribution receive 50% of BF transfers, while Lindert et al. (2007) report that the bottom 20% receive 73% of BF transfers). Thus, we believe the assumption that the income distribution of BF households is the same as the income distribution for the bottom 50% of income earners in Brazil is highly conservative.

44. See footnote 6.

5.3 Robustness to Identification Assumption 1

Identification Assumption 1 may not hold if optimization frictions are large. For example, consider households who, in response to the reform, jump below the threshold by changing their labor supply but face large labor market frictions. These households may not be able to perfectly jump and bunch at the new notch, but may instead jump to an income level below R$63. Alternatively, under the original policy, some households may wish to bunch at R$70, but are only able to locate at an income below R$63 due to labor market frictions; if these households are able to update their income to an income between R$(70,77], they may move to the new notch as a result of the reform. Hence, large labor market frictions may imply that the distribution below R$63 is impacted by the reform. To test for this possibility (or equivalently, to allow for larger frictions), we relax Identification Assumption 1 by instead assuming that the number of people reporting in bins ≤R$56 are unaffected by the reform, allowing for the possibility that the number reporting in bin R$(56, 63] is affected by the reform. We augment Equation (15) by setting $treat_x = 1$ if $x \in \{63, 70, 77\}$ and 0 otherwise, i.e., we include R$(56, 63] as a treatment bin. Figure 7 plots the actual and counterfactual paths for $\log\left(N_{(56,63],t}\right)$, $\log\left(N_{(63,70],t}\right)$, and $\log\left(N_{(70,77],t}\right)$. It does not appear that the reform had a significant impact on the number reporting in R$(56, 63] and the impact of the reform on our two original treatment bins is robust to including R$(56,63] as a treatment bin. Using this relaxation of Identification Assumption 1, we find that the MVPF is bounded between 0.82 and 1.07 (see Table 7 in Appendix C.6).



Note: This figure the log number of one single individual households reporting incomes in R$(56, 63], R$(63, 70], and R$(70, 77] over time along with the counterfactual paths had the reform not happened. The counterfactual paths are estimated using Equation (15) (where we set $treat_x = 1$ if $x \in \{63, 70, 77\}$, i.e., we have three treatment bins: R$(56, 63], R$(63, 70], and R$(70, 77]). We assume the difference in the (log) number reporting in any two bins follows a cubic time trend in absence of the reform (i.e., we set $K = 3$ in Equation (15)). Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

FIGURE 7: COUNTERFACTUAL PATHS FOR THREE TREATMENT BINS USING EQUATION (15)

5.4 Robustness to Identification Assumption 2

Identification Assumption 2 essentially boils down to assuming that the differences over time between the (log) number of people reporting in each income bin (which are determined by unknown, complex structural relationships) are well approximated by polynomials which would have persisted in absence of the reform. Identification Assumption 2 therefore implies that if, for example, bin R$(35,42] experiences a 1% deviation from its bin-specific polynomial trend, bin R$(63,70] would also experience a 1% deviation from its bin-specific polynomial trend in absence of the reform. However, one may argue that, for example, a structural change in the economy which leads to a 1% deviation from the bin-specific polyomial time trend for bin R$(35,42] would lead to a 2% deviation for bin R$(63,70]. As such, we relax Identification Assumption 2 to instead assume that the log number reporting in each bin evolves according to:

$$\log\left(N_{(x-7,x],t}\right) = \gamma_x h(t) + \sum_{k=0}^{K} \alpha_{k,x} t^k + \nu_{x,t} \quad \text{for } x \in \{7, 14, ..., 77\}$$

Thus, we now assume that deviations from underlying bin-specific polynomials (in logs or, equivalently, percentage terms) are not necessarily the same across all bins. This relaxation of Identification Assumption 2 allows for a structural change in the economy which creates a 1% deviation from bin-specific polynomial trend for R$(35,42] to create a $(\gamma_{70}/\gamma_{42})$% deviation from bin-specific polynomial trend for bin R$(63,70]. Under this relaxed version of Identification Assumption 2, we estimate the following non-linear least squares regression:[45]

$$\log\left(N_{(x-7,x],t}\right) = \gamma_x \delta_t + \sum_{k=0}^{K} \alpha_{k,x} t^k + [\beta_{1,x} post_t + \beta_{2,x} post_t \times t] \times treat_x + \epsilon_{xt} \text{ for } x \in \{7, ..., 77\} \qquad (16)$$

Essentially, running Regression (16) will estimate a common set of month dummies, δ_t, as well as bin-specific factors, γ_x, which multiply these common month dummies for each bin x. We show in Appendix C.7 that our results are very robust to this relaxation of Identification Assumption 2.

5.5 Robustness to Household Composition

We also consider the possibility that some of the estimated behavioral impact of the reform for single individual households may be coming from households misreporting their family composition. For example, a two adult household can receive greater benefits if they report to be two separate one adult households as benefits are paid out per-household as opposed to per-capita.[46] Hence, the reform may have increased incentives to misreport family composition as well as income.[47] From a theory perspective, Proposition 3 holds even if misreporting responses occur on the family composition margin. However, such a behavioral response may affect the validity of our identification strategy (for example, it may no longer be reasonable

45. Clearly, Identification Assumption 2 is nested within this more general functional form.

46. E.g., a household with two adults and no children that has a combined per-capita income of R$60 is eligible for R$70 in transfers pre-reform and R$77 in transfers post-reform. However, if this household were to report that they were actually two single individual households with incomes of R$60, they would *each* be eligible for R$70 in transfers pre-reform and R$77 in transfers post-reform.

47. However, the ability of households to misreport the number of members is limited as individuals must provide government issued IDs for all family members to be on the registry. Moreover, household composition is arguably more verifiable than income for many households given the large informal sector in Brazil. Thus, we suspect that households are more likely to misreport income than family composition.

to assume that the distribution below R$63 is unaffected by the reform if these new "single individual" households enter at income levels well below the threshold). Thus, we re-do our main analysis restricting to single individual households whose composition does not vary over the sample period. Our point estimates for the lower bound for the MVPF still lie between 0.88 and 0.91, while our point estimates for the upper bound are now smaller lying between 1.01 and 1.05; see Appendix C.8.

Finally, our main analysis is restricted to single individual households due to the fact that the schedule for households with more than one individual features a notch at R$70 and a kink below R$70 (e.g., this kink is at R$35 for two adult households with no children; see Section 3). Consequently, the 2014 BF reform led to both a change in the notch *and* a change in the kink for households with more than one individual. Thus, for these households, it is not necessarily reasonable to make Identification Assumption 1 as we might expect the reform to affect the number of people reporting in bins around the kink. We nonetheless proceed with estimating the bounds on the MVPF for two adult households without children, noting that our identification assumptions are imperfect in this setting. For these households we see the same pattern of behavioral responses to the reform and estimate lower bounds for the MVPF between 0.88 and 0.96 and upper bounds for the MVPF between 1.08 and 1.12; see Appendix C.9.[48] [49]

5.6 Placebo Tests

We can *partially* test the validity of our two identification assumptions via placebo tests. Together, Identification Assumptions 1 and 2 imply that all of the income bins below R$63 should evolve approximately according to a common time trend plus a bin-specific polynomial throughout the entire analysis period. To test this, we re-estimate Equation (15) but only include bins below R$63 (i.e., $x \in \{7, 14, ..., 63\}$) and randomly assign some of these bins to be "treatment" bins (i.e., $treat_x = 1$). If both of our identification assumptions are correct, the "treatment effects" from these placebo regressions should be close to zero, i.e., the post-reform deviation of the "treated" bins relative to the post-reform deviation of the "control" bins should be close to 0. There are a total of 510 different ways to assign our nine bins below R$63 "treatment" status.[50]

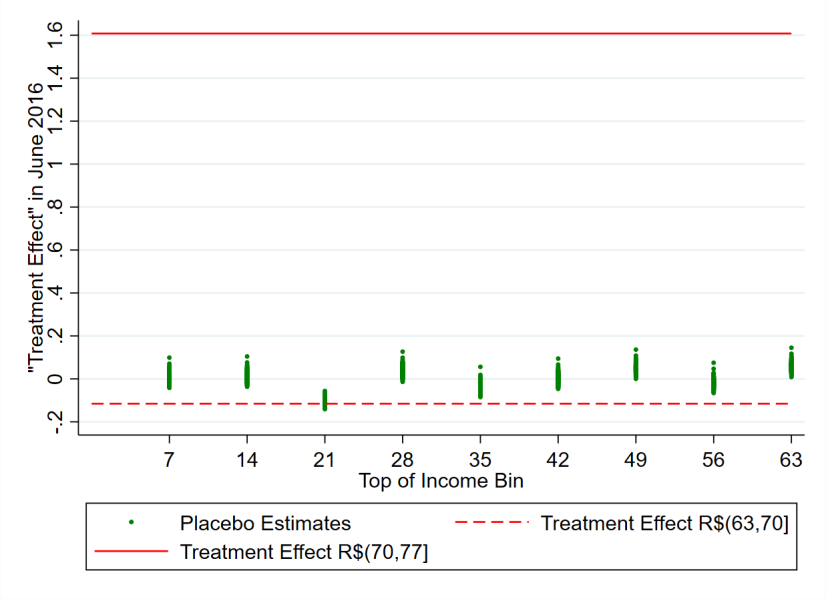
Figure 8 plots the results from all 510 such regressions, showing the estimated "treatment effects" in June 2016: $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}\bar{t}$ for $x \in \{7, .., 63\}$, where $\bar{t}$ represents June 2016. Each bin gets assigned $treat_x = 1$ 255 times meaning we have 255 "treatment effect" estimates for each bin. Figure 8 shows a fairly tight clustering of these treatment effects around zero: the mean "treatment effect" (in absolute value) across all bins is around 4% (0.04 log points). Hence, the "treatment effects" are relatively small for bins below R$63.[51] This suggests that Identification Assumptions 1 and 2 are fairly reasonable.

48. Note these bounds are for the MVPF associated with changing the location and level of the notch only, i.e., we ignore any welfare impacts of changing the level and location of the kink.

49. In Appendix C.10 we show strong suggestive evidence of a behavioral response to the change in the basic benefit notch for households with children, but we do not attempt to bound the MVPF for these households. This is because households with children may also receive the "variable benefit" which is a cash transfer made conditional on investing in the health and education of their children. Both the level and location of the notch associated with the variable benefit also changed with the June 2014 reform. Thus, the WTP of households with children for the June 2014 reform needs to account for the WTP for changes to the variable benefit schedule in addition to changes in the basic benefit schedule. This exercise is beyond the scope of the current paper.

50. There are $\binom{9}{1}$ ways to pick one "treatment" bin, $\binom{9}{2}$ ways to pick two "treatment" bins; thus, there are a total of $\sum_{i=1}^{8}\binom{9}{i}$= 510 possible regressions.

51. We find a negative "average treatment effect" of -0.105 log points for R$(14, 21]. However, R$(14, 21] seems implausibly far away from the notch to be impacted by the reform, especially given that other bins around R$(14,21] do not appear to be impacted by the reform. Hence, we interpret this negative "average treatment effect" simply as random variation.



Note: This figure shows the placebo "treatment effects" (in logs) for each income bin $R\$(x-7,x]$ obtained from estimating Regression (15) with only bins $x \in \{7,14,...,63\}$ and every possible assignment of $treat_x \in \{0,1\}$ for each $x \in \{7,14,...,63\}$. There are $510 = \sum_{i=1}^{8} \binom{9}{i}$ such regressions and we plot the values of $\hat{\beta}_{1,x} + \hat{\beta}_{2,x} \times \bar{t}$ for each bin R\$(x-7,x] whenever bin R\$(x-7,x] is assigned to be a treatment bin ($\bar{t}$ represents June, 2016). We use bin-specific cubic polynomials in each regression, i.e., $K = 3$. The solid red line shows the estimated treatment effect for R\$(70,77] and the dashed red line shows the estimated treatment effect for R\$(63,70], both from our preferred specification ($J = 3$).

FIGURE 8: PLACEBO ANALYSIS: CALCULATING "TREATMENT EFFECTS" FOR EACH CONTROL BIN

Finally, we can also use these placebo tests as an alternative gauge of the statistical significance of our estimated treatment effects for income bins R\$(70,77] and R\$(63,70] from Regression (15). Under Identification Assumptions 1 and 2, the post-reform deviations from bin-specific trend plotted in Figure 8 are due to randomness. Hence, the distribution of "treatment effects" in Figure 8 can be used to construct p-values for the estimated treatment effects for income bins R\$(70,77] and R\$(63,70] from Regression (15) in a manner akin to randomization inference.[52] In Figure 8, the solid red line shows the treatment effect of about 1.6 log points for R\$(70,77] while the dashed red line shows the treatment effect of about -0.12 log points for R\$(63,70], both from our preferred specification ($K = 3$). The treatment effect for R\$(70,77] is an order of magnitude larger than any estimated "treatment effect" from placebo regressions, giving us high certainty that the number of individuals reporting incomes in R\$(70,77] was impacted by the reform. Similarly, the treatment effect for R\$(63,70] is larger (in absolute magnitude) than 97.8% of the placebo "treatment effects", implying a p-value of 0.022 against the null hypothesis that R\$(63,70] was not impacted by the reform.

6 CONCLUSION

This paper proposes a theoretical framework and empirical strategy to gauge the welfare impacts of reforms that change notches in social welfare programs. We show that we can use two empirically observable objects to bound the MVPF of the reform: the number of households who jump down to the new notch and the number of households bunching at the old notch who move toward the new notch. These bounds on the MVPF can then be used to construct bounds for the welfare impact of changing the notch. We estimate bounds for the MVPF of a reform to the Bolsa Família program that changed the structure of the notch,

52. See Imbens and Rubin (2021) or Young (2018) for discussions of randomization inference.

finding a lower bound of 0.9 and an upper bound of 1.12. Back-of-the-envelope calculations suggest that this reform was likely welfare improving.

More broadly, we believe that this paper highlights a new manner in which reduced-form evidence on jumping and bunching can be used to inform policy. Given the ubiquity of notches, we hope that the methods developed in this paper will be useful for analyzing reforms in a variety other contexts such Medicaid, income-dependent tax credits, or firm tax schedules. Finally, the techniques in this paper may be helpful for bounding welfare impacts of other government programs which do not necessarily feature notches.

References

Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?"

Bachas, Pierre, and Mauricio Soto. 2021. "Corporate Taxation under Weak Enforcement." *American Economic Journal: Economic Policy*, 13(4): 3671.

Bastagli, Francesca. 2008. "The design, implementation and impact of conditional cash transfers targeted on the poor: An evaluation of Brazil's Bolsa Família." *Thesis submitted to the London School of Economics for the degree of Doctor of Philosophy.*

Bergstrom, Katy, and William Dodds. 2021. "Optimal taxation with multiple dimensions of heterogeneity." *Journal of Public Economics*, 200: 104442.

Best, Michael Carlos, and Henrik Jacobsen Kleven. 2017. "Housing Market Responses to Transaction Taxes: Evidence From Notches and Stimulus in the U.K." *The Review of Economic Studies*, 85(1): 157193.

Best, Michael Carlos, James S Cloyne, Ethan Ilzetzki, and Henrik J Kleven. 2020. "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches." *The Review of Economic Studies*, 87(2): 656690.

Bilinski, Alyssa, and Laura Hatfield. 2019. "Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions." *Working Paper.*

Carril, Rodrigo. 2021. "Rules Versus Discretion in Public Procurement." *Working Paper.*

De Brauw, Alan, Daniel O. Gilligan, John Hoddinott, and Shalini Roy. 2015. "Bolsa Família and Household Labor Supply." *Economic Development and Cultural Change*, 63(3): 423457.

Feldstein, Martin. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics*, 81(4): 674680.

Ferman, Bruno, and Cristine Pinto. 2019. "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity." *The Review of Economics and Statistics*, 101(3): 452467.

Gazola Hellmann, Aline. 2015. "How does Bolsa Familia Work? Best Practices in the Implementation of Conditional Cash Transfer Programs in Latin America and the Caribbean." *Inter-American Development Bank Technical Note*, 856.

Gerard, Franois, Joana Naritomi, and Joana Silva. 2021. "Cash Transfers and Formal Labor Markets: Evidence from Brazil." *Policy Research Working Papers.*

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies*." *The Quarterly Journal of Economics*, 135(3): 12091318.

Imbens, Guido, and Donald B. Rubin. 2021. *Causal inference for statistics, social, and biomedical sciences: an introduction.* Cambridge University Press.

Imbens, Guido W., and Michal Kolesár. 2016. "Robust Standard Errors in Small Samples: Some Practical Advice." *Review of Economics and Statistics*, 98(4): 701712.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics*, 8(1): 435464.

Kleven, Henrik J., and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *The Quarterly Journal of Economics*, 128(2): 669723.

Lindert, Kathy, Anja Linder, Jason Hobbs, and Bénédicte de la Brière. 2007. "The Nuts and Bolts of Brazil's Bolsa Família Program: Implementing Conditional Cash Transfers in a Decentralized Context." *World Bank Social Protection Discussion Paper*, 0709.

Lockwood, Ben. 2020. "Malas notches." *International Tax and Public Finance*, 27(4): 779804.

Mayshar, Joram. 1991. "Taxation with Costly Administration." *The Scandinavian Journal of Economics*, 93(1): 75.

Mora, Ricardo, and Iliana Reggio. 2013. "Treatment Effect Identification Using Alternative Parallel Assumptions." *Working Paper*.

Outreville, J. Francois. 2014. "Risk Aversion, Risk Behavior and Demand for Insurance: A Survey." *Journal of Insurance Issues*, 37(2): 158186.

Rambachan, Ashesh, and Jonathan Roth. 2020. "An Honest Approach to Parallel Trends." *Working Paper*.

Roodman, David, Morten Ørregaard Nielsen, James G. Mackinnon, and Matthew D. Webb. 2019. "Fast and wild: Bootstrap inference in Stata using boottest." *The Stata Journal: Promoting communications on statistics and Stata*, 19(1): 460.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180212.

Slemrod, Joel. 2001. "A General Model of the Behavioral Response to Taxation." *International Tax and Public Finance*, 8: 119128.

Veras Soares, Fabio. 2011. "Brazil's Bolsa Família: A Review." *Economic and Political Weekly*, 46(21): 5560.

Wolfers, Justin. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review*, 96(5): 18021820.

Yelowitz, A. S. 1995. "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions." *The Quarterly Journal of Economics*, 110(4): 909939.

Young, Alwyn. 2018. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *The Quarterly Journal of Economics*, 134(2): 557598.

A Proofs Appendix

A.1 Optimal Reported Incomes in Baseline Model

Type $\mu = 2$ households with $y < \tau$ clearly prefer to report $y \leq \hat{y} \leq \tau$ (and are indifferent between reporting income levels in this range as only under-reporting is costly). We assume they break this indifference by reporting at y. Anyone with $y > \tau$ prefers misreporting to τ over misreporting to an income less than τ because $v' > 0$. By definition, those with $y = y^c(\mathbf{p})$ are indifferent between misreporting at the threshold τ and truthfully reporting (we break their indifference by saying they will misreport at τ), where $y^c(\mathbf{p})$ solves $y^c(\mathbf{p}) + b - v(y^c(\mathbf{p}) - \tau) = y^c(\mathbf{p})$.[53] Those with $\tau < y \leq y^c(\mathbf{p})$ prefer misreporting to τ over truthfully reporting and those with $y > y^c(\mathbf{p})$ prefer the reverse by the fact that $v' > 0$.

A.2 Proof of Lemma 1

Let us begin by calculating the WTP for the mechanical households, which are the households who report an income $\leq \tau$ under both $\mathbf{p}$ and $\mathbf{p}'$. Because households with $y > \tau$ will never report $\leq \tau$ under policy $\mathbf{p}'$ as $v' > 0$, mechanical households are those with $y \leq \tau$; they always report truthfully and receive the benefit. Hence, their utility is equal to $y + b'$ under policy $\mathbf{p}'$ and equal to $y + b$ under policy $\mathbf{p}$. Clearly, their WTP for the reform (or compensating variation) is equal to $b' - b = \Delta b$. Because only mechanical households report under τ given policy $\mathbf{p}'$, the number of such households is given by $M = G(\tau, \mathbf{p}')$.

Next, let us discuss the WTP for bunching households. Bunching households are those with type $\mu = 2$ and $y \in (\tau, y^c(\mathbf{p})]$. These individuals have utility $y + b - v(y - \tau)$ under policy $\mathbf{p}$. Bunching households with $y \in (\tau, \tau']$ will simply report truthfully after the program, getting utility $y + b'$. Those with $y \in (\tau', y^c(\mathbf{p})]$ will bunch at the new threshold, getting utility $y + b' - v(y - \tau')$.[54] Hence, we know that the WTP (or compensating variation) satisfies:

$$y + b - v(y - \tau) = y + b' - WTP - v(y - \hat{y}^*(y, 0; \mathbf{p}'))$$

where $\hat{y}^*(y, 0; \mathbf{p}')$ equals y if $y \in (\tau, \tau']$ and equals τ' if $y \in (\tau', y^c(\mathbf{p})]$. Equivalently:

$$WTP = b' - b + v(y - \tau) - v(y - \hat{y}^*(y, 0; \mathbf{p}'))$$

Given that $v(y - \tau) - v(y - \hat{y}^*(y, 0; \mathbf{p}')) > 0$ (as $v' > 0$ and $\hat{y}^*(y, 0; \mathbf{p}') > \tau$ for bunching households), we know $WTP \geq \Delta b$. Moreover, because $v(y - \hat{y}^*(y, 0; \mathbf{p}')) \geq 0$ and $y + b - v(y - \tau) \geq y$ for bunching households (or, equivalently, $v(y - \tau) \leq b$), we know that $WTP \leq b' - b + b = b'$ for bunching households. Only bunching and mechanical households report incomes $\leq \tau$ under policy $\mathbf{p}$; under policy $\mathbf{p}'$ bunching households strictly increase their reported income. Hence, the number of bunching individuals is given by $B = G(\tau; \mathbf{p}) - G(\tau; \mathbf{p}')$.

Next, let us discuss threshold households with income $y \in (\tau, \tau']$ and $\mu = 1$ (note, if $y^c(\mathbf{p}) \leq \tau'$, then type $\mu = 2$ households with $y \in (y^c(\mathbf{p}), \tau']$ are also threshold households). These households all receive utility y under policy $\mathbf{p}$ and receive utility $y + b'$ under policy $\mathbf{p}'$. Hence, their WTP is given by b'. Definitionally, threshold households are those reporting incomes in $(\tau, \tau']$ under policy $\mathbf{p}$ so that the total number of such households is given by $T = G(\tau'; \mathbf{p}) - G(\tau; \mathbf{p})$.

53. The existence of a unique $y^c(\mathbf{p})$ follows from $v' > 0$.
54. Note, if $y^c(\mathbf{p}) \leq \tau'$, i.e., the change in τ is large, all bunching households will report truthfully under $\mathbf{p}'$.

Finally, there are jumping households with type $\mu = 2$ and income $y \in (y^c(\mathbf{p}), y^c(\mathbf{p}')]$ (or $y \in (\tau', y^c(\mathbf{p}')]$ if $y^c(\mathbf{p}) \leq \tau'$). These households receive utility y under policy $\mathbf{p}$ and receive utility $y + b' - v(y - \hat{y}^*(y, 0; \mathbf{p}'))$ where $\hat{y}^*(y, 0; \mathbf{p}') \leq \tau'$. For these individuals we have:

$$y = y + b' - WTP - v(y - \hat{y}^*(y, 0; \mathbf{p}'))$$

Because $y + b' - v(y - \hat{y}^*(y, 0; \mathbf{p}')) \geq y$ by revealed preference, $WTP \geq 0$. And because $v(y - \hat{y}^*(y, 0; \mathbf{p}')) \geq 0$, $WTP \leq b'$ for jumping individuals. Jumping households are those who report incomes above τ' under policy $\mathbf{p}$ but below τ' under policy $\mathbf{p}'$. Because all households who report incomes under τ' under policy $\mathbf{p}$ also report incomes under τ' under policy $\mathbf{p}'$, the number of jumping households is given by $J = G(\tau'; \mathbf{p}') - G(\tau'; \mathbf{p})$.

Putting this all together we get:

$$\Delta b(M + B) + b'T \leq \text{Total WTP} \leq \Delta bM + b'(B + T + J)$$

A.3 Proof of Proposition 3

Proof. We start with proving the lower bound for $\frac{\frac{1}{\lambda}[\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p})]}{[b'G(\tau'; \mathbf{p}') - bG(\tau; \mathbf{p})]}$. First, note that by revealed preference, we have the following:

$$u\left(y\left(x^*\left(\theta, \mathbf{p}\right), \theta\right) + b\mathbb{1}\left(\hat{y}(x^*\left(\theta, \mathbf{p}\right), \theta) \leq \tau\right), x^*(\theta, \mathbf{p}); \theta\right) \geq u\left(y\left(x, \theta\right) + b\mathbb{1}\left(\hat{y}(x, \theta) \leq \tau\right), x; \theta\right)$$

Put simply, optimal decisions conditional on any given θ under $\mathbf{p}$, $x^*(\theta, \mathbf{p})$, yield weakly higher decisions than any other set of decisions x that one could make. This yields the following bound on welfare under policy $\mathbf{p}' = \{b', \tau'\}$, which ensues by evaluating utility under policy $\mathbf{p}'$, but holding household decisions constant at their values under policy $\mathbf{p}$ (i.e., by revealed preference):

$$\begin{aligned}
\mathcal{W}(\mathbf{p}') &= \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta, \mathbf{p}'), \theta\right) + b'\mathbb{1}\left(\hat{y}\left(x^*\left(\theta, \mathbf{p}'\right), \theta\right) \leq \tau'\right), x^*(\theta, \mathbf{p}'); \theta\right) dF(\theta) - \lambda b' G(\tau'; \mathbf{p}') \\
&\geq \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta, \mathbf{p}), \theta\right) + b'\mathbb{1}\left(\hat{y}\left(x^*\left(\theta, \mathbf{p}\right), \theta\right) \leq \tau'\right), x^*(\theta, \mathbf{p}); \theta\right) dF(\theta) - \lambda b' G(\tau'; \mathbf{p}')
\end{aligned}$$

So as to slightly reduce some cumbersome notation, let us define:

$$y^*(\theta, \mathbf{p}) \equiv y\left(x^*\left(\theta, \mathbf{p}\right), \theta\right)$$

$$\hat{y}^*(\theta, \mathbf{p}) \equiv \hat{y}(x^*\left(\theta, \mathbf{p}\right), \theta)$$

For the reform from $\mathbf{p} = \{b, \tau\}$ to $\mathbf{p}' = \{b', \tau'\}$ with $\mathbf{p}' - \mathbf{p} = \{\Delta b, \Delta\tau\}$ and $\Delta\tau > 0$:

$$\begin{aligned}
\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p}) \geq & \int_{\theta: \hat{y}^*(\theta, \mathbf{p}) \leq \tau} \phi(\theta) \left[u\left(y^*(\theta, \mathbf{p}) + b', x^*(\theta, \mathbf{p}); \theta\right) - u\left(y^*(\theta, \mathbf{p}) + b, x^*(\theta, \mathbf{p}); \theta\right)\right] dF(\theta) \\
& + \int_{\theta: \hat{y}^*(\theta, \mathbf{p}) \in (\tau, \tau']} \phi(\theta) \left[u\left(y^*(\theta, \mathbf{p}) + b', x^*\left(\theta, \mathbf{p}\right); \theta\right) - u\left(y^*(\theta, \mathbf{p}), x^*\left(\theta, \mathbf{p}\right); \theta\right)\right] dF(\theta) \\
& - \lambda[b' G(\tau'; \mathbf{p}') - bG(\tau; \mathbf{p})]
\end{aligned} \tag{17}$$

Next, define $\eta_{\{\hat{y}^*(\theta,\mathbf{p})\leq\tau\}}$ as the government's average welfare weight on the households who optimally report incomes $\hat{y}^* \leq \tau$ under policy $\mathbf{p}$. $\eta_{\{\hat{y}^*(\theta,\mathbf{p})\leq\tau\}}$ captures the average welfare gain from giving these households an extra \$1:

$$\eta_{\{\hat{y}^*(\theta,\mathbf{p})\leq\tau\}} = \frac{\int\limits_{\theta:\hat{y}^*(\theta,\mathbf{p})\leq\tau} \phi(\theta)\frac{1}{b'-b}\left[u\left(y^*(\theta,\mathbf{p})+b', x^*(\theta,\mathbf{p});\theta\right) - u\left(y^*(\theta,\mathbf{p})+b, x^*(\theta,\mathbf{p});\theta\right)\right] dF(\theta)}{G(\tau,\mathbf{p})}$$

And define $\eta_{\{\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']\}}$ as the government's average welfare weight of giving a dollar to the households who optimally report incomes $\hat{y}^* \in (\tau,\tau']$ under policy $\mathbf{p}$. $\eta_{\{\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']\}}$ captures the average welfare gain from giving these households an extra \$1:[55]

$$\eta_{\{\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']\}} = \frac{\int\limits_{\theta:\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']} \phi(\theta)\frac{1}{b'}\left[u\left(y^*(\theta,\mathbf{p})+b', x^*(\theta,\mathbf{p});\theta\right) - u\left(y^*(\theta,\mathbf{p}), x^*(\theta,\mathbf{p});\theta\right)\right] dF(\theta)}{G(\tau';\mathbf{p}) - G(\tau,\mathbf{p})}$$

Note, by the mean value theorem, $\eta_{\{\hat{y}^*(\theta,\mathbf{p})\leq\tau\}}$ and $\eta_{\{\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']\}}$ are equal to average social marginal utilities of consumption for their respective groups of households. Next, let us define the aggregate welfare weight, η_L, which equals the weighted average welfare weight of giving a dollar to all households, where the weights are determined by the lower bound of WTP for the reform:

$$\eta_L = \frac{\eta_{\{\hat{y}^*(\theta,\mathbf{p})\leq\tau\}}\Delta bG(\tau;\mathbf{p}) + \eta_{\{\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']\}}b'[G(\tau';\mathbf{p}) - G(\tau,\mathbf{p})]}{\Delta bG(\tau;\mathbf{p}) + b'[G(\tau';\mathbf{p}) - G(\tau,\mathbf{p})]}$$

Then, dividing Equation (23) through by the budgetary effect multiplied by λ, we have (recall we assume the budgetary effect is > 0):

$$\begin{aligned}
\frac{\frac{1}{\lambda}\left[\mathcal{W}\left(\mathbf{p}'\right) - \mathcal{W}(\mathbf{p})\right]}{\left[b'G\left(\tau';\mathbf{p}'\right) - bG\left(\tau;\mathbf{p}\right)\right]} &\geq \frac{\eta_{\{\hat{y}^*(\theta,\mathbf{p})\leq\tau\}}\Delta bG(\tau;\mathbf{p}) + \eta_{\{\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']\}}b'\left[G(\tau';\mathbf{p}) - G(\tau;\mathbf{p})\right]}{\lambda\left[b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p})\right]} - 1 \\
&= \frac{\eta_L}{\lambda}\frac{\Delta bG(\tau;\mathbf{p}) + b'\left[G(\tau';\mathbf{p}) - G(\tau;\mathbf{p})\right]}{\left[b'G\left(\tau';\mathbf{p}'\right) - bG\left(\tau;\mathbf{p}\right)\right]} - 1 \\
&= \omega_L MVPF_L - 1
\end{aligned}$$

where $\omega_L = \eta_L/\lambda$ and $MVPF_L$ is given by:

$$MVPF_L = \frac{\Delta bG(\tau;\mathbf{p}) + b'[G(\tau';\mathbf{p}) - G(\tau,\mathbf{p})]}{b'G\left(\tau';\mathbf{p}'\right) - bG\left(\tau;\mathbf{p}\right)} = 1 - \frac{b'\left[G_t(\tau';\mathbf{p}') - G_t(\tau';\mathbf{p})\right]}{b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p})}$$

Next, we prove the upper bound for $\frac{\frac{1}{\lambda}[\mathcal{W}(\mathbf{p}')-\mathcal{W}(\mathbf{p})]}{[b'G(\tau';\mathbf{p}')-bG(\tau;\mathbf{p})]}$. We use identical revealed preference logical to bound welfare under policy $\mathbf{p} = \{b,\tau\}$ by evaluating utility under policy $\mathbf{p}$, but holding household decisions

55. We have used $\int\limits_{\theta:\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']} dF(\theta) = G(\tau';\mathbf{p}) - G(\tau,\mathbf{p})$.

constant at their values under policy $\mathbf{p}'$:

$$\begin{aligned}\mathcal{W}(\mathbf{p}) &= \int_{\Theta} \phi(\theta) u\left(y^*(\theta,\mathbf{p}) + b\mathbb{1}\left(\hat{y}^*(\theta,\mathbf{p}) \leq \tau\right), x^*(\theta,\mathbf{p});\theta\right) dF(\theta) - \lambda b G(\tau;\mathbf{p}) \\ &\geq \int_{\Theta} \phi(\theta) u\left(y^*(\theta,\mathbf{p}') + b\mathbb{1}\left(\hat{y}^*(\theta,\mathbf{p}') \leq \tau\right), x^*(\theta,\mathbf{p}');\theta\right) dF(\theta) - \lambda b G(\tau;\mathbf{p})\end{aligned}$$

Hence, for the reform from $\mathbf{p} = \{b, \tau\}$ to $\mathbf{p}' = \{b', \tau'\}$ with $\mathbf{p}' - \mathbf{p} = \{\Delta b, \Delta\tau\}$ and $\Delta\tau > 0$:

$$\begin{aligned}\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p}) \leq & \int_{\theta:\hat{y}^*(\theta,\mathbf{p}')\leq\tau} \phi(\theta)\left[u\left(y^*(\theta,\mathbf{p}') + b', x^*(\theta,\mathbf{p}');\theta\right) - u\left(y^*(\theta,\mathbf{p}') + b, x^*(\theta,\mathbf{p}');\theta\right)\right] dF(\theta) \\ & + \int_{\theta:\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']} \phi(\theta)\left[u\left(y^*(\theta,\mathbf{p}') + b', x^*\left(\theta,\mathbf{p}'\right);\theta\right) - u\left(y^*(\theta,\mathbf{p}'), x^*\left(\theta,\mathbf{p}'\right);\theta\right)\right] dF(\theta) \\ & - \lambda[b' G(\tau';\mathbf{p}') - b G(\tau;\mathbf{p})]\end{aligned} \tag{18}$$

Next, define $\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\leq\tau\}}$ as the government's average welfare weight on the households who optimally report incomes $\hat{y}^* \leq \tau$ under policy $\mathbf{p}'$. $\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\leq\tau\}}$ captures the average welfare gain from giving these households an extra \$1:

$$\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\leq\tau\}} = \frac{\int_{\theta:\hat{y}^*(\theta,\mathbf{p}')\leq\tau} \phi(\theta)\frac{1}{b'-b}\left[u\left(y^*(\theta,\mathbf{p}') + b', x^*(\theta,\mathbf{p}');\theta\right) - u\left(y^*(\theta,\mathbf{p}') + b, x^*(\theta,\mathbf{p}');\theta\right)\right] dF(\theta)}{G(\tau,\mathbf{p}')}$$

And define $\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']\}}$ as the government's average welfare weight of giving a dollar to the households who optimally report incomes $\hat{y}^* \in (\tau, \tau']$ under policy $\mathbf{p}'$. $\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']\}}$ captures the average welfare gain from giving these households an extra \$1:[56]

$$\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']\}} = \frac{\int_{\theta:\hat{y}^*(\theta,\mathbf{p})\in(\tau,\tau']} \phi(\theta)\frac{1}{b'}\left[u\left(y^*(\theta,\mathbf{p}') + b', x^*(\theta,\mathbf{p}');\theta\right) - u\left(y^*(\theta,\mathbf{p}'), x^*(\theta,\mathbf{p}');\theta\right)\right] dF(\theta)}{G(\tau';\mathbf{p}') - G(\tau,\mathbf{p}')}$$

Again, by the mean value theorem, $\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\leq\tau\}}$ and $\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']\}}$ are equal to average social marginal utilities of consumption for their respective groups of households. Next, let us define the aggregate welfare weight, η_U, which equals the weighted average welfare weight of giving a dollar to all households, where the weights are determined by the upper bound of WTP for the reform:

$$\eta_U = \frac{\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\leq\tau\}}\Delta b G(\tau;\mathbf{p}') + \eta_{\{\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']\}} b'[G(\tau';\mathbf{p}') - G(\tau,\mathbf{p}')]}{\Delta b G(\tau;\mathbf{p}') + b'[G(\tau';\mathbf{p}') - G(\tau,\mathbf{p}')]}$$

Then, dividing Equation (24) through by the budgetary effect multiplied by λ, we have (recall we assume the

56. We have used $\int_{\theta:\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']} dF(\theta) = G(\tau';\mathbf{p}') - G(\tau,\mathbf{p}')$.

budgetary effect is > 0):

$$\begin{aligned}\frac{\frac{1}{\lambda}\left[\mathcal{W}\left(\mathbf{p}'\right)-\mathcal{W}(\mathbf{p})\right]}{\left[b'G\left(\tau';\mathbf{p}'\right)-bG\left(\tau;\mathbf{p}\right)\right]} &\leq \frac{\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\leq\tau\}}\Delta bG(\tau;\mathbf{p}')+\eta_{\{\hat{y}^*(\theta,\mathbf{p}')\in(\tau,\tau']\}}b'\left[G(\tau';\mathbf{p}')-G(\tau;\mathbf{p}')\right]}{\lambda\left[b'G(\tau';\mathbf{p}')-bG(\tau;\mathbf{p})\right]}-1\\ &=\frac{\eta_{U,t}}{\lambda}\frac{\Delta bG(\tau;\mathbf{p}')+b'\left[G(\tau';\mathbf{p}')-G(\tau;\mathbf{p}')\right]}{\left[b'G\left(\tau';\mathbf{p}'\right)-bG\left(\tau;\mathbf{p}\right)\right]}-1\\ &=\omega_U MVPF_U-1\end{aligned}$$

where $\omega_U=\eta_U/\lambda$ and $MVPF_U$ is given by:

$$MVPF_U=\frac{\Delta bG(\tau;\mathbf{p}')+b'[G(\tau';\mathbf{p}')-G(\tau,\mathbf{p}')]}{b'G\left(\tau';\mathbf{p}'\right)-bG\left(\tau;\mathbf{p}\right)}=1+\frac{b\left[G_t(\tau;\mathbf{p})-G_t(\tau;\mathbf{p}')\right]}{b'G(\tau';\mathbf{p}')-bG(\tau;\mathbf{p})}$$

□

A.4 Proposition 3 Cannot Be Improved

Proposition 4. *Without further assumptions on primitives, the bounds in Proposition 3 cannot be improved.*

Proof. To show that one cannot construct tighter bounds than Proposition 3 without additional structure on the agent's problem, we provide examples for which these bounds are attained. In particular, we create examples for which (1) bunching households have a WTP of Δb and jumping households have a WTP of 0 as well as (2) bunching and jumping households both have a WTP of b'.[57]

Example 1: Consider our baseline, misreporting model in Section 2.1 with $v(0)=0$, $v'>0$ and $v''>0$. Suppose that $\Delta b,\Delta\tau>0$. Consider a distribution for the misreporting types ($\mu=2$) with a mass point at $y=\tau$, no density on $(\tau,y^{c'})$, and another mass point at $y=y^{c'}$, where $y^{c'}$ solves $b'=v(y^{c'}-\tau')$. All bunching housholds therefore have $y=\tau$. Thus, the utility change for bunching households is equal to:

$$\tau+b'-[\tau+b-v(\tau-\tau)]=\Delta b$$

Hence, the WTP for bunching households is Δb. Because there are no individuals with incomes $\in(\tau,y^{c'})$, all jumping households have $y=y^{c'}$. Hence, all the jumping households report truthfully (and do not get the benefit) under the original policy $\mathbf{p}$. We break their indifference by assuming that they jump to the threshold τ' under the new policy $\mathbf{p}'$. The change in utility for jumpers is therefore given by:

$$y^{c'}+b'-v(y^{c'}-\tau')-y^{c'}=b'-b'=0$$

Thus, each jumping household's WTP is equal to 0. Hence, the total WTP for the reform equals $(M+B)\Delta b+Tb'=MVPF_{LB}$.

57. Note that our examples require mass points of the income distribution. But one can approximate our example scenarios arbitrarily well with smooth income distributions; hence, we can get arbitrarily close to the cases when either (1) all bunching households have a WTP of Δb and all jumping households have a WTP of 0 or (2) all bunching and jumping households both have a WTP of b'.

Example 2: Consider our baseline, misreporting model in Section 2.1 with $v(0) = 0$, $v' > 0$ and $v'' > 0$. Suppose that $\Delta b, \Delta\tau > 0$. Let y^c solve $b = v(y^c - \tau)$ and $y^{c'}$ solve $b' = v(y^{c'} - \tau')$. Finally, suppose that $\tau' - \tau$ is large enough so that $\tau' > y^c$. Consider a distribution for the misreporting types ($\mu = 2$) with no density on $[\tau, y^c)$, a mass point at $y = y^c$, and no density on $(y^c, y^{c'}]$. For the $\mu = 2$ individuals with $y = y^c$, they are indifferent between bunching at τ and reporting truthfully under policy $\mathbf{p}$. In the former case, they are bunching individuals (and there are no jumping individuals) and in the latter case they are jumping individuals (and there are no bunching individuals). Regardless, the utility gain for these households is equal to:

$$y^c + b' - [y^c + b - v(y^c - \tau)] = b'$$

Thus the total WTP for the reform will equal $M\Delta b + (B + T + J)b' = MVPF_{UB}$.

□

A.5 Incorporating Misperceptions of the Schedule

We now assume that households do not necessarily understand how the policy impacts their consumption. Households solve the following problem:

$$\begin{aligned} &\max_{x \in X} \; u\left(c, x; \theta\right) \\ &\text{s.t. } c = f(y(x,\theta), \hat{y}(x,\theta), \mathbf{p}, \theta) \end{aligned} \tag{19}$$

In words, households make decisions under the assumption that their consumption is some function of their true income y, their reported income $\hat{y}$, the $\mathbf{p}$, and state variables θ. For instance, this framework allows for households to misperceive the threshold τ or the benefit level b (e.g., $f(y(x,\theta), \hat{y}(x,\theta), \mathbf{p}, \theta) = y(x,\theta) + (b + \theta_1)\mathbb{1}(\hat{y}(x,\theta) \leq (\tau + \theta_2))$, where θ_1, θ_2 are state variables for the household). Total welfare is still given by:

$$\mathcal{W}(\mathbf{p}) = \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta, \mathbf{p}), \theta\right) + b\mathbb{1}\left(\hat{y}\left(x^*\left(\theta, \mathbf{p}\right), \theta\right) \leq \tau\right), x^*(\theta, \mathbf{p}); \theta\right) dF(\theta) - \lambda b G(\tau; \mathbf{p}) \tag{20}$$

In order for Proposition 3 to hold under the more general model with misperceptions (Problem (19), we need to make two additional assumptions. We need to assume that when the policy changes from $\mathbf{p}$ to $\mathbf{p}'$, household behavioral re-optimization improves welfare, *on average*. In other words, misperceptions of the policy reform cannot be so severe that households make themselves worse off by responding to the reform. Mathematically, we require that:

$$\begin{aligned} \mathcal{W}(\mathbf{p}') &= \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta, \mathbf{p}'), \theta\right) + b'\mathbb{1}\left(\hat{y}\left(x^*\left(\theta, \mathbf{p}'\right), \theta\right) \leq \tau'\right), x^*(\theta, \mathbf{p}'); \theta\right) dF(\theta) - \lambda b' G(\tau'; \mathbf{p}') \\ &\geq \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta, \mathbf{p}), \theta\right) + b'\mathbb{1}\left(\hat{y}\left(x^*\left(\theta, \mathbf{p}\right), \theta\right) \leq \tau'\right), x^*(\theta, \mathbf{p}); \theta\right) dF(\theta) - \lambda b' G(\tau'; \mathbf{p}') \end{aligned}$$

Note, that the previous inequality holds by revealed preference if households correctly perceive the schedule (i.e., behavioral re-optimization can only improve utility). If agents can misperceive the schedule, we simply need to assume that behavioral responses improve welfare. Correspondingly, our second assumption is that

if, hypothetically, the policy were to change from $\mathbf{p}'$ to $\mathbf{p}$, household behavioral re-optimization would also improve welfare, *on average*. Mathematically, this amounts to assuming:

$$\begin{aligned}\mathcal{W}(\mathbf{p}) &= \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta,\mathbf{p}),\theta\right) + b\mathbb{1}\left(\hat{y}\left(x^*\left(\theta,\mathbf{p}\right),\theta\right) \leq \tau\right), x^*(\theta,\mathbf{p});\theta\right) dF(\theta) - \lambda b G(\tau;\mathbf{p}) \\ &\geq \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta,\mathbf{p}'),\theta\right) + b\mathbb{1}\left(\hat{y}\left(x^*\left(\theta,\mathbf{p}'\right),\theta\right) \leq \tau\right), x^*(\theta,\mathbf{p}');\theta\right) dF(\theta) - \lambda b G(\tau;\mathbf{p})\end{aligned}$$

If we are willing to make these two assumptions, the rest of the proof to Proposition 3 goes through, so that we can bound the welfare impact of changing notches if individuals misperceive the schedule. Hence, we can state:

Proposition 5. *Suppose households solve Problem (19), welfare is given by Equation (20) and $\tau' > \tau$. If we assume:*

$$\mathcal{W}(\mathbf{p}') \geq \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta,\mathbf{p}),\theta\right) + b'\mathbb{1}\left(\hat{y}\left(x^*\left(\theta,\mathbf{p}\right),\theta\right) \leq \tau'\right), x^*(\theta,\mathbf{p});\theta\right) dF(\theta) - \lambda b' G(\tau';\mathbf{p}')$$

and

$$\mathcal{W}(\mathbf{p}) \geq \int_{\Theta} \phi(\theta) u\left(y\left(x^*(\theta,\mathbf{p}'),\theta\right) + b\mathbb{1}\left(\hat{y}\left(x^*\left(\theta,\mathbf{p}'\right),\theta\right) \leq \tau\right), x^*(\theta,\mathbf{p}');\theta\right) dF(\theta) - \lambda b G(\tau;\mathbf{p})$$

Then as long as $b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p}) > 0$ have:

$$\omega_L MVPF_L - 1 \leq \frac{\frac{1}{\lambda}\left[\mathcal{W}(\mathbf{p}') - \mathcal{W}(\mathbf{p})\right]}{b'G(\tau';\mathbf{p}') - bG(\tau;\mathbf{p})} \leq \omega_U MVPF_U - 1$$

where $MVPF_L$ is given by Equation (7), $MVPF_U$ is given by Equation (8), and ω_L (ω_U) captures the weighted average money-metric welfare gain from giving a dollar to mechanical, bunching, threshold, and jumping households, where the weights are determined by the lower bound (upper bound) of each group's WTP.

A.6 Discounted Welfare Impact of Reform

Suppose households have several decisions variables at time t denoted by the vector x_t (within a potentially limited choice set X_t). Household decisions are made conditional on state variables denoted by the vector $\theta_t \in \Theta_t$ and the policy $\mathbf{p}$. Households get the benefit b if their reported income $\hat{y}_t$, which can be a function of decision variables x_t of included as a decision variable, is below τ. Household income, denoted y_t, is also potentially a function of decisions x_t.[58] Households in period t solve:

$$\begin{aligned} V(\theta_t) = &\max_{x_t \in X_t} \; u\left(c_t, x_t; \theta_t\right) + \beta \mathbb{E}_{\theta_{t+1}|\theta_t, x_t}[V(\theta_{t+1})] \\ &\text{s.t. } c_t = y_t(x_t,\theta_t) + b\mathbb{1}(\hat{y}_t(x_t,\theta_t) \leq \tau) \end{aligned} \tag{21}$$

58. For example, in a dynamic misreporting model, $x_t = \hat{y}_t$ and θ_t could include current income y_t, aversion to misreporting μ, prior reported income $\hat{y}_{t-1}$, and a parameter governing expected future income growth. Households may also make savings decisions if assets are a state variable in θ_t, current savings is included in x_t, and c_t represents post-transfer income.

where c_t denotes consumption in period t, β is a discount factor, and $\mathbb{E}_{\theta_{t+1}|\theta_t,x_t}[V(\theta_{t+1})]$ represents the expected value of starting period $t+1$ with state variables θ_{t+1}, noting that the expectation over θ_{t+1} may be impacted by current state variables θ_t and current decisions x_t. Equivalently, we can write out individual utility from the perspective of time period 0 as:

$$\begin{aligned}&\sum_{t=0}^{T}\beta^t\mathbb{E}_{\theta_1|\theta_0,x_0}[\dots\mathbb{E}_{\theta_t|\theta_{t-1},x_{t-1}}[u\left(c_t,x_t;\theta_t\right)]\dots]\\&=\sum_{t=0}^{T}\beta^t\mathbb{E}_{\theta_t|\theta_0,\{x_t\}}[u\left(c_t,x_t;\theta_t\right)]\end{aligned}$$

where $\mathbb{E}_{\theta_t|\theta_0,\{x_t\}}$ represents the expectation over θ_t from the perspective of time period 0 (taking into account the impact of all conditional decisions $\{x_t\}$ between time 0 and time t on the underlying expectations). The equality above follows from the law of iterated expectations.

So as to slightly reduce some cumbersome notation, let us define:

$$y_t^*(\theta_t,\mathbf{p})\equiv y_t\left(x_t^*\left(\theta_t,\mathbf{p}\right),\theta_t\right)$$

$$\hat{y}_t^*(\theta_t,\mathbf{p})\equiv \hat{y}_t(x_t^*\left(\theta_t,\mathbf{p}\right),\theta_t)$$

Using this notation, we again assume total discounted welfare is given by a weighted discounted sum of utilities, with welfare weights given by $\phi(\theta_0)$, less the total discounted budgetary cost of the policy multiplied by a shadow value of public funds λ:

$$\begin{aligned}&\sum_{t=0}^{T}\beta^t\mathcal{W}_t(\mathbf{p})=\\&\int_{\Theta_0}\phi(\theta_0)\left[\sum_{t=0}^{T}\beta^t\mathbb{E}_{\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p})\}}[u\left(y_t^*(\theta_t,\mathbf{p})+b\mathbb{1}\left(\hat{y}_t^*(\theta_t,\mathbf{p})\leq\tau\right),x_t^*(\theta_t,\mathbf{p});\theta_t)]dF(\theta_0)\right]-\lambda\sum_{t=0}^{T}\beta^t bG_t(\tau;\mathbf{p})\end{aligned}\tag{22}$$

where β represents the governments discount rate, λ represents the shadow value of public funds in time $t=0$ (so that the shadow value of public funds in future periods equals $\beta^t\lambda$), and:

$$G_t(\tau;\mathbf{p})=\int_{\Theta_0}\mathbb{E}_{\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p})\}}[\mathbb{1}(\hat{y}_t^*(\theta_t,\mathbf{p})\leq\tau)]dF(\theta_0)$$

represents the *expected* number of households receiving the benefit under policy $\mathbf{p}$ in period t. More generally, we define:

$$G_t(z;\mathbf{p})=\int_{\Theta_0}\mathbb{E}_{\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p})\}}[\mathbb{1}(\hat{y}_t^*(\theta_t,\mathbf{p})\leq z)]dF(\theta_0)$$

as the expected number of households with a reported income below z under policy $\mathbf{p}$ in time period t.

This setup allows us to bound the welfare impacts of a policy reform over many time periods:

Proposition 6. *Suppose households solve Problem (21), total welfare is given by Equation (22), and* $\tau'>\tau$.

Defining:

$$MVPF_{L,T} \equiv 1 - \frac{\sum_{t=0}^{T} \beta^t b' \left[G_t(\tau'; \mathbf{p}') - G_t(\tau'; \mathbf{p})\right]}{\sum_t \beta^t [b' G_t(\tau'; \mathbf{p}') - b G_t(\tau; \mathbf{p})]}$$

$$MVPF_{U,T} \equiv 1 + \frac{\sum_{t=0}^{T} \beta^t b \left[G_t(\tau; \mathbf{p}) - G_t(\tau; \mathbf{p}')\right]}{\sum_t \beta^t [b' G_t(\tau'; \mathbf{p}') - b G_t(\tau; \mathbf{p})]}$$

Then as long as $\sum_t \beta^t [b' G_t(\tau'; \mathbf{p}') - b G_t(\tau; \mathbf{p})] > 0$ *have:*

$$\omega_{L,T} MVPF_{L,T} - 1 \leq \frac{\frac{1}{\lambda} \sum_{t=0}^{T} \beta^t [\mathcal{W}_t(\mathbf{p}') - \mathcal{W}_t(\mathbf{p})]}{\sum_{t=0}^{T} \beta^t \left[b' G_t(\tau'; \mathbf{p}') - b G_t(\tau; \mathbf{p})\right]} \leq \omega_{U,T} MVPF_{U,T} - 1$$

where $\omega_{L,T}$ *(*$\omega_{U,T}$*) captures the discounted weighted average money-metric welfare gain from giving a dollar to mechanical, bunching, threshold, and jumping households, where the weights are determined by the lower bound (upper bound) of each group's discounted WTP.*

Proof. We start with proving the lower bound for $\frac{\frac{1}{\lambda} \sum_{t=0}^{T} \beta^t [\mathcal{W}_t(\mathbf{p}') - \mathcal{W}_t(\mathbf{p})]}{\sum_{t=0}^{T} \beta^t [b' G_t(\tau'; \mathbf{p}') - b G_t(\tau; \mathbf{p})]}$. First, note that by revealed preference, we have the following:

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \mathbb{E}_{\theta_t|\theta_0, \{x_t^*(\theta_t, \mathbf{p})\}} [u\left(y_t^*(\theta_t, \mathbf{p}) + b \mathbb{1}\left(\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau\right), x_t^*(\theta_t, \mathbf{p}); \theta_t\right)] \\
&\equiv \sum_{t=0}^{T} \beta^t \mathbb{E}_{\theta_t|\theta_0, \{x_t^*(\theta_t, \mathbf{p})\}} [u\left(y_t\left(x_t^*\left(\theta_t, \mathbf{p}\right), \theta_t\right) + b \mathbb{1}\left(\hat{y}_t(x_t^*\left(\theta_t, \mathbf{p}\right), \theta_t) \leq \tau\right), x_t^*(\theta_t, \mathbf{p}); \theta_t\right)] \\
&\geq \sum_{t=0}^{T} \beta^t \mathbb{E}_{\theta_t|\theta_0, \{x_t\}} [u\left(y_t\left(x_t, \theta_t\right) + b \mathbb{1}\left(\hat{y}_t(x_t, \theta_t) \leq \tau\right), x_t; \theta_t\right]
\end{aligned}$$

Put simply, optimal decisions conditional on any given θ_t under $\mathbf{p}$, $x^*(\theta_t, \mathbf{p})$, yield weakly higher decisions than any other set of decisions x_t that one could make. This yields the following bound on welfare under policy $\mathbf{p}' = \{b', \tau'\}$, which ensues by evaluating utility under policy $\mathbf{p}'$, but holding household decisions constant at their values under policy $\mathbf{p}$ (i.e., by revealed preference):

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \mathcal{W}_t(\mathbf{p}') = \\
&\int_{\Theta_0} \phi(\theta_0) \left[\sum_{t=0}^{T} \beta^t \mathbb{E}_{\theta_t|\theta_0, \{x_t^*(\theta_t, \mathbf{p}')\}} [u\left(y_t^*(\theta_t, \mathbf{p}) + b' \mathbb{1}\left(\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau'\right), x_t^*(\theta_t, \mathbf{p}'); \theta_t\right)] dF(\theta_0)\right] - \lambda \sum_{t=0}^{T} \beta^t b G_t(\tau'; \mathbf{p}') \\
&\geq \int_{\Theta_0} \phi(\theta_0) \left[\sum_{t=0}^{T} \beta^t \mathbb{E}_{\theta_t|\theta_0, \{x_t^*(\theta_t, \mathbf{p})\}} [u\left(y_t^*(\theta_t, \mathbf{p}) + b' \mathbb{1}\left(\hat{y}_t(x_t^*\left(\theta_t, \mathbf{p}\right), \theta_t) \leq \tau'\right), x_t^*(\theta_t, \mathbf{p}); \theta_t\right)] dF(\theta_0)\right] - \lambda \sum_{t=0}^{T} \beta^t b G_t(\tau'; \mathbf{p}')
\end{aligned}$$

Hence, for the reform from $\mathbf{p} = \{b, \tau\}$ to $\mathbf{p}' = \{b', \tau'\}$ with $\mathbf{p}' - \mathbf{p} = \{\Delta b, \Delta\tau\}$ and $\Delta\tau > 0$:

$$\begin{aligned}
&\sum_{t=0}^{T} \beta^t \left[\mathcal{W}_t\left(\mathbf{p}'\right) - \mathcal{W}_t(\mathbf{p})\right] \geq \\
&\sum_{t=0}^{T} \beta^t \int_{\Theta_0} \int_{\theta_t : \hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau} \phi(\theta_0) \left[u\left(y_t^*(\theta_t, \mathbf{p}) + b', x_t^*(\theta_t, \mathbf{p}); \theta_t\right) - u\left(y_t^*(\theta_t, \mathbf{p}) + b, x_t^*(\theta_t, \mathbf{p}); \theta_t\right)\right] dF(\theta_t | \theta_0, \{x_t^*(\theta_t, \mathbf{p})\}) dF(\theta_0) \\
&+ \sum_{t=0}^{T} \beta^t \int_{\Theta_0} \int_{\theta_t : \hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']} \phi(\theta_0) \left[u\left(y_t^*(\theta_t, \mathbf{p}) + b', x_t^*(\theta_t, \mathbf{p}); \theta_t\right) - u\left(y_t^*(\theta_t, \mathbf{p}), x_t^*(\theta_t, \mathbf{p}); \theta_t\right)\right] dF(\theta_t | \theta_0, \{x_t^*(\theta_t, \mathbf{p})\}) dF(\theta_0) \\
&- \lambda \sum_{t=0}^{T} \beta^t [b' G_t(\tau'; \mathbf{p}') - b G_t(\tau; \mathbf{p})]
\end{aligned} \tag{23}$$

Next, define $\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau\}}$ as the government's average expected welfare weight on the households who optimally report incomes $\leq \tau$ under policy $\mathbf{p}$ at time t. $\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau\}}$ captures the average expected welfare gain from giving these households an extra \$1:

$$\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau\}} = \frac{\int_{\Theta_0} \int_{\theta_t : \hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau} \phi(\theta_0) \frac{1}{b' - b} \left[u\left(y_t^*(\theta_t, \mathbf{p}) + b', x_t^*(\theta_t, \mathbf{p}); \theta_t\right) - u\left(y_t^*(\theta_t, \mathbf{p}) + b, x_t^*(\theta_t, \mathbf{p}); \theta_t\right)\right] dF(\theta_t | \theta_0, \{x_t^*(\theta_t, \mathbf{p})\}) dF(\theta_0)}{G_t(\tau, \mathbf{p})}$$

And define $\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']\}}$ as the government's average expected welfare weight of giving a dollar to the households who optimally report incomes $\in (\tau, \tau']$ under policy $\mathbf{p}$ at time t. $\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']\}}$ captures the average expected welfare gain from giving these households an extra \$1:[59]

$$\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']\}} = \frac{\int_{\Theta_0} \int_{\theta_t : \hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']} \phi(\theta_0) \frac{1}{b'} \left[u\left(y_t^*(\theta_t, \mathbf{p}) + b', x_t^*(\mathbf{p}); \theta_t\right) - u\left(y_t^*(\theta_t, \mathbf{p}), x_t^*(\theta_t, \mathbf{p}); \theta_t\right)\right] dF(\theta_t | \theta_0, \{x_t^*(\theta_t, \mathbf{p})\}) dF(\theta_0)}{G_t(\tau'; \mathbf{p}) - G_t(\tau, \mathbf{p})}$$

Note, by the mean value theorem, $\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau\}}$ and $\eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']\}}$ are equal to average expected social marginal utilities of consumption for their respective groups of households. Next, let us define an aggregate discounted welfare weight, $\eta_{L,t}$, which equals the weighted average discounted expected welfare weight of giving a dollar to all households, where the weights are determined by the lower (discounted) bound of expected WTP for the reform:

$$\eta_{L,T} = \frac{\sum_{t=0}^{T} \beta^t \eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \leq \tau\}} \Delta b G_t(\tau; \mathbf{p}) + \sum_{t=0}^{T} \beta^t \eta_{\{\hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']\}} b' [G_t(\tau'; \mathbf{p}) - G_t(\tau, \mathbf{p})]}{\sum_{t=0}^{T} \beta^t [\Delta b G_t(\tau; \mathbf{p}) + b' [G_t(\tau'; \mathbf{p}) - G_t(\tau, \mathbf{p})]]}$$

Then, dividing Equation (23) through by the budgetary effect multiplied by λ, we have (recall we assume the

59. We have used $\int_{\Theta_0} \int_{\theta_t : \hat{y}_t^*(\theta_t, \mathbf{p}) \in (\tau, \tau']} dF(\theta_t | \theta_0, \{x_t^*(\theta_t, \mathbf{p})\}) dF(\theta_0) = G_t(\tau'; \mathbf{p}) - G_t(\tau, \mathbf{p})$.

budgetary effect is > 0):

$$\frac{\sum_{t=0}^{T}\beta^t\frac{1}{\lambda}\left[\mathcal{W}_t\left(\mathbf{p}'\right)-\mathcal{W}_t(\mathbf{p})\right]}{\sum_{t=0}^{T}\beta^t\left[b'G_t\left(\tau';\mathbf{p}'\right)-bG_t\left(\tau;\mathbf{p}\right)\right]}\geq\frac{\sum_{t=0}^{T}\beta^t\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p})\leq\tau\}}\Delta bG_t(\tau;\mathbf{p})+\sum_{t=0}^{T}\beta^t\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p})\in(\tau,\tau']\}}b'\left[G_t(\tau';\mathbf{p})-G_t(\tau;\mathbf{p})\right]}{\sum_{t=0}^{T}\beta^t\lambda\left[b'G_t(\tau';\mathbf{p}')-bG_t(\tau;\mathbf{p})\right]}-1$$

$$=\frac{\eta_{L,T}}{\lambda}\frac{\sum_{t=0}^{T}\beta^t[\Delta bG_t(\tau;\mathbf{p})+b'\left[G_t(\tau';\mathbf{p})-G_t(\tau;\mathbf{p})\right]]}{\sum_{t=0}^{T}\beta^t\left[b'G_t\left(\tau';\mathbf{p}'\right)-bG_t\left(\tau;\mathbf{p}\right)\right]}-1$$

$$=\omega_{L,T}MVPF_{L,T}-1$$

where $\omega_{L,T}=\eta_{L,T}/\lambda$ and $MVPF_{L,T}$ is given by:

$$MVPF_{L,T}=\frac{\sum_{t=0}^{T}\beta^t[\Delta bG_t(\tau;\mathbf{p})+b'[G_t(\tau';\mathbf{p})-G_t(\tau,\mathbf{p})]]}{\sum_{t=0}^{T}\beta^t[b'G_t\left(\tau';\mathbf{p}'\right)-bG_t\left(\tau;\mathbf{p}\right)]}=1-\frac{\sum_{t=0}^{T}\beta^tb'\left[G_t(\tau';\mathbf{p}')-G_t(\tau';\mathbf{p})\right]}{\sum_{t=0}^{T}\beta^t[b'G_t(\tau';\mathbf{p}')-bG_t(\tau;\mathbf{p})]}$$

Next, we prove the upper bound for $\frac{\frac{1}{\lambda}\sum_{t=0}^{T}\beta^t[\mathcal{W}_t(\mathbf{p}')-\mathcal{W}_t(\mathbf{p})]}{\sum_{t=0}^{T}\beta^t[b'G_t(\tau';\mathbf{p}')-bG_t(\tau;\mathbf{p})]}$. We use identical revealed preference logical to bound welfare under policy $\mathbf{p}=\{b,\tau\}$ by evaluating utility under policy $\mathbf{p}$, but holding household decisions constant at their values under policy $\mathbf{p}'$:

$$\sum_{t=0}^{T}\beta^t\mathcal{W}_t(\mathbf{p})=\int_{\Theta_0}\phi(\theta_0)\left[\sum_{t=0}^{T}\beta^t\mathbb{E}_{\theta_t|\theta_0,\{x_t^*(\mathbf{p})\}}[u\left(y_t^*(\theta_t,\mathbf{p})+b\mathbb{1}\left(\hat{y}_t^*(\theta_t,\mathbf{p})\leq\tau\right),x_t^*(\theta_t,\mathbf{p});\theta_t)]dF(\theta_0)\right]-\lambda\sum_{t=0}^{T}\beta^tbG_t(\tau;\mathbf{p})$$

$$\geq\int_{\Theta_0}\phi(\theta_0)\left[\sum_{t=0}^{T}\beta^t\mathbb{E}_{\theta_t|\theta_0,\{x_t^*(\mathbf{p}')\}}[u\left(y_t^*(\theta_t,\mathbf{p}')+b\mathbb{1}\left(\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\right),x_t^*(\theta_t,\mathbf{p}');\theta_t)]dF(\theta_0)\right]-\lambda\sum_{t=0}^{T}\beta^tbG_t(\tau;\mathbf{p})$$

Hence, for the reform from $\mathbf{p}=\{b,\tau\}$ to $\mathbf{p}'=\{b',\tau'\}$ with $\mathbf{p}'-\mathbf{p}=\{\Delta b,\Delta\tau\}$ and $\Delta\tau>0$:

$$\begin{aligned}
&\sum_{t=0}^{T}\beta^t\left[\mathcal{W}_t\left(\mathbf{p}'\right)-\mathcal{W}_t(\mathbf{p})\right]\leq\\
&\sum_{t=0}^{T}\beta^t\int_{\Theta_0}\int_{\theta_t:\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau}\phi(\theta_0)\left[u\left(y_t^*(\theta_t,\mathbf{p}')+b',x_t^*(\theta_t,\mathbf{p}');\theta_t\right)-u\left(y_t^*(\theta_t,\mathbf{p}')+b,x_t^*(\theta_t,\mathbf{p}');\theta_t\right)\right]dF(\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p}')\})dF(\theta_0)\\
&+\sum_{t=0}^{T}\beta^t\int_{\Theta_0}\int_{\theta_t:\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']}\phi(\theta_0)\left[u\left(y_t^*(\theta_t,\mathbf{p}')+b',x_t^*(\theta_t,\mathbf{p}');\theta_t\right)-u\left(y_t^*(\theta_t,\mathbf{p}'),x_t^*(\theta_t,\mathbf{p}');\theta_t\right)\right]dF(\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p}')\})dF(\theta_0)\\
&-\lambda\sum_{t=0}^{T}\beta^t[b'G_t(\tau';\mathbf{p}')-bG_t(\tau;\mathbf{p})]
\end{aligned}\tag{24}$$

Next, define $\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\}}$ as the government's average expected welfare weight on the households who optimally report incomes $\leq\tau$ under policy $\mathbf{p}'$ at time t. $\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\}}$ captures the average expected welfare gain

from giving these households an extra \$1:

$$\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\}} = \frac{\int_{\Theta_0}\int_{\theta_t:\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau} \phi(\theta_0)\frac{1}{b'-b}\left[u\left(y_t^*(\theta_t,\mathbf{p}')+b',x_t^*(\theta_t,\mathbf{p}');\theta_t\right)-u\left(y_t^*(\theta_t,\mathbf{p}')+b,x_t^*(\theta_t,\mathbf{p}');\theta_t\right)\right]dF(\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p}')\})dF(\theta_0)}{G_t(\tau,\mathbf{p}')}$$

And define $\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']\}}$ as the government's average expected welfare weight of giving a dollar to the households who optimally report incomes $\in(\tau,\tau']$ under policy $\mathbf{p}'$ at time t. $\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']\}}$ captures the average expected welfare gain from giving these households an extra \$1:[60]

$$\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']\}} = \frac{\int_{\Theta_0}\int_{\theta_t:\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']} \phi(\theta_0)\frac{1}{b'}\left[u\left(y_t^*(\theta_t,\mathbf{p}')+b',x_t^*(\theta_t,\mathbf{p}');\theta_t\right)-u\left(y_t^*(\theta_t,\mathbf{p}'),x_t^*(\theta_t,\mathbf{p}');\theta_t\right)\right]dF(\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p}')\})dF(\theta_0)}{G_t(\tau';\mathbf{p}')-G_t(\tau,\mathbf{p}')}$$

Again, by the mean value theorem, $\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\}}$ and $\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']\}}$ are equal to average expected social marginal utilities of consumption for their respective groups of households. Next, let us define an aggregate discounted welfare weight, $\eta_{U,t}$, which equals the weighted average discounted expected welfare weight of giving a dollar to all households, where the weights are determined by the upper (discounted) bound of expected WTP for the reform:

$$\eta_{U,T} = \frac{\sum_{t=0}^{T}\beta^t\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\}}\Delta bG_t(\tau;\mathbf{p}') + \sum_{t=0}^{T}\beta^t\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']\}}b'[G_t(\tau';\mathbf{p}')-G_t(\tau,\mathbf{p}')]}{\sum_{t=0}^{T}\beta^t[\Delta bG_t(\tau;\mathbf{p}')+b'[G_t(\tau';\mathbf{p}')-G_t(\tau,\mathbf{p}')]]}$$

Then, dividing Equation (24) through by the budgetary effect multiplied by λ, we have (recall we assume the budgetary effect is >0):

$$\begin{aligned}
\frac{\frac{1}{\lambda}\sum_{t=0}^{T}\beta^t\left[\mathcal{W}_t\left(\mathbf{p}'\right)-\mathcal{W}_t(\mathbf{p})\right]}{\sum_{t=0}^{T}\beta^t\left[b'G_t\left(\tau';\mathbf{p}'\right)-bG_t\left(\tau;\mathbf{p}\right)\right]} &\leq \frac{\sum_{t=0}^{T}\beta^t\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\leq\tau\}}\Delta bG_t(\tau;\mathbf{p}') + \sum_{t=0}^{T}\beta^t\eta_{\{\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']\}}b'\left[G_t(\tau';\mathbf{p}')-G_t(\tau;\mathbf{p}')\right]}{\lambda\sum_{t=0}^{T}\beta^t\left[b'G_t(\tau';\mathbf{p}')-bG_t(\tau;\mathbf{p})\right]} - 1 \\
&= \frac{\eta_{U,T}}{\lambda}\frac{\sum_{t=0}^{T}\beta^t[\Delta bG_t(\tau;\mathbf{p}')+b'\left[G_t(\tau';\mathbf{p}')-G_t(\tau;\mathbf{p}')\right]]}{\sum_{t=0}^{T}\beta^t\left[b'G_t\left(\tau';\mathbf{p}'\right)-bG_t\left(\tau;\mathbf{p}\right)\right]} - 1 \\
&= \omega_{U,T}MVPF_{U,T} - 1
\end{aligned}$$

where $\omega_{U,T}=\eta_{U,T}/\lambda$ and $MVPF_{U,T}$ is given by:

$$MVPF_{U,T} = \frac{\sum_{t=0}^{T}\beta^t[\Delta bG_t(\tau;\mathbf{p}')+b'[G_t(\tau';\mathbf{p}')-G_t(\tau,\mathbf{p}')]]}{\sum_{t=0}^{T}\beta^t[b'G_t\left(\tau';\mathbf{p}'\right)-bG_t\left(\tau;\mathbf{p}\right)]} = 1+\frac{\sum_{t=0}^{T}\beta^t b\left[G_t(\tau;\mathbf{p})-G_t(\tau;\mathbf{p}')\right]}{\sum_{t=0}^{T}\beta^t[b'G_t(\tau';\mathbf{p}')-bG_t(\tau;\mathbf{p})]}$$

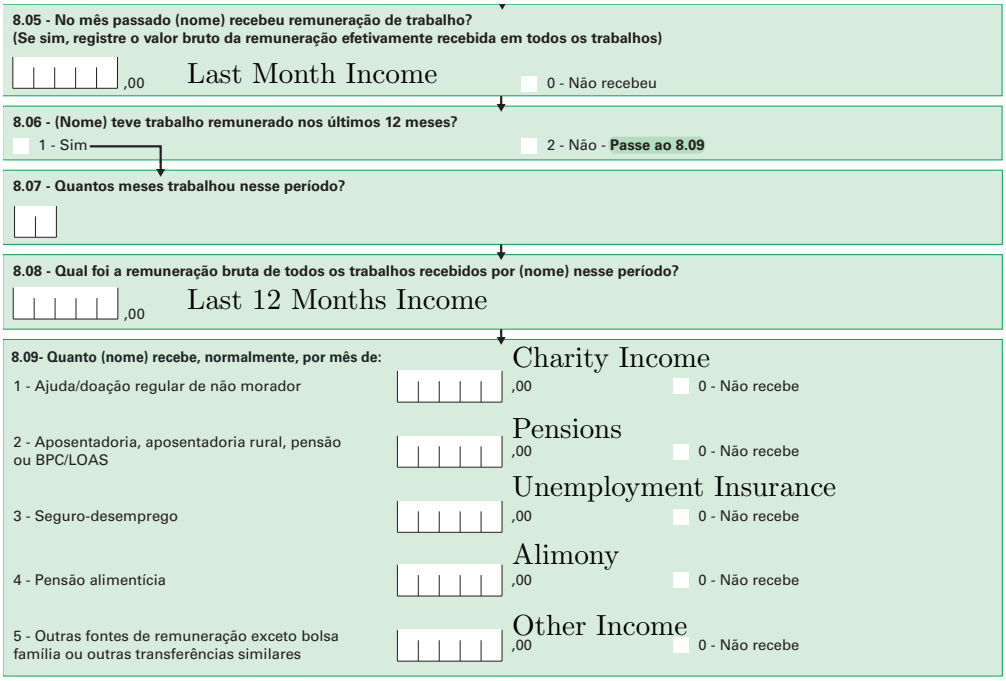
60. We have again used $\int_{\Theta_0}\int_{\theta_t:\hat{y}_t^*(\theta_t,\mathbf{p}')\in(\tau,\tau']} dF(\theta_t|\theta_0,\{x_t^*(\theta_t,\mathbf{p})\})dF(\theta_0) = G_t(\tau';\mathbf{p}')-G_t(\tau,\mathbf{p}')$.

□

B Bolsa Família Program Appendix

B.1 Bolsa Família Questionnaire

Figure 9 shows the entries on the questionnaire used to calculate each individual's total monthly income (household per-capita income will be calculated via summing total individual monthly incomes across all members of a household divided by the number of members in the household). Questions 8.05-8.08 relate to determining last month's labor income and the labor income over the last 12 months. The computer will then calculate an individual's minimum monthly labor income via taking the minimum between the individual's labor income last month and the individual's average monthly labor income over the last 12 months. Question 8.09 relates to determining the average monthly income from five other income sources: charity, pensions, unemployment insurance, alimony, and other. An individual's total monthly income is then equal to their monthly income from these five sources plus their minimum monthly labor income.



Note: The figure depicts the income categories reported by applicants for each member of the household. Each category has been translated into English in the figure. This is a print out of the screen seen by interviewers on their computers when filling in applicants' information.

Figure 9: Income Questionnaire

B.2 Other Social Security Programs Based on the Cadastro Único

This appendix describes other programs that set their eligibility based on information from the Cadastro Único database.

Benefício de Pretação Continuada (BPC): This benefit targets the elderly (above 65 years of age) and disabled. It gives a minimum wage to all households with per-capita income up to a quarter of the minimum

wage. Table 4 reports the minimum wage and BPC threshold across all years of the analysis. The Brazilian Social Security System administers its own exam to define eligibility for this program.

Table 4: Minimum Wage and BPC Eligibility Thresholds

Year	Minimum Wage	BPC Threshold
2011	545.00	136.25
2012	622.00	155.50
2013	678.00	169.50
2014	724.00	181.00
2015	788.00	197.00
2016	880.00	220.00

Carteira do Idoso: This "Elderly Card" guarantees to all individuals 60 years of age or older and with income up to two times the minimum wage at least a 50% discount on any interstate trip by road, rail, or waterway.

Créditos Instalação do Programa Nacional de Reforma Agrária: Households with per-capita income up to three times the minimum wage and that are living in camping grounds get points in a system that selects beneficiaries to be settled through the Brazilian land reform.

Facultativo de Baixa Renda: This is an option to contribute to social security at a lower rate (5% of the minimum wage). The individual cannot have any income and household income must be below two times the minimum wage.

Identidade Jovem (ID Jovem): Discounts for cultural events and trips by road, rail, or waterway for individuals between 15 and 29 years of age living in a household with up to two times the minimum wage.

Isenção de taxas de inscrição em concursos públicos: Since 2008, households with per-capita income up to half of the minimum wage or total income of up to three times the minimum wage are exempt from public tender registration payment.

Política Nacional Assistência Técnica Rural — PNATER Brasil Sem Miséria: Technical assistance for households working on activities for their own consumption in rural areas.

Programa Água para Todos — Programa Nacional de Universalização do Acesso e Uso da Água: Since July 2011, the government has installed cisterns to ensure access to clean water for all Brazilians, with priority going to those who satisfy the criteria for BF program.

Bolsa Estiagem: This is a benefit of at least R$80 per month to households with total income up to two times the minimum wage that live in areas hit by natural disasters.

Programa Bolsa Verde — Programa de Apoio à Conservação Ambiental: Since October 2011, this program transfers R$300 every 3 months to households in extreme poverty (first threshold of BF) and that follow the requirements for using natural resources.

Programa Cisternas: This program aims to provide cisterns to low-income families registered in the Cadastro Único.

Programa de Erradicação do Trabalho Infantil: This program transfers benefits similar to the BF (R$25 and R$40 per child per month in municipalities with less and more than 250,000 inhabitants, respectively) to households whose incomes are above the BF threshold with working children (up to 16 years of age)

conditional on these children attending school 85% of the time instead of working.

Programa de Fomento às Atividades Produtivas Rurais: Since 2012, the government has made a one-time transfer of around R$2,400 to families that are eligible for the BF program and work on agricultural activities or belong to native or traditional communities.

Programa Minha Casa Minha Vida: Households with total monthly income up to R$1,416.67 have access to a subsidized credit line to purchase a house.

Programa Nacional de Crédito Fundiário: Households with total monthly income up to R$2,500 have access to a subsidized credit line to purchase land for production.

Serviços Socioassistenciais: MDS offers social services to poor individuals who have suffered any type of violence or neglect.

Sistema de Seleção Unificada — Sisu/Lei de Cotas: Since 2016, all federal universities in the country reserve some seats for students coming from families with per-capita income up to 1.5 times the minimum wage.

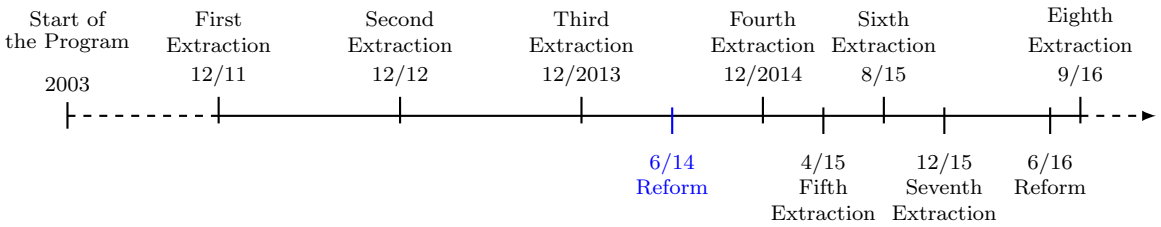
Tarifa Social de Energia Elétrica: Households with monthly per-capita income of up to a half the minimum wage have access to a discounted electricity price.

Tefone Popular — Acesso Individual Classe Especial: The government offers a landline with lower prices for individuals registered in the Cadastro Único database.

Distribuição de Conversores de TV Digital: Since the September of 2015, MDS has offered digital converters to beneficiaries of the program, which helps them transition from open TV to the new system.

B.3 Bolsa Família Program and Data Extraction Timeline

Figure 10 presents the data extraction timeline relative to when the BF program started and the June 2014 reform. As can be seen, the program started in 2003, the extractions we have span the months between December 2011 to September 2016, and the reform we study occurred in June 2014. Note, there was also another reform in June 2016.



Note: The figure describes the timeline of the program and the data extractions. BF started in 2003, and the reform we study occurred in June 2014. The final dataset is constructed from 8 extractions from December 2011 until September of 2016. Each extraction contains the the most recent information on each household as of the extraction date. Note, there was another reform in June 2016.

FIGURE 10: TIMELINE

B.4 Bolsa Família Schedule for Households with Children

At the beginning of our dataset (December 2011), the BF program has two eligibility thresholds in the per-capita monthly income distribution for households with children: the extreme-poverty line (R$70) and

the poverty line (R$140). Households with per-capita income below the extreme-poverty line are eligible for the constant basic benefit (R$70 per-month), a variable benefit proportional to the number of family members between 0 and 15 years old (R$32 per-child, per-month), and a teenager benefit proportional to the number of members aged 16 or 17 years old (R$38 per-teenager, per-month). Households with per-capita income between the extreme-poverty and poverty thresholds are only eligible for the variable and teenager benefits. Households with per-capita income above the second threshold are not eligible for any BF cash transfers. Moreover, the total variable benefit was capped at R$160 (5 children per household) and the total teenager benefit was capped at R$76 (two teenagers per household).[61]

The reform this paper studies, which occurred in June 2014, increased the extreme poverty threshold from R$70 to R$77 and the poverty threshold from R$140 to R$154. The basic benefit was raised from R$70 to R$77, the benefit per child from R$32 to R$35, and the benefit per teenager from R$38 to R$42. Note that the thresholds are based on per-capita income but the benefits are denominated in raw amounts. This reform was announced on national television by the president in April 2014.

Table 5 summarizes these aspects of the schedule before (first column) and after (second column) the reform for households with children.

Table 5: Summary of Bolsa Família Schedule for Families with Children Before and After June 2014 Reform

	Before	After
Extreme-Poverty Threshold	70	77
Poverty Threshold	140	154
Basic Benefit (for those in extreme-poverty)	70	77
Variable Benefit Per Child 15 or Younger (max 5) (for those in poverty)	32	35
Teenager Benefit Per Teen 16-18 (max 2) (for those in poverty)	38	42

Note: The first two rows correspond to the the extreme-poverty and poverty thresholds, respectively. These are measured in monthly, per-capita income. I.e., before the reform a household is below the extreme-poverty threshold if their monthly, per-capita income is below R$70. The third, fourth, and fifth rows display the benefits given to households; these are denoted in monthly amounts. I.e., before the reform a household below the extreme-poverty threshold receives R$70 per-month in the basic benefit.

Between December 2011 and February 2013 there were three other reforms to the BF program, which successively instituted a guaranteed minimum income of R$70 per-capita (along with an associated negative income tax) for several groups of households. This guaranteed minimum income was instituted in June 2012 for households with children below 6 years of age, in November 2012 for households with children below 15 years of age, and in February 2013 for all remaining households. These reforms thus created a kink (which varies with household composition) in the benefit schedule as a function of reported, per-capita household income for households with children as well as for two adults households with no children. For example, if a two adult household without kids prior to June 2014 has a reported per-capita income of R$20, they get R$70 from the basic benefit and then get an additional R$30 to bring them up to the guaranteed minimum per-capita income of R$70. Mathematically, prior to June 2014 two adult households without kids face a

61. Households with children need to fulfill three additional conditions to receive the variable benefit and/or teenager benefit: (1) children must maintain a minimum of 85% school attendance between ages 6 and 15 and 75% school attendance between 16 and 17; (2) households must keep track of their children's vaccines; and (3) parents must maintain at least 85% attendance in a social-education program if the household has violated child labor laws in the past. All conditionalities were held constant during the analysis period.

benefit schedule as a function of reported per-capita income, $\hat{y}$, equal to:

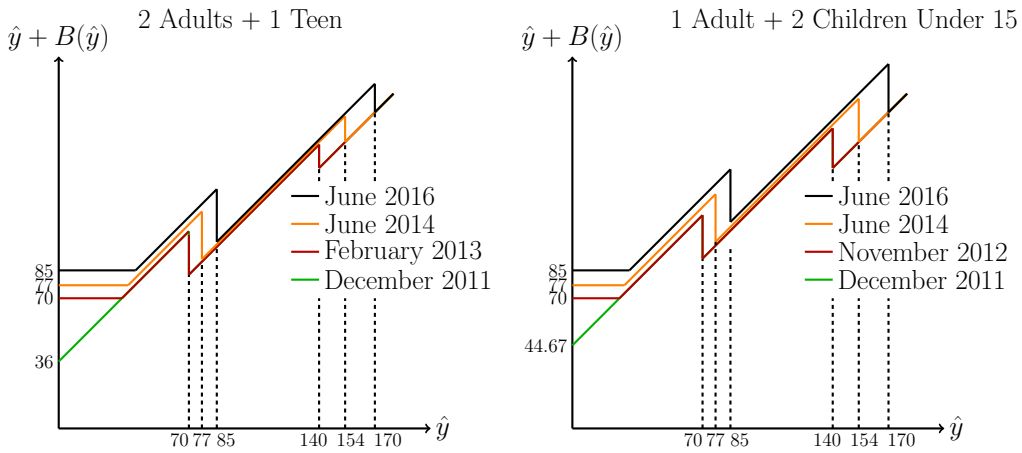
$$B(\hat{y}) = 70\mathbb{1}(\hat{y} \leq 70) + \max\{0, 70 - 2 \times \hat{y}\}$$

This benefit schedule therefore has a kink at R$35. The kink then changed slightly with the June 2014 reform as the guaranteed minimum income was raised to R$77 per-capita. For example, after June 2014 two adult households without kids face a benefit schedule with a kink at R$38.5:

$$B(\hat{y}) = 77\mathbb{1}(\hat{y} \leq 77) + \max\{0, 77 - 2 \times \hat{y}\}$$

Thus, the June 2014 reform potentially impacted the reported income distribution around the kink. Because the kink is located below the first BF notch of R$70, Identification Assumption 1 is less likely to hold for these households as households with reported incomes under R$63 may respond to this changing kink. Finally, a reform in June 2016 further increased the extreme-poverty threshold to R$85 (per-capita, per-month), the poverty threshold to R$170 (per-capita, per-month), the basic benefit to R$85 (per-month), the variable benefit R$39 (per-child, per-month), and the teenager benefit to R$46 (per-teenager, per-month).

For purposes of illustration, Figure 11 plots how the benefit schedules for two particular household compositions varied over time.

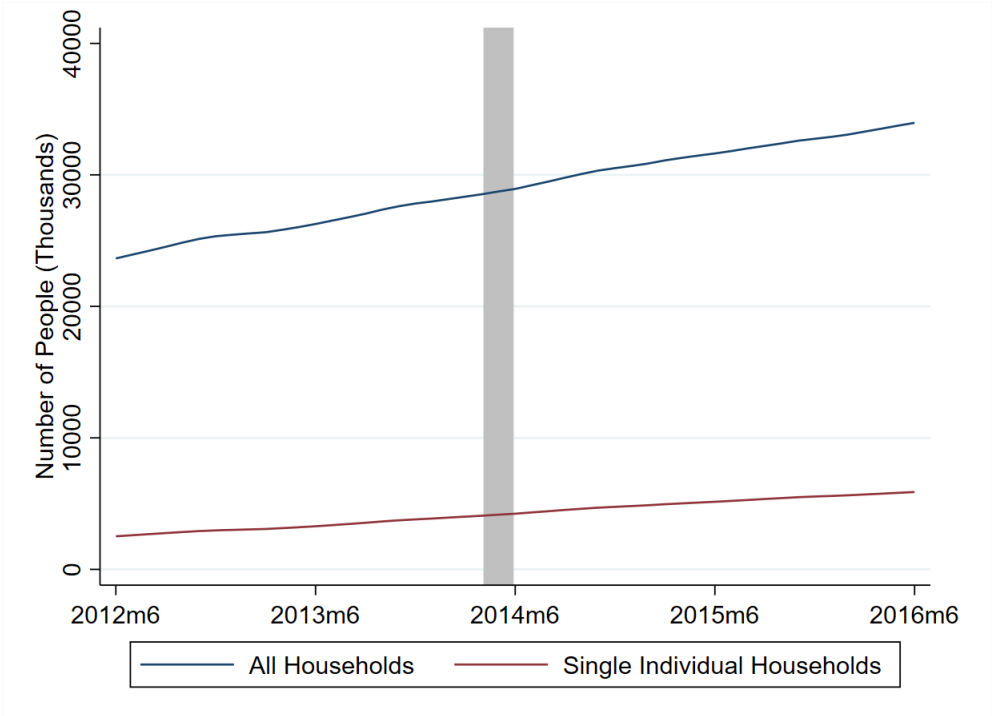


Note: $\hat{y}$ denotes the reported, per-capita, monthly household income. $B(\hat{y})$ denotes the monthly, per-capita benefits a household receives if they report $\hat{y}$. These benefits will also depend on household composition. For example, a household with 2 adults and 1 teenager reporting $\hat{y} = 0$ in December 2011 will receive R$70 in the basic benefit and R$38 in the teenager benefit. Thus $\hat{y} + B(\hat{y}) = (70 + 38)/3 = 36$.

FIGURE 11: BOLSA FAMÍLIA SCHEDULE REFORMS FOR TWO EXAMPLE HOUSEHOLD COMPOSITIONS

C Data Appendix

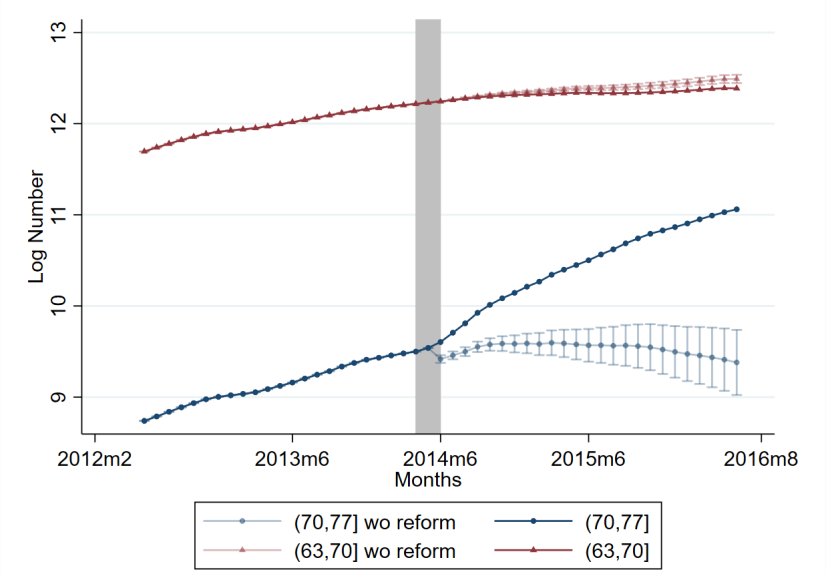
C.1 Number of Individuals on the Cadastro Único Registry Over Time



Note: This figure shows the raw number of individuals on the Cadastro Único Registry over time. The timing of the reform is indicated by the gray, shaded region.

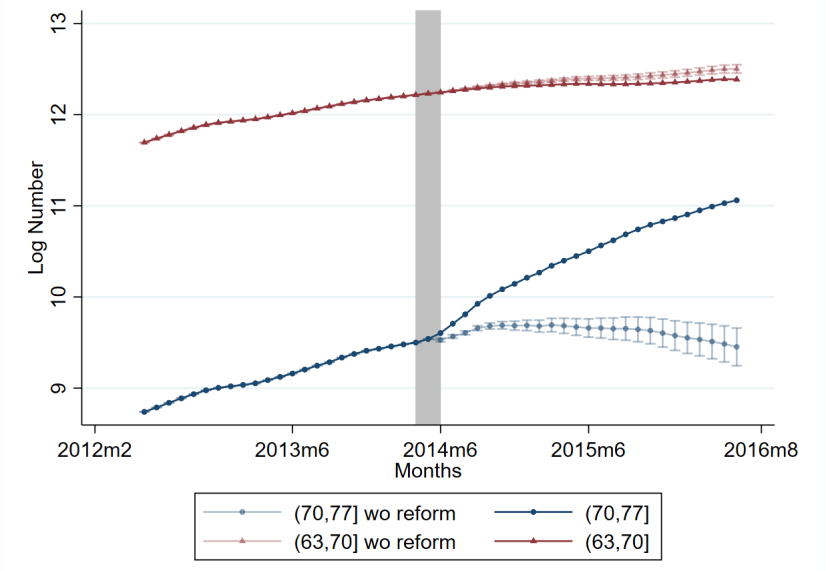
Figure 12: Number of Individuals on the Cadastro Único Registry

C.2 *Actual and Counterfactual Paths Estimated from Equation (15) with Different Polynomial Degrees*



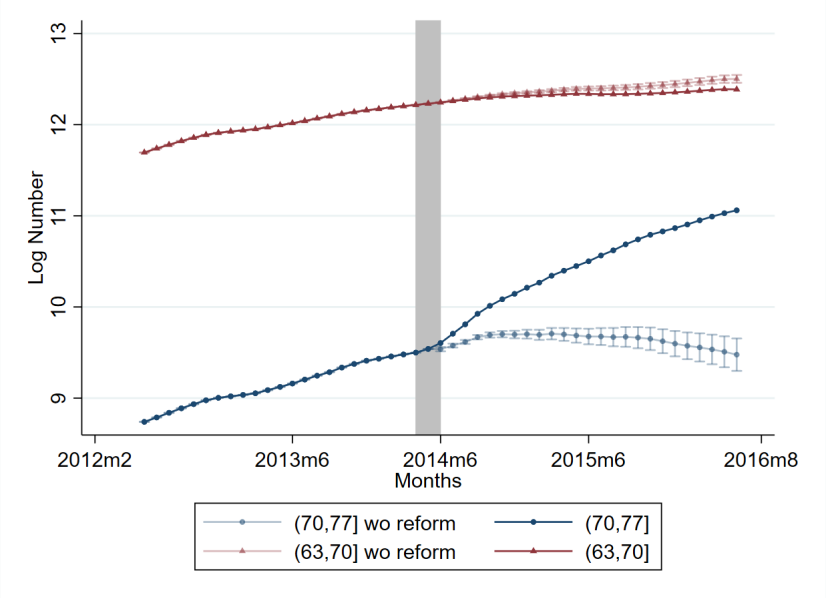
Note: This figure shows the log number of single individual households reporting incomes in R$(63, 70] and R$(70, 77] over time along with the counterfactual paths had the reform not happened. The counterfactual paths are equal to the actual number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}t$, estimated using Equation (15) where we set $treat_x = 1$ if $x \in \{70, 77\}$ and $K = 2$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

FIGURE 13: ACTUAL AND COUNTERFACTUAL PATHS FOR TREATMENT BINS, $J = 2$



Note: This figure shows the log number of single individual households reporting incomes in R$(63, 70] and R$(70, 77] over time along with the counterfactual paths had the reform not happened. The counterfactual paths are equal to the actual number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}t$, estimated using Equation (15) where we set $treat_x = 1$ if $x \in \{70, 77\}$ and $K = 4$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

FIGURE 14: ACTUAL AND COUNTERFACTUAL PATHS FOR TREATMENT BINS, $J = 4$

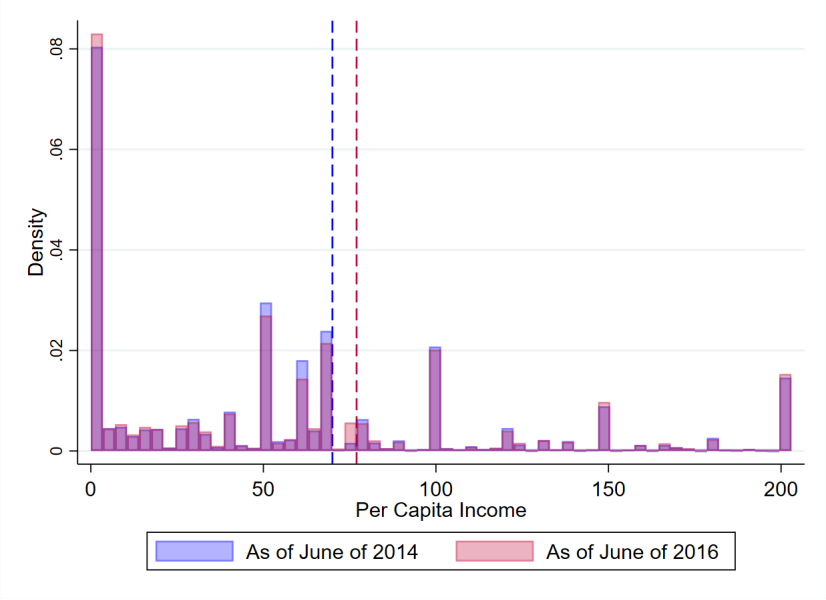


Note: This figure shows the log number of single individual households reporting incomes in R$(63, 70] and R$(70, 77] over time along with the counterfactual paths had the reform not happened. The counterfactual paths are equal to the actual number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}t$, estimated using Equation (15) where we set $treat_x = 1$ if $x \in \{70, 77\}$ and $K = 5$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

FIGURE 15: ACTUAL AND COUNTERFACTUAL PATHS FOR TREATMENT BINS, $J = 5$

C.3 Histograms of Reported Income Distribution

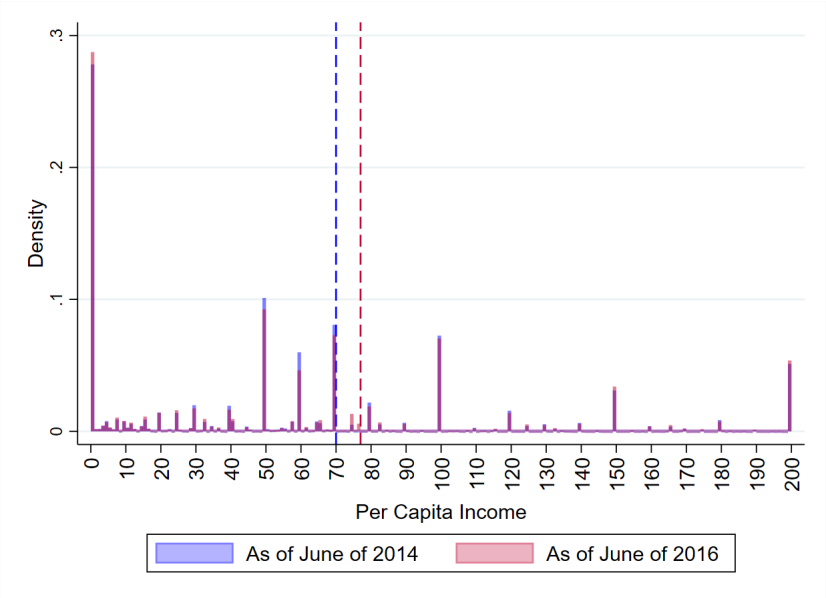
Figure 16 plots the distribution of reported incomes in June 2014 and June 2016 for single individual households split into seven increment bins. While these histograms show the pre- and post-reform distributions, they *should not* be used to make inferences about the causal impact of the reform due to significant underlying time trends in the reported income distribution.



Note: This figure shows the number of single individual households that report incomes in seven increment bins in the period before the reform (June 2014) and two years after the reform (June 2016). The extreme poverty threshold before June 2014 is shown with a dashed blue line and the extreme poverty threshold after June 2014 is shown with a dashed red line.

FIGURE 16: HISTOGRAM OF SINGLE INDIVIDUAL HOUSEHOLD REPORTED INCOMES PRE- AND POST-REFORM

Figure 17 shows a more granular view of the same reported income distributions for single individual households split into one increment bins. There is substantial bunching at 0 mod 50 numbers and a lesser degree of bunching at 0 mod 10 numbers:

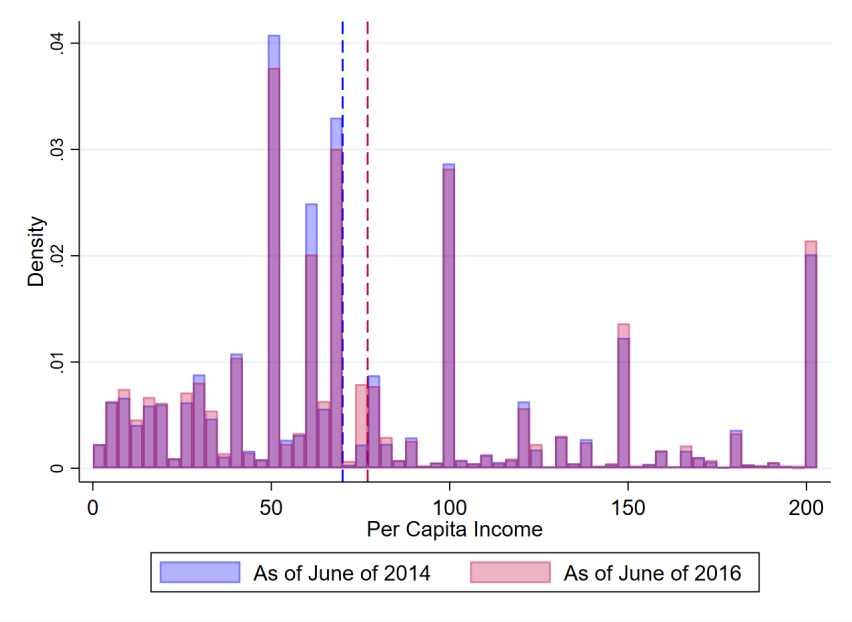


Note: This figure shows the number of single individual households that report incomes in one increment bins in the period before the reform (June 2014) and two years after the reform (June 2016). The extreme poverty threshold before June 2014 is shown with a dashed blue line and the extreme poverty threshold after June 2014 is shown with a dashed red line.

FIGURE 17: GRANULAR HISTOGRAM OF SINGLE INDIVIDUAL HOUSEHOLD REPORTED INCOMES PRE- AND POST-REFORM

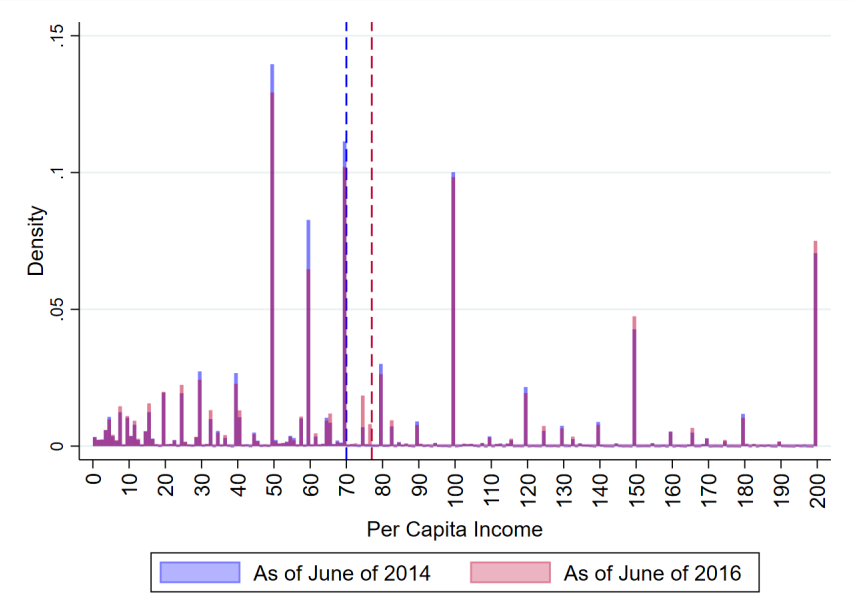
Finally, for better visualization of the reported income distribution for per-capita incomes>0, we repeat

Figures 16 and 17 restricting to households with strictly positive per-capita incomes:



Note: This figure shows the number of single individual households that report incomes in seven increment bins in the period before the reform (June 2014) and two years after the reform (June 2016). The extreme poverty threshold before June 2014 is shown with a dashed blue line and the extreme poverty threshold after June 2014 is shown with a dashed red line.

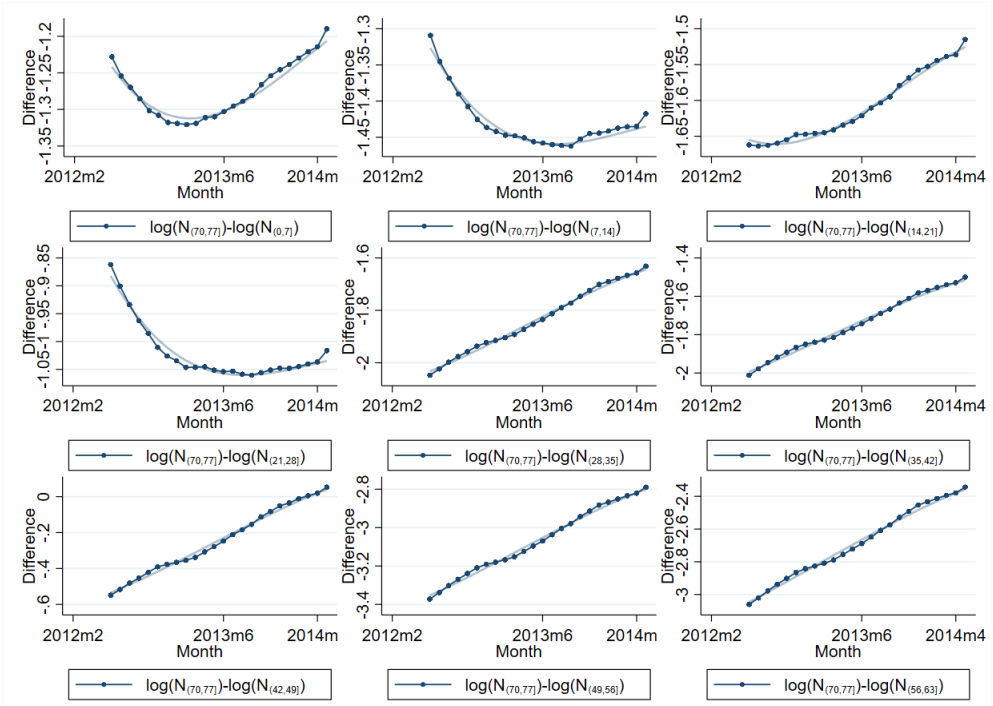
FIGURE 18: HISTOGRAM OF SINGLE INDIVIDUAL HOUSEHOLD REPORTED INCOMES PRE- AND POST-REFORM (STRICTLY POSITIVE INCOMES ONLY)



Note: This figure shows the number of single individual households that report incomes in one increment bins in the period before the reform (June 2014) and two years after the reform (June 2016). The extreme poverty threshold before June 2014 is shown with a dashed blue line and the extreme poverty threshold after June 2014 is shown with a dashed red line.

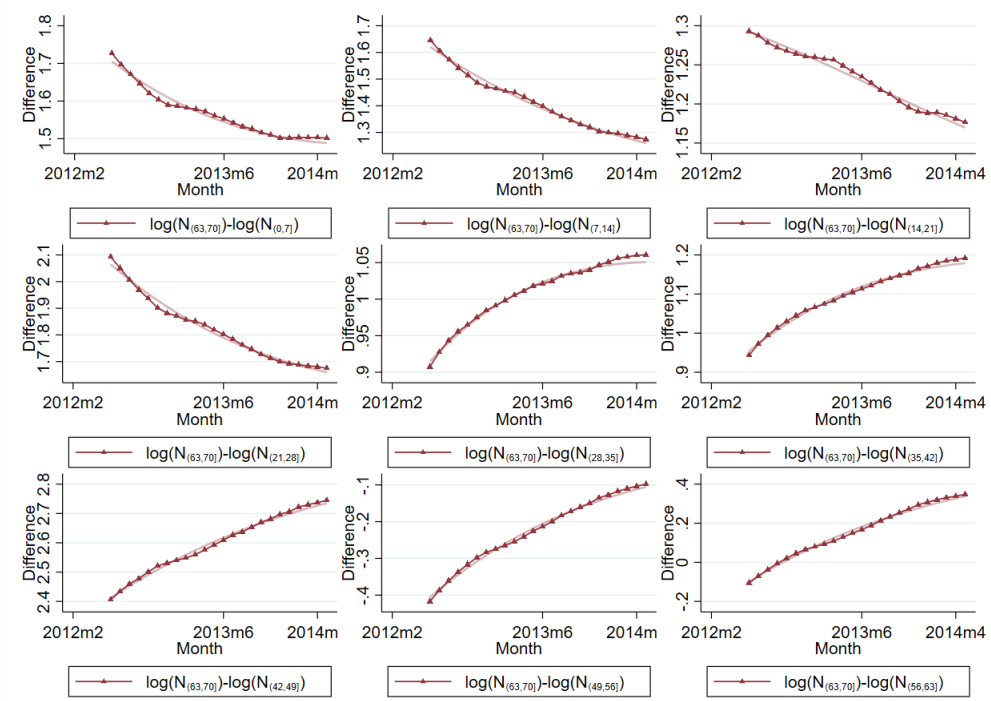
FIGURE 19: GRANULAR HISTOGRAM OF SINGLE INDIVIDUAL HOUSEHOLD REPORTED INCOMES PRE- AND POST-REFORM (STRICTLY POSITIVE INCOMES ONLY)

C.4 Pre-Reform Differences



Note: This figure shows how the difference between the log number of people reporting incomes in R$(70,77], denoted $\log\left(N_{(70,77]}\right)$, and the log number of people reporting in R$(x-7,x], denoted $\log\left(N_{(x-7,x]}\right)$, varied prior to the reform in June, 2014. Each plot also includes the cubic trend estimated from Equation (15).

FIGURE 20: PRE-REFORM DIFFERENCES BETWEEN $N_{(70,77]}$ AND $N_{(x-7,x]}$ FOR $x \in \{7, 14, ..., 63\}$

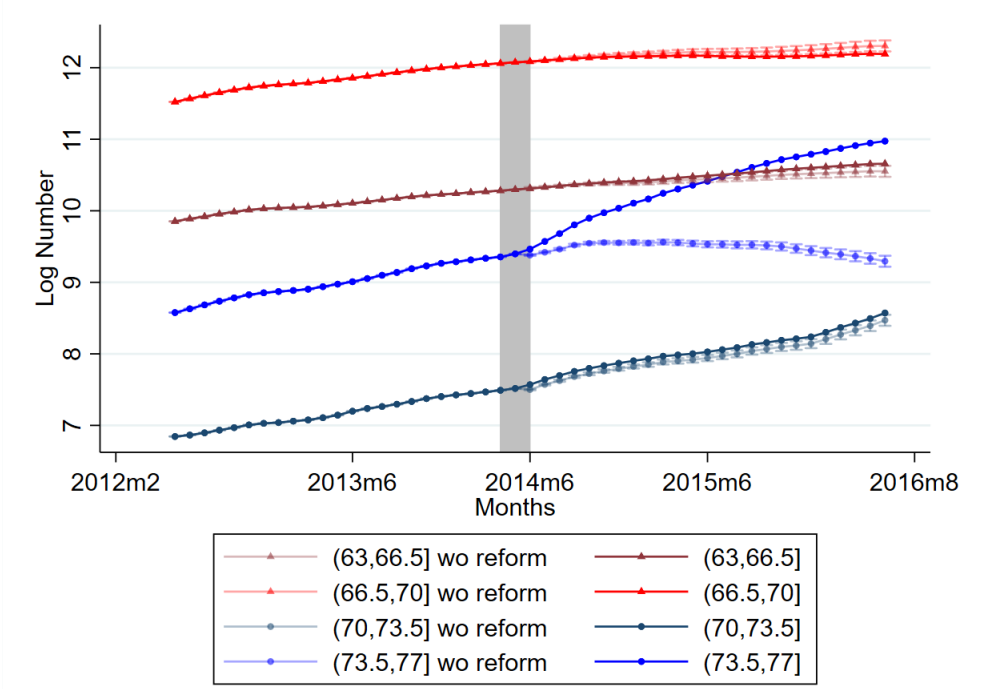


Note: This figure shows how the difference between the log number of people reporting incomes in R$(63,70], denoted $\log\left(N_{(63,70]}\right)$, and the log number of people reporting in R$(x-7,x], denoted $\log\left(N_{(x-7,x]}\right)$, varied prior to the reform in June, 2014. Each plot also includes the cubic trend estimated from Equation (15).

FIGURE 21: PRE-REFORM DIFFERENCES BETWEEN $N_{(63,70]}$ AND $N_{(x-7,x]}$ FOR $x \in \{7, 14, ..., 63\}$

C.5 Results with Smaller Bin Sizes

This Appendix contains results using a bin size of 3.5 rather than a bin size of 7 as in the main text. We estimate Equation (15) for all $x \in \{3.5, 7, 10.5, ..., 77\}$. Estimated counterfactuals for bins R$(63,66.5], R$(66.5,70], R$(70,73.5], and R$(73.5,77] are shown in Figure 22. The estimated treatment effects are large for R$(66.5,70] and R$(73.5,77] but small for R$(63,66.5] and R$(70,73.5] suggesting that bunching is relatively precise. The estimated effects and corresponding lower and upper bounds for the MVPF are shown in Table 6. The estimated MVPF bounds are relatively similar to our main results in Table 3.



Note: This figure shows the log number of single individual households reporting incomes in R$(63,66.5], R$(66.5,70], R$(70,73.5], and R$(73.5,77] over time along with the counterfactual paths had the reform not happened. The counterfactual paths are equal to the actual number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}t$, estimated using Equation (15) where we set $treat_x = 1$ if $x \in \{66.5, 70, 73.5, 77\}$ and $K = 3$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

FIGURE 22: ACTUAL AND COUNTERFACTUAL PATHS FOR TREATMENT BINS USING EQUATION (15) WITH SMALLER BIN SIZES

Table 6: IMPACTS OF REFORM AND MVPF BOUNDS IN JUNE 2016 ESTIMATED FROM EQUATION (15) USING SMALLER BIN SIZES

	(1)	(2)	(3)	(4)	(5)	(6)
Polynomial Degree, K	$\Delta N_{(63,66.5],\bar{t}}$	$\Delta N_{(66.5,70],\bar{t}}$	$\Delta N_{(70,73.5],\bar{t}}$	$\Delta N_{(73.5,77],\bar{t}}$	$MVPF_{L,\bar{t}}$	$MVPF_{U,\bar{t}}$
Quadratic, $K = 2$	5,362	-20,258	563	49,646	0.84	1.06
	(1,831)	(4,331)	(517)	(1,884)	(0.02)	(0.02)
Cubic, $K = 3$	4,270	-23,606	514	47,450	0.87	1.08
	(1,501)	(8,659)	(187)	(427)	(0.04)	(0.05)
Quartic, $K = 4$	-8,170	-10,852	3,421	48,206	0.85	1.08
	(3,253)	(6,265)	(194)	(1,177)	(0.03)	(0.03)
Quintic, $K = 5$	-5,393	-12,304	3,338	47,911	0.85	1.07
	(1,139)	(6,133)	(141)	(1,884)	(0.03)	(0.03)

Note: Columns (1), (2), (3), and (4) show the estimated impacts of the reform on the number of single individual households reporting incomes in bins R$(63,66.5], R$(66.5,70], R$(70,73.5], and R$(73.5,77] for June 2016. Estimates are calculated using Equation (15) with various polynomial degrees $K \in \{2, 3, 4, 5\}$. Columns (5) and (6) show the estimated upper and lower bounds for the MVPF for June 2016, calculated using Equations (10) and (11). Standard errors are presented in parentheses and are computed from the delta method using the clustered standard errors estimated in Equation (15).

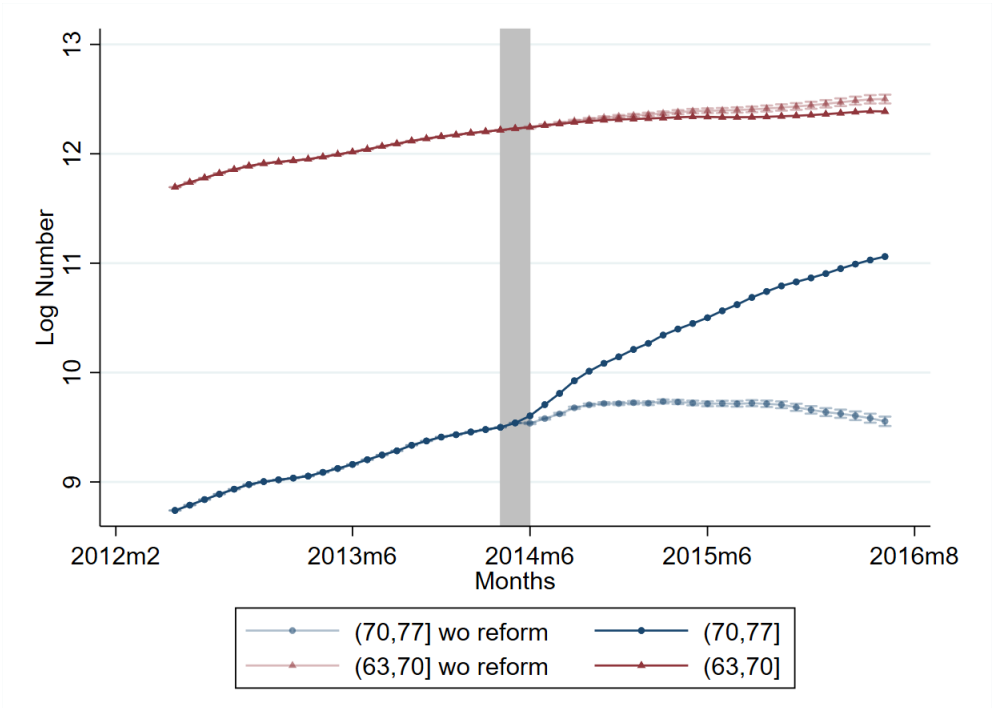
C.6 Results Using R$(56,63] as a Treatment Bin

Table 7: Impacts of Reform and MVPF Bounds in June 2016 Estimated from Equation (15) Using R$(56,63] as Treatment Bin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Polynomial Degree, K	$\Delta N_{(56,63],\bar{t}}$	$\Delta N_{(63,70],\bar{t}}$	$\Delta N_{(70,77],\bar{t}}$	$B_{\bar{t}}$	$J_{\bar{t}}$	$MVPF_{L,\bar{t}}$	$MVPF_{U,\bar{t}}$
Quadratic, $K=2$	13,173	-23,591	51,879	10,418	41,461	0.82	1.04
	(5,309)	(6,397)	(2,144)	(9,607)	(9,995)	(0.04)	(0.04)
Cubic, $K=3$	8,579	-25,719	49,340	17,140	32,199	0.85	1.07
	(5,381)	(3,580)	(682)	(7,006)	(7,119)	(0.03)	(0.03)
Quartic, $K=4$	9,825	-27,331	50,968	17,507	33,462	0.85	1.07
	(8,922)	(6,275)	(1,338)	(11,592)	(11,759)	(0.05)	(0.05)
Quintic, $K=5$	12,489	-26,669	50,683	14,180	36,503	0.83	1.06
	(10,645)	(6,397)	(1,175)	(12,709)	(12,841)	(0.05)	(0.06)

Note: Columns (1) (2), and (3) show the estimated impacts of the reform on the number of single individual households reporting incomes in bins R$(56,63], R$(63,70], and R$(70,77] for June 2016: $\Delta N_{(56,63],\bar{t}}$, $\Delta N_{(63,70],\bar{t}}$, and $\Delta N_{(70,77],\bar{t}}$. Estimates are calculated using Equation (15) with various polynomial degrees $K \in \{2,3,4,5\}$. Columns (4) and (5) show the estimated number of bunching and jumping households for June 2016, $B_{\bar{t}}$ and $J_{\bar{t}}$, calculated using Equations (12) and (13). Columns (6) and (7) show the estimated upper and lower bounds for the MVPF for June 2016, calculated using Equations (10) and (11). Standard errors are presented in parentheses and are computed from the delta method from the clustered standard errors estimated in Equation (15).

C.7 Results Estimated from Nonlinear Least Squares Equation (16)



Note: This figure shows the log number of single individual households reporting incomes in R$(63, 70] and R$(70, 77] over time along with the counterfactual paths had the reform not happened. The counterfactual paths are equal to the actual number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}t$, estimated using Equation (16) where we set $treat_x = 1$ if $x \in \{70, 77\}$ and $K = 3$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

Figure 23: Counterfactual Paths for Treatment Bins from Equation (16)

Table 8: Impacts of Reform and MVPF Bounds in June 2016 Estimated from Equation (16)

	(1)	(2)	(3)	(4)	(5)	(6)
Polynomial Degree, K	$\Delta N_{(63,70],\bar{t}}$	$\Delta N_{(70,77],\bar{t}}$	$B_{\bar{t}}$	$J_{\bar{t}}$	$MVPF_{L,\bar{t}}$	$MVPF_{U,\bar{t}}$
Quadratic, $K = 2$	-29,038	50,724	29,038	21,686	0.90	1.13
	(13,711)	(635)	(13,711)	(13,723)	(0.06)	(0.07)
Cubic, $K = 3$	-28,834	49,505	28,834	20,671	0.90	1.13
	(5,485)	(309)	(5,485)	(5,532)	(0.02)	(0.03)
Quartic, $K = 4$	-23,825	50,824	23,825	26,999	0.87	1.10
	(9,358)	(485)	(9,358)	(9,429)	(0.04)	(0.04)
Quintic, $K = 5$	-24,571	51,228	24,571	26,658	0.87	1.11
	(8,311)	(426)	(8,311)	(8,374)	(0.04)	(0.04)

Note: Columns (1) and (2) show the estimated impacts of the reform on the number of single individual households reporting incomes in bins R$(63,70] and R$(70,77] for June 2016: $\Delta N_{(63,70],\bar{t}}$ and $\Delta N_{(70,77],\bar{t}}$. Estimates are calculated from the nonlinear least squares regression given by Equation (16) with various polynomial degrees $J \in \{2, 3, 4, 5\}$. Columns (3) and (4) show the estimated number of bunching and jumping households for June 2016, $B_{\bar{t}}$ and $J_{\bar{t}}$, calculated using Equations (12) and (13). Columns (5) and (6) show the estimated upper and lower bounds for the MVPF for June 2016, calculated using Equations (10) and (11). Standard errors are presented in parentheses and are computed from the delta method from the clustered standard errors estimated in Equation (15).

C.8 Results for Households with Constant Composition

In this appendix, we present results for our sample of single individual households restricted to those who do not change their reported family composition throughout the analysis period (June 2012 to June 2016). To do so, we drop any household that reports a composition change over the analysis period. Moreover, we drop any household who enters the registry post June-2014 as we cannot tell whether these new entrants are truly new households or are pre-existing households (with multiple adults) pretending to be separate households so as to increase the amount of benefits they receive. In other words, we restrict our sample to single individual households who were (a) on the registry prior to June 2014, and (b) do not report a change in composition over the analysis period. This reduces our sample from 1,938,653 single individual households with incomes below R$77 in June 2016 to 1,039,573 single individual households with incomes below R$77 in June 2016. Table 9 presents results for this exercise. We find very similar estimates for the MVPF lower bound and slightly smaller estimates for the MVPF upper bound. Note that the estimated numbers of jumpers and bunchers are smaller than in Table 3 due to the smaller sample size.

Table 9: Impacts of Reform and MVPF Bounds in June 2016 Estimated from Equation (15), Constant Composition Households

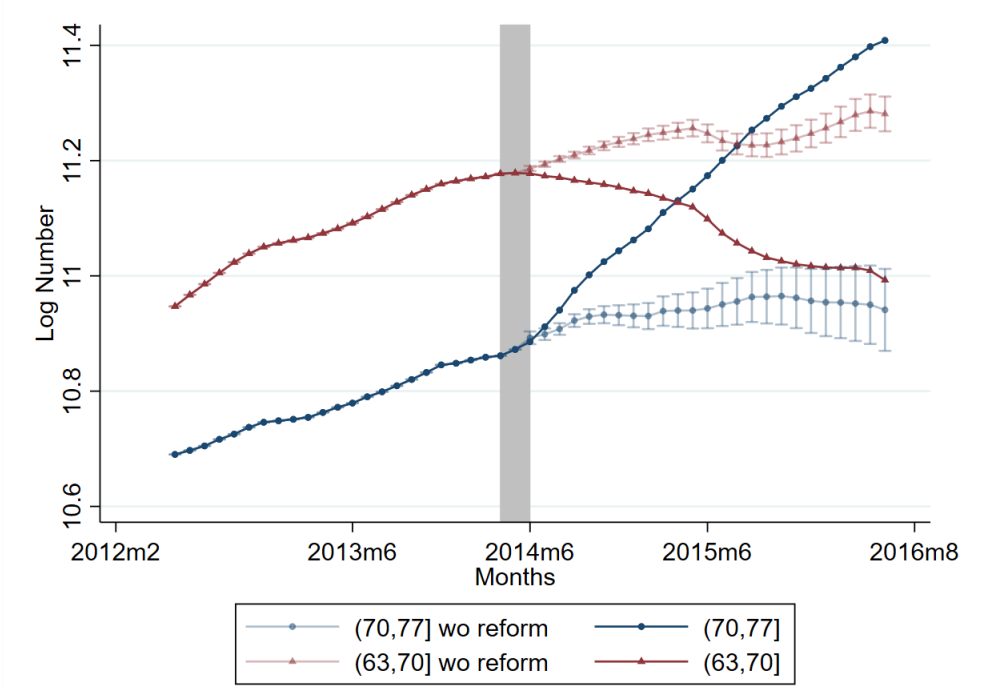
	(1)	(2)	(3)	(4)	(5)	(6)
Polynomial Degree, K	$\Delta N_{(63,70],\bar{t}}$	$\Delta N_{(70,77],\bar{t}}$	$B_{\bar{t}}$	$J_{\bar{t}}$	$MVPF_{L,\bar{t}}$	$MVPF_{U,\bar{t}}$
Quadratic, $K = 2$	-4,714	19,019	4,714	14,305	0.88	1.04
	$(3,071)$	$(1,106)$	$(3,071)$	$(3,404)$	(0.03)	(0.02)
Cubic, $K = 3$	-6,847	17,713	6,847	10,866	0.91	1.05
	(2,345)	(619)	$(2,345)$	$(2,524)$	(0.02)	(0.02)
Quartic, $K = 4$	-862	13,722	862	12,860	0.89	1.01
	(3,483)	$(1,696)$	$(3,483)$	$(3,984)$	(0.03)	(0.03)
Quintic, $K = 5$	-1,586	14,679	1,586	13,094	0.89	1.01
	(3,302)	(1,188)	(3,302)	(3,661)	(0.03)	(0.03)

Note: Columns (1) and (2) show the estimated impacts of the reform on the number of single individual households reporting incomes in bins R$(63,70] and R$(70,77] for June 2016: $\Delta N_{(63,70],\bar{t}}$ and $\Delta N_{(70,77],\bar{t}}$. Estimates are calculated from Equation (15) with various polynomial degrees $K \in \{2,3,4,5\}$, restricting the sample to single individual households who do not change their reported household composition throughout the analysis period. Columns (3) and (4) show the estimated number of bunching and jumping households for June 2016, $B_{\bar{t}}$ and $J_{\bar{t}}$, calculated using Equations (12) and (13). Columns (5) and (6) show the estimated upper and lower bounds for the MVPF for June 2016, calculated using Equations (10) and (11). Standard errors are presented in parentheses and are computed from the delta method from the clustered standard errors estimated in Equation (15).

C.9 Results for Two Adult Households with No Children

In this appendix, we present results for households with two adults and no children. As discussed in Section 3.2 and Appendix B.4, the guaranteed minimum income, which was instituted in February 2013 for two adult households without kids, generates a kink in the benefit schedule that occurs at a per-capita household income of R$35. Hence, after February 2013 and prior to June 2014, two adult families in extreme poverty (i.e., with household incomes below R$70) get the basic benefit of R$70 plus additional funds to bring their per-capita income up to R$70. For example, a two adult family with a combined household income of R$40 gets R$70 *plus* an additional R$30 (giving them a total income of R$140) to reach the guaranteed minimum per-capita income of R$70. This generates a kink in the benefit schedule for two adult families at the per-capita household income of R$35 after February 2013 and prior to June 2014.

The June 2014 reform not only changed the extreme poverty threshold and the basic benefit both from R$70 to R$77 but it also changed the guaranteed minimum income from R$70 per-capita to R$77 per-capita. Hence, the June 2014 reform changed the location and level of both the notch and the kink in the benefit schedule (for example, the location of the notch moved from R$70 to R$77 per-capita while the location of the kink moved from R$35 to R$38.5 per-capita). Identification Assumption 1 is now harder to justify given that the June 2014 reform changed incentives for households to locate around R$35 by changing the kink in the benefit schedule from R$35 to R$38.5. However, many studies have shown that behavioral responses to kinks are typically very small; kinks generally induce substantially less bunching than do notches (Kleven, 2016). Hence, we simply ignore the presence of the kink and use all seven-increment income bins below R$63 as control groups just as in our main analysis. Figure 24 shows the analogue of Figure 6 for two adult households with no kids. There is a clear increase in the number of households locating in R$(70, 77] and a clear decrease in the fraction locating between R$(63, 70] just as for single individual households.



Note: This figure shows the log number of single individual households reporting incomes in R$(63, 70] and R$(70, 77] over time along with the counterfactual paths had the reform not happened. The sample is restricted to two adult households without children. The counterfactual paths are equal to the actual number of people reporting in the given interval minus the causal impact of the reform, $\hat{\beta}_{1,x} + \hat{\beta}_{2,x}t$, estimated using Equation (15) where we set $treat_x = 1$ if $x \in \{70, 77\}$ and $K = 3$. Confidence intervals are constructed from clustered standard errors at the bin level. The timing of the reform is indicated by the gray, shaded region.

FIGURE 24: COUNTERFACTUAL PATHS FOR TREATMENT BINS, TWO ADULT HOUSEHOLDS WITH NO CHILDREN

Results from estimating Equation (15) for various polynomial degrees K can be found in Table 10. The MVPF bounds are roughly similar in magnitude as for the single individual households discussed in Section 5 but are a bit more sensitive to the degree of polynomial used.[62]

62. The bounds for the MVPF are the sufficient statistic for the welfare impact of the reform, ignoring any welfare impacts of the increased guaranteed minimum income.

Table 10: Impacts of Reform and Efficiency Loss as of June, 2016 Estimated from Equation (15), Two Adult Households with No Children

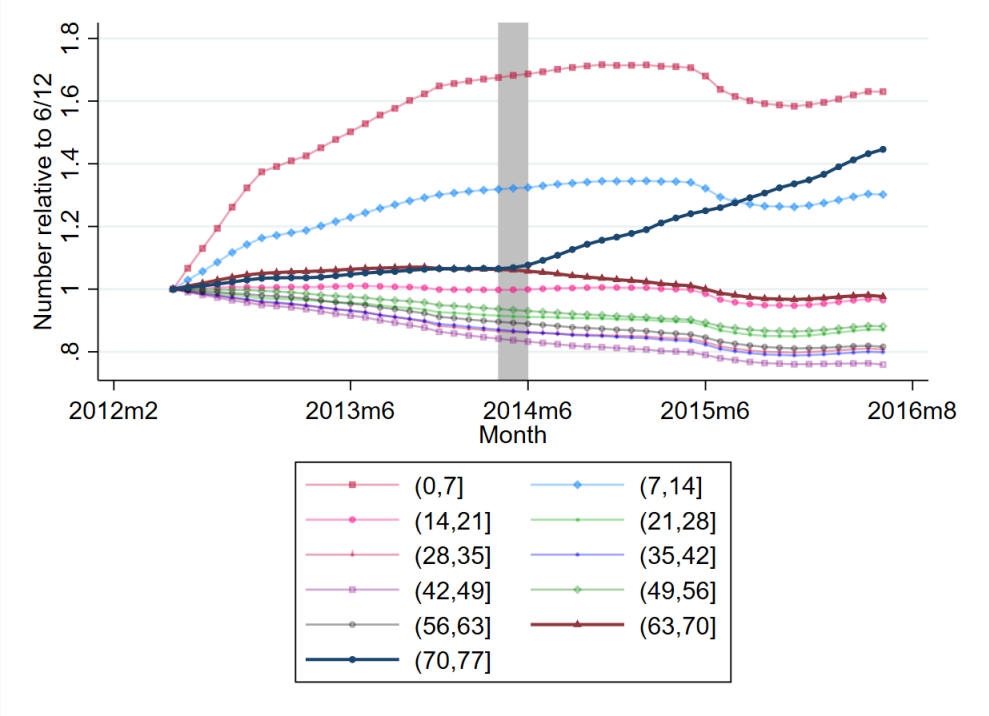
	(1)	(2)	(3)	(4)	(5)	(6)
Polynomial Degree, K	$\Delta N_{(63,70],\bar{t}}$	$\Delta N_{(70,77],\bar{t}}$	$B_{\bar{t}}$	$J_{\bar{t}}$	$MVPF_{L,\bar{t}}$	$MVPF_{U,\bar{t}}$
Quadratic, $K = 2$	-19,897	37,055	19,897	17,157	0.88	1.12
	(1,795)	(3,592)	(1,795)	(4,292)	(0.03)	(0.01)
Cubic, $K = 3$	-19,862	33,645	19,862	13,783	0.91	1.12
	(1,221)	(2,048)	(1,221)	(2,536)	(0.02)	(0.01)
Quartic, $K = 4$	-14,347	21,043	14,347	6,696	0.96	1.08
	(2,100)	(5,011)	(2,100)	(5,610)	(0.04)	(0.01)
Quintic, $K = 5$	-14,157	25,714	14,157	11,557	0.93	1.08
	(1,883)	(2,625)	(1,883)	(3,467)	(0.02)	(0.01)

Note: Columns (1) and (2) show the estimated impacts of the reform on the number of single individual households reporting incomes in bins R$(63,70] and R$(70,77] for June 2016: $\Delta N_{(63,70],\bar{t}}$ and $\Delta N_{(70,77],\bar{t}}$. Estimates are calculated from Equation (15) with various polynomial degrees $K \in \{2, 3, 4, 5\}$, restricting the sample to two adult households without children. Columns (3) and (4) show the estimated number of bunching and jumping households for June 2016, $B_{\bar{t}}$ and $J_{\bar{t}}$, calculated using Equations (12) and (13). Columns (5) and (6) show the estimated upper and lower bounds for the MVPF for June 2016, calculated using Equations (10) and (11). Standard errors are presented in parentheses and are computed from the delta method from the clustered standard errors estimated in Equation (15).

C.10 Results for Households with Kids

The benefit schedule for families with children is substantially more complex than the benefit schedule for households without children, as discussed in detail Appendix B.4. Prior to June 2014, in addition to the guaranteed minimum income and the basic benefit, there was also a variable per-child benefit for households below the poverty threshold of R$140 per-capita. The June 2014 reform led to changes in the levels and locations of the guaranteed minimum income kink, the basic benefit notch, and the variable benefit notch. For example, the poverty threshold was raised from R$140 per-capita to R$154 per-capita and the levels of the variable benefits were also increased by around 10% (see Appendix B.4 for more details).

Estimating the MVPF of the reform for households with children will require estimating the WTP for the all of the different changes to the BF schedule. Estimating the WTP for the change to the variable benefit schedule would be particularly difficult given that this benefit is made conditional on investments in children. For example, suppose parents are under-investing in their children's education and the reform increases school attendance. Then we would need to estimate the childrens' WTP for the increased education they receive as a result of the reform to the variable benefit schedule. Thus, we leave calculating the MVPF of the reform for households with children to future work. We do, however, show strong evidence that households with kids respond to the reform. In particular, Figure 25 shows *prima facie* evidence that the reform increased the number of households reporting incomes in R$(70, 77].



Note: This figure shows the number of households with kids that report incomes in the various bins. The number in each bin is normalized to 1 in June, 2012. The timing of the reform is indicated by the gray, shaded region.

FIGURE 25: NUMBER OF HOUSEHOLDS WITH KIDS REPORTING INCOMES IN VARIOUS BINS, NORMALIZED TO 1 IN JUNE, 2012