STAFF REPORTS

NO. 902
NOVEMBER 2019

REVISED
JULY 2021

Latent Heterogeneity in the Marginal Propensity to Consume

Daniel J. Lewis | Davide Melcangi | Laura Pilossoph

FEDERAL RESERVE BANK of NEW YORK

Latent Heterogeneity in the Marginal Propensity to Consume

Daniel J. Lewis, Davide Melcangi, and Laura Pilossoph

Federal Reserve Bank of New York Staff Reports, no. 902
November 2019; revised July 2021
JEL classification: D12, D91, E21, E32, E62

Abstract

We estimate the unconditional distribution of the marginal propensity to consume (MPC) using clustering regression and the 2008 stimulus payments. Since we do not measure heterogeneity as the variation of MPCs with observables, we can recover the full distribution of MPCs. Households spent at least one quarter of the rebate, and individual households used rebates for different goods. While many observables are individually correlated with our estimated MPCs, these relationships disappear when tested jointly, except for nonsalary income and the average propensity to consume. Household observables explain at most one quarter of MPC variation, highlighting the role of unobserved heterogeneity.

Key words: marginal propensity to consume, consumption, tax rebate, heterogeneous treatment effects, clustering, c-means

---

Lewis, Melcangi, Pilossoph: Federal Reserve Bank of New York (emails: daniel.lewis@ny.frb.org, davide.melcangi@ny.frb.org, laura.pilossoph@ny.frb.org). This paper was previously circulated under the title “A New Approach to Estimating Heterogeneous Effects and the Distribution of the Marginal Propensity to Consume.” The authors thank René Chalom and Meghana Gaur for excellent research assistance. For helpful comments, they also thank Sara Casella, Richard Crump, Marco Del Negro, Keshav Dogra, Domenico Giannone, Simon Gilchrist, Ulrich Müller, Mikkel Plagborg-Møller, Isaac Sorkin, and Mary Wootters, as well as various seminar and conference participants.

To view the authors’ disclosure statements, visit https://www.newyorkfed.org/research/staff_reports/sr902.html.

1 Introduction

Recent work highlights the importance of heterogeneity in the marginal propensity to consume (MPC) out of transitory income shocks for fiscal policy, the transmission of monetary policy, and welfare.[1] Despite their importance, estimates of the *distribution* of MPCs are largely elusive. Even with plausibly identified transitory income shocks, estimating individual-level MPCs requires panel data with long horizons, which are typically not available; it also usually requires the unappealing assumption that an individual's marginal propensity to consume (MPC) is time invariant.[2] The existing literature, therefore, has followed one of two avenues: estimating a fully structural model and simulating a distribution of MPCs, or grouping observations by some presupposed observable characteristics and estimating group-specific MPCs out of transitory income shocks.[3] However, because both of these approaches require taking a stance on the source of MPC heterogeneity, they may fail to uncover the true degree of heterogeneity, miss other relevant dimensions of heterogeneity that predict an individual's MPC, or both.

In this paper, we estimate the distribution of MPCs directly. We adopt a Gaussian mixture linear regression (GMLR) (e.g., Quandt (1972)), which jointly (i) groups households together that have similar latent consumption responses to the 2008 tax rebate and (ii) provides estimates of the MPCs within these groups. Specifically, the algorithm takes a standard regression of consumption changes on the tax rebate receipt and basic controls (Johnson, Parker, and Souleles (2006), Parker, Souleles, Johnson, and McClelland (2013)), but allows the coefficient on the rebate to be heterogeneous across unknown groups; the groups as well as their rebate coefficients are then jointly estimated.

This approach offers four advantages over existing efforts to recover the distribution of MPCs. First, it allows us to estimate the full unconditional distribution of MPCs, which can be driven both by latent factors and observable characteristics, broadly defined; understanding the range of such a distribution casts light on whether there is potential value, in principle, in attempting to target fiscal transfers to households more likely to spend

[1]The MPC distribution is a crucial object in Heterogeneous Agent New Keynesian (HANK) models of monetary policy (see Kaplan, Moll, and Violante (2018)). For example, Auclert (2019) shows that the response of aggregate consumption to monetary policy shocks depends on the covariance of the distribution of MPCs with the cyclicality of income, net nominal position, and unhedged interest rate exposure.

[2]Nearly all theories of MPC heterogeneity have some form of state dependence. For example, in Carroll (1992) the MPC is a declining function of gross household wealth.

[3]For the former, see for instance Kaplan and Violante (2014) and Carroll, Slacalek, Tokuoka, and White (2017). For the latter, Fagereng, Holm, and Natvik (2016) exploit randomized lottery winnings to identify transitory income shocks, and subsequently group observations on observables to estimate group-level MPCs. See also Johnson et al. (2006), Blundell, Pistaferri, and Preston (2008), Parker et al. (2013), Kaplan, Violante, and Weidner (2014), and Crawley and Kuchler (2018).

the funds. Standard methods that rely on sample splitting by observable characteristics can recover only the extent to which the MPC varies with the chosen household characteristics, as opposed to a true distribution, and cannot recover heterogeneity in MPCs associated with latent factors (by definition) or different observables. Indeed, we find that the majority of MPC heterogeneity can be attributed to such latent traits; household observable characteristics explain only a small portion of the variance in MPCs. Second, because our approach does not require taking an *ex ante* stand on what observables correlate with MPC heterogeneity, we can "let the data speak" by investigating *ex post* which observables predict the recovered individual MPCs. Third, we show formally that our approach potentially overcomes the loss of statistical power which appears to affect the sample-splitting approach in existing studies. We find that a household's MPC is correlated with various observable characteristics individually, and that these relationships are generally statistically significant, where previous studies obtained null results in the same data. Finally, by estimating household-level MPCs we are able to project them on various explanatory variables *jointly.* When doing so, we find that the majority of aforementioned significant univariate relationships disappear, leaving non-salary income and the average propensity to consume (APC, average consumption divided by income) as exceptions, both correlating positively with the MPC. By the same token, we can quantify the share of MPC variations explained jointly by observables, finding it lies below a quarter.

Our contribution hinges on the fact that clustering algorithms like the one we adopt assign individuals to groups not based on observable characteristics, but based on how well each set of estimated group-specific parameters describes the observations within the group. This feature allows us to bypass the *ex ante* decision of which observables matter for MPC heterogeneity, and instead estimate the heterogeneity directly first. GMLR specifies a linear regression model with different parameters for each group or "cluster". It is a probabilistic clustering approach, in which individuals are not assigned to groups in a binary fashion, but instead have posterior weights derived from a Gaussian distribution of regression errors. Conditional on these weights, GMLR simply represents a weighted least squares (WLS) regression. When the panel dimension present in "hard clustering" approaches like that in Bonhomme and Manresa (2015) or Bonhomme et al. (2019) is absent, as is the case in our empirical setting, it is unrealistic to think that group assignment can be determined binarily in the presence of noise, so such continuous posterior weights are desirable to represent the level of uncertainty that exists in the assignment. Despite this uncertain assignment, GMLR estimates are consistent and asymptotically normal (e.g., Desarbo and Cron (1988)).

Applying our estimator to study the MPC distribution using the 2008 Economic Stimulus Act, we uncover a substantial degree of heterogeneity. In particular, households spend at least one fourth of the rebate within one quarter, with some households displaying an MPC above one. Generally, the share of households with a particular MPC declines as the MPC increases. At the same time, we estimate sizable MPCs even at the bottom of the distribution. Fagereng et al. (2016) and Olafsson and Pagel (2018) find evidence of similar behavior in Norway and Iceland respectively, and recent papers have proposed models of limited cognitive perception to rationalize such behavior (e.g., Ilut and Valchev (2020)).

We next estimate the distribution of MPCs for specific spending categories. For nondurable expenditure, a lower bound for the MPC of 16 cents on the dollar emerges. On the other hand, spending on durables is dichotomous, with a nontrivial fraction of households not spending any of the rebate on durables, while a significant share spent the entire rebate on such goods. Finally, since our approach provides household level, good-specific MPCs, and we compute their correlation. We find that households with higher nondurable MPCs also display higher durable MPCs, although the correlation is weak (~0.12).

Having characterized the distribution of marginal propensities to consume, we recover its observable drivers. Historically, the literature has found mixed empirical evidence and generally weak relationships between MPCs and observable household characteristics, with the possible exception of liquid wealth.[4] This is likely due to a loss of statistical power when re-estimating the MPC with interactions or sample splits. We formally show that our approach may allow us to sidestep these issues. In practice, we indeed find that our estimated MPCs are significantly correlated, individually, with many observable drivers, despite the fact that we use the same identification strategy and dataset that previously delivered insignificant relationships. For example, we find that homeowners have significantly higher MPCs than renters, and households with a mortgage display even greater marginal propensities to consume than outright homeowners. The correlations hold for all expenditure categories that we consider.

Our estimates for household-level MPCs also allow us to study multivariate relationships without further losses of power. We find that only two observables are robust to the inclusion of additional regressors and positively correlate with MPCs: households' non-salary income and their APC. Kueng (2018) also finds that high-income households have higher MPCs in Alaska Permanent Fund data. We highlight how our result cru-

[4] Parker et al. (2013) find statistically insignificant differences by age, income, and liquid wealth. Broda and Parker (2014) and Fagereng et al. (2016) find significant relationships for the latter.

cially hinges on the non-salary component of income, such as business and financial income. Examining how MPCs vary jointly with income and the APC, we uncover three groups of households. "Poor-savers", with low total income and a low APC, have the lowest MPCs. Households with high total income and a low APC, or vice versa, display intermediate marginal propensities to consume. The greatest MPCs are found among "rich-spenders", who not only have high total income, but also typically spend a large portion of it. This group of households has not received much attention in models of consumption and savings.

Importantly, our best array of observable predictors is able to explain at most one quarter of the variation in estimated MPCs. With the vast majority of heterogeneity unexplained by standard observables, our results suggest that a relevant portion of MPC heterogeneity is driven by latent household traits. For example, heterogeneity in discount rates and/or intertemporal elasticities of substitution (Aguiar, Boar, and Bils (2019)) would deliver heterogeneity in MPCs, and is further supported by the aforementioned significance of APCs in predicting MPCs, as APCs can also be a function of the same unobserved traits. This type of unobserved heterogeneity could never be recovered by simply splitting the sample on observable characteristics and estimating within-subsample homogeneous MPCs, as is typically done in the literature.[5]

Our results have four important policy implications. First, we find that 2008 Economic Stimulus payments increased spending of all households, at least in partial equilibrium. Second, the fact that we uncover considerable MPC heterogeneity suggests that, in principle, aggregate spending could be further increased by targeting fiscal transfers to high MPC households. Our robustly significant correlations suggest that it might be desirable to target relatively higher-income households to maximize the aggregate consumption effects of stimulus checks.[6] Such a strategy may imply a tension between the stimulus and relief/insurance purposes of lump-sum transfers.[7] However, since we find that observable characteristics predict little of the variation in MPCs, it is likely that any attempt at targeting will only exploit a limited share of the overall variation in MPCs, given the information available to fiscal authorities.

[5]This is true unless preference heterogeneity is explicitly elicited in survey questions so that it can be used as an observable control. Using Nielsen panel data, Parker (2017) finds that the MPC out of the tax rebate is indeed strongly correlated with a self-reported measure of impatience.

[6]Stimulus checks were phased out for households whose income was above $150,000 ($75,000 for single filers), implying that higher earners did not receive the rebate. Thus, our findings on the positive correlation of MPCs with total income are limited to households within the income range of stimulus checks recipients.

[7]Shapiro and Slemrod (2009) find that low-income individuals were more likely to use the 2008 rebates to pay off debt. Similar patterns have been observed for the CARES act transfers in 2020, see Armantier et al. (2020). Our analysis focuses on the consumption effects of fiscal transfers but is consistent with a potential distributional trade-off.

We also believe our findings can be used to discipline heterogeneous agent models in three ways. First, our estimated full distribution of MPCs is an agnostic target, regardless of a model's characteristics. Second, individual correlations between MPCs and observable characteristics are crucial objects for many heterogeneous agent models, and we are able to estimate them with statistical precision. Third, we provide researchers with an explicit number for the joint importance of observable and unobservable drivers for the distribution of MPCs.

Our paper is related to an extensive literature estimating the marginal propensity to consume out of transitory income shocks, and a smaller, complementary literature examining how it varies across households. As previously mentioned, the vast majority of existing papers study *observable* drivers of the MPC; we relate our findings to this literature. A burgeoning literature has turned its attention to unobserved household traits and preference heterogeneity. Our findings corroborate the importance of this dimension, recently highlighted by Alan, Browning, and Ejrnaes (2018), Parker (2017), Aguiar et al. (2019), and Gelman (2019).

Our approach allows us to flexibly and non-parametrically combine observed and unobserved MPC heterogeneity. In this respect, Misra and Surico (2014) is closest in spirit to our work.[8] They estimate a quantile regression of consumption responses to the 2008 tax rebate using the same data, and find substantial heterogeneity. However, quantile regression estimates the role of regressors at specific points in the overall conditional distribution of the dependent variable. In Supplement B, we show how this approach is sensitive to the correlation of MPC heterogeneity with other forms of heterogeneity, since other factors may be quantitatively larger drivers of the conditional distribution of consumption than the tax rebate.

The paper proceeds as follows. In Section 2 we describe our empirical strategy based on the 2008 tax rebate. In Section 3, we formulate the problem at hand and describe our clustering approach in detail. We also compare our methodology – recovering unconditional heterogeneity in the MPC and then regressing on observables – to previous approaches, stipulating correlates of the MPC and using them to estimate interacted regressions. Our results are outlined in Section 4, where we provide estimates of the distribution of MPCs for various consumption categories. Section 4.3 discusses observable characteristics that are correlated with the estimated MPCs. Section 4.5 shows the longer-run consumption responses to stimulus checks. Section 5 concludes.

[8]Other papers have used the "reported preference" approach, eliciting MPC heterogeneity directly from responses to survey questions. Recent examples include Sahm, Shapiro, and Slemrod (2010), Jappelli and Pistaferri (2014) and Fuster, Kaplan, and Zafar (2018).

2 Empirical methodology

In order to estimate the marginal propensity to consume, and how it varies across households, we consider an off-the-shelf well-identified quasi-natural experiment: the 2008 Economic Stimulus Act (ESA), as studied by Parker et al. (2013), among others. Between April and July of 2008, $100 billion in tax rebates was sent to approximately 130 million US tax filers.[9] Importantly, the timing of rebate receipt was determined by the last two digits of the recipient's Social Security Number (SSN), making the timing of receipt random. As in Parker et al. (2013), we exploit the randomized timing of the rebate receipt, but instead estimate heterogeneous marginal propensities to consume rather than a homogeneous marginal propensity to consume.

Our data come from the 2008 Consumer Expenditure Survey (CEX), which contains comprehensive and detailed measures of household-level consumption expenditures. The 2008 CEX wave also has supplemental questions on the ESA, including the amount of each stimulus payment received. While CEX expenditures are reported at the quarterly frequency, new households enter the survey at each month, making the frequency of our data monthly. Since we depart from Parker et al. (2013) by allowing for treatment heterogeneity, we present their homogeneous specification first as a useful benchmark, introducing our generalizations thereafter.

2.1 Homogeneous MPC

Parker et al. (2013) consider the following specification:

$$\Delta C_j = \beta' \omega_j + \lambda R_j + \alpha + \epsilon_j, j = 1, \ldots, N, \tag{1}$$

where ΔC_j is the first difference of consumption expenditure for $j = (i, t)$, corresponding to household i in quarter t.[10] ω_j is a set of controls including month dummies aimed at absorbing common time effects such as aggregate shocks, as well as seasonal factors.[11] The independent variable of interest is R_j, which denotes the amount of the tax rebate

[9] We defer to Parker et al. (2013) and Sahm et al. (2010) for an exhaustive discussion of the Economic Stimulus Act.

[10] To maintain consistent notation throughout the paper, we refer to j as the (i, t) combination of household i in quarter t. While we have information on the same households i in different periods t, identification is not obtained by comparing individual responses over time, as in Parker et al. (2013). We do not exploit any limited panel structure, except to construct consumption changes for the left-hand-side variable. We return to this point below.

[11] In Parker et al. (2013), the other controls are age, change in number of adults in the household, and change in the number of children in the household. The controls we will use are the same, but additionally include age squared.

received by each household. λ is then interpreted as the causal effect of the rebate on expenditures, where identification is achieved by comparing expenditure changes of households that received the rebate in a certain period to expenditure changes of households that did not receive the rebate in the same period.[12]

2.2 Heterogeneous MPCs

We depart from the homogeneous specification in Equation (1) and allow for heterogeneity in expenditure responses to the tax rebate across households. In particular, we augment the specification in Parker et al. (2013) as follows:

$$\Delta C_j = \beta' \omega_j + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \left(\lambda_g R_j + \alpha_g\right) + \epsilon_j, j = 1, \ldots, N, \tag{2}$$
$$\forall g = 1, \ldots, G,\ E\left[\epsilon_j \mid \omega_j, R_j, j \in g\right] = 0,$$

where $\mathbf{1}\left[j \in g\right]$ is an indicator that takes a value of 1 if household i in period t belongs to a certain group $g = 1, \ldots, G$. That is, we assume that heterogeneity in responses to the rebate can be summarized with G groups, characterized by the vector of coefficients $\left\{\alpha_g, \lambda_g\right\}$. We include group-specific intercepts, α_g, to correctly interpret λ_g as a marginal propensity to consume. For example, since we do not observe quarterly changes in income, failing to include group-specific level effects may bias MPC estimates due to heterogeneity in income changes unrelated to the tax rebate. We assume that the usual conditional mean independence assumption holds separately within each group; in practice, we instrument for R_j with an indicator for the timing of rebate receipt as in Parker et al. (2013) to make this assumption plausible. Our object of interest is $\lambda = \begin{pmatrix} \lambda_1 & \ldots & \lambda_G \end{pmatrix}'$, which describes MPC heterogeneity, while $\mathbf{1}\left[j \in g\right]$ encodes the group membership of each household. The vector of coefficients, combined with $1\left[j \in g\right]$, gives an approximation of the MPC distribution. In the next section we discuss our clustering methodology to jointly estimate λ and $\mathbf{1}\left[j \in g\right]$.

[12] Kaplan and Violante (2014) discuss why λ may not correctly measure the marginal propensity to consume out of a transitory income shock, but is instead better thought of as a "rebate coefficient". We address these issues in Supplement C.5.

3 A clustering approach to MPC estimation

To estimate group-specific MPCs according to (2), we must somehow assign individuals to groups. Previous papers have grouped individuals based on observable characteristics and estimated MPCs within those groups, but doing so presupposes the determinants of the MPCs *a priori*, and rather than a true distribution of MPCs, simply measures how the MPC changes with the chosen household characteristics. Instead, we do not suppose to know these determinants in advance (there is indeed considerable empirical and theoretical disagreement on this point), but rather aim to investigate correlates of the MPC *ex post*, requiring us to remain agnostic while recovering the MPCs.[13] Moreover, we are interested in recovering the full degree of "latent" MPC heterogeneity, including variation not associated with observables. For these reasons, we group individuals on the basis of their heterogeneous MPCs themselves (and potentially other group-specific parameters). Clustering methods are tailored to this goal.[14]

3.1 Gaussian Mixture Linear Regression

The model (2) can be rewritten more compactly as

$$y_j = \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \psi_g^{G\prime} x_j + \epsilon_j, \tag{3}$$

where $y_j = \Delta C_j$, $x_j = \begin{pmatrix} 1 & R_j & \omega_j{}' \end{pmatrix}'$, and the elements of ψ_g^G corresponding to ω_j are restricted to be constant across g. In this section, we include the G superscript on ψ_g^G to make explicit the dependence of the parameter values on the specified number of groups, but omit it subsequently for compactness. In general, clustering models posit an objective function of the form

$$Q\left(Y, X; \Psi^G\right) = \sum_{i=1}^{N} \sum_{g=1}^{G} w_g\left(y_j, x_j, \Psi^G\right)\left(y_j - \psi_g^{G\prime} x_j\right)^2 / \sigma_g^2, \tag{4}$$

[13] We cover the extensive literature on MPC heterogeneity in Sections 1 and 4.3.

[14] Apart from clustering approaches, quantile regression is used by Misra and Surico (2014) to characterize heterogeneous responses to the 2008 tax rebate. Quantile regression differs from clustering; because quantile regression computes relationships at percentiles of the overall conditional distribution, the estimated MPC distribution depends on the correlation of MPCs with other forms of heterogeneity. If the "ranking" of the conditional distribution is mostly driven by factors other than responsiveness to the rebate (like fixed effects or other covariates), and these factors are uncorrelated with the rebate, heterogeneity of the MPC distribution will be underestimated in the presence of noise. We provide a simple example in Supplement B.

where the vector Ψ^G collects group-specific parameter vectors, ψ_g^G, σ_g^2 are group-specific variances of ϵ_j, y_j is a scalar outcome, x_j is a $k \times 1$ vector of explanatory variables, and $w_g(\cdot)$ are group membership weights that sum to 1. Exchanging the order of summations, minimizing (4), conditional on $w_g(\cdot)$, constitutes G weighted least squares (WLS) problems, with weights $w_g(\cdot)$. When $\sigma_g^2 \equiv \sigma^2$, minimizing jointly over $w_g(\cdot)$ and Ψ^G delivers the "hard K-means" algorithm considered by Bonhomme and Manresa (2015). In this algorithm, each observation has binary weights, assigned with certainty to whichever group minimizes its residual. However, in the cross-sectional setting we consider, and short-panel settings in general, the panel dimension is not long enough to meaningfully diminish the noise in group assignment, so one must treat group membership as probabilistic.[15]

To accommodate probabilistic group membership, we postulate a likelihood for Y, which gives rise to continuous posterior weights, $w_g(\cdot)$. The standard parametric choice is a Gaussian mixture regression or "switching regression" (e.g., Quandt (1972); Desarbo and Cron (1988)). In particular, the probability of observing (x_j, y_j) is given by

$$\Pr\left(y_j, x_j; \Psi^G, \Sigma\right) = \sum_{g=1}^{G} \pi_g \phi\left(y_j; \psi_g^{G'} x_j, \sigma_g^2\right),$$

where π_g is the unconditional probability that any observation belongs to group g, Σ^G collects σ_g^2 across groups, and $\phi\left(y_j; \psi_g^{G'} x_j, \sigma_g^2\right)$ is the p.d.f. evaluated for mean $\psi_g^{G'} x_j$ and variance σ_g^2. For such a model, the complete-data (where group membership is known) likelihood is

$$L\left(Y, X, D; \theta^G\right) = \prod_{j=1}^{N} \prod_{g=1}^{G} \pi_g^{d_{jg}} \phi\left(y_j; \psi_g^{G'} x_j, \sigma_g^2\right)^{d_{jg}},$$

where D collects d_{jg}, binary indicators for whether observation j is a member of group g, and θ collects Ψ^G, Σ, and π_g across groups. Since d_{jg} are latent variables, L cannot be maximized directly. Integrating $\mathcal{L} = \log L$ over d_{jg} (conditional on (y_j, x_j)) yields

$$E_{D|Y,X}\left[\mathcal{L}\left(Y, X, D; \theta\right)\right] = \sum_{j=1}^{N} \sum_{g=1}^{G} \gamma_{jg} \left(\log \pi_g + \log \phi\left(y_j; \psi_g^{G'} x_j, \sigma_g^2\right)\right),$$

[15] Bonhomme et al. (2021) investigate how estimating models like (2) can also serve to discretize and recover continuous forms of unobserved heterogeneity, but their results are also tailored to large$-T$ panel settings.

where

$$\gamma_{jg} = \Pr\left(d_{jg} = 1 \mid y_j, x_j\right) = \frac{\pi_g \phi\left(y_j; \psi_g^{G'} x_j, \sigma_g^2\right)}{\sum_{h=1}^{G} \pi_h \phi\left(y_j; \psi_g^{G'} x_j, \sigma_h^2\right)}.$$

Thus, Ψ^G can be obtained from maximizing

$$\begin{aligned} l_\phi\left(Y, X; \theta^G\right) &= \sum_{j=1}^{N} \sum_{g=1}^{G} \gamma_{jg} \log \phi\left(y_j; \psi_g^{G'} x_j, \sigma_g^2\right) \\ &= \sum_{j=1}^{N} \sum_{g=1}^{G} \gamma_{jg} \left[\log\left(\frac{1}{\sqrt{2\pi\sigma_g^2}}\right) - \frac{\left(y_j - \psi_g^{G'} x_j\right)^2}{2\sigma_g^2}\right]. \end{aligned}$$

Conditional on γ_g and Σ^G, this implies minimizing

$$Q_\phi\left(Y, X; \Psi^G\right) = \sum_{j=1}^{N} \sum_{g=1}^{G} \gamma_{jg} \left(y_j - \psi_g^{G'} x_j\right)^2 / \sigma_g^2,$$

which is exactly (4), but with $w_g\left(y_j, x_j, \Psi^G\right)$ specialized to posterior weights γ_{jg}. In practice, given the dependence of γ_{jg} on $\pi_1, \ldots \pi_g, \Psi^G$, and Σ^G, the model is solved using the Expectation-Maximization (EM) algorithm (Dempster et al. (1977)), where the E-step updates the posterior weights γ_{jg} conditional on a set of parameters and the M-step updates $\pi_1, \ldots \pi_g, \Psi^G$, and Σ^G as in WLS. For a detailed discussion of the GMLR problem and its implementation via the EM algorithm, see Desarbo and Cron (1988) or Jones and McLachlan (1992).

An advantage of the GMLR approach to regression-based clustering is that asymptotic properties of the estimator (consistency and asymptotic normality) follow immediately under regularity conditions from standard maximum likelihood results. This means that analytical inference on Ψ^G (and other parameters) is straightforward. Desarbo and Cron (1988) provide a discussion of these inference results; the more detailed discussion for Gaussian mixture models found in McLachlan and Basford (1988) easily extends to GMLR as well. We use the observed information approach for inference on Ψ^G, with analytical formulas adapted to accommodate the presence of common coefficients across groups in (2).

Gaussian mixture instrumental variables regression In our empirical setting, the value of the rebate an individual receives is potentially endogenous, so we turn to an instrumental variables extension of GMLR, "GMIVR". In particular, we consider a two-stage least squares (TSLS) estimator, $\hat{\theta}^{TSLS}$. Our instrument, denoted z_j, is an indicator for rebate receipt in a given quarter (which is based on the last two digits of an individual's Social Security number), and is independent of individual characteristics as well as group structure by construction.

One possible concern is that heterogeneity is not just present in the second stage, but in the first stage as well. However, the group structure in the first stage may not be the same as that in the second, and forcing the group structure to align could bias estimates in both stages. As a solution, we estimate a homogeneous first stage omitting controls, ω_j. We show in Supplement A.1 that doing so leads to unbiased estimates of λ_g in the second stage, regardless of heterogeneity in the first stage, given standard assumptions for our setting.[16] In particular, we estimate

$$R_j = a + \Pi z_j + u_j, \tag{5}$$

and generate $\tilde{R}_j = a + \Pi z_j$. We define $\tilde{x}_j = \begin{pmatrix} 1 & \tilde{R}_j & \omega_j{}' \end{pmatrix}'$, and estimate (3), replacing x_j with $\tilde{x}_j$, as the second stage.

It is straightforward to adjust inference for $\hat{\theta}^{TSLS}$ to account for the fact that $\tilde{R}_j$ is a generated regressor by augmenting the log-likelihood with a second component corresponding to the first stage and computing the score and observed information based on this augmented object.

Choosing the number of groups In all clustering models, it is necessary to choose G, the number of groups; we use the Bayesian Information Criterion (BIC). In particular, the BIC for a candidate number of groups, $\tilde{G}$, is given by

$$BIC\left(\tilde{G}\right) = k_{\theta^{\tilde{G}}} \log N - 2 l_{\phi}\left(\hat{\theta}^{\tilde{G}}\right),$$

where $k_{\theta^{\tilde{G}}}$ is the number of unique parameters in $\theta^{\tilde{G}}$, and $l_{\phi}\left(\hat{\theta}^{\tilde{G}}\right)$ is the maximized incomplete-data log-likelihood for $\tilde{G}$ groups,

$$l_{\phi}\left(\hat{\theta}^{\tilde{G}}\right) = \sum_{i=1}^{N} \log \sum_{g=1}^{\tilde{G}} \hat{\pi}_g \phi\left(y_j; \hat{\psi}_g^{\tilde{G}\prime} x_j, \hat{\sigma}_g^2\right).$$

[16] Including ω_j in the first stage, as is conventional, leads to bias in λ_g in the second stage if heterogeneity is not fully modeled in the first stage and ω_j is correlated with such heterogeneity. Intuitively, unmodeled heterogeneity in the first stage behaves like a set of omitted variables, which bias the first-stage coefficients on ω_j and thus the predicted rebate values. As a robustness check, we consider alternative specifications with different first stage configurations in Supplement C.2 and show that the impact of the form of the first stage is small in practice.

Under regularity conditions, the BIC is consistent for the true value of G (see, e.g., Celeux et al. (2018)). To complement the BIC, we compare the selected G to that obtained from $K-$fold cross validation, and ensure that the chosen model is compatible with both criteria.

3.2 Power comparison of interacted and clustering regressions

After estimating the unconditional distribution of MPCs using the GMIVR estimator described above, we also examine the correlation between the MPC and observable household characteristics. We do so using regressions of the form

$$\hat{\lambda}_j = c + \mu' F_j + v_j, \tag{6}$$

where $\hat{\lambda}_j = \sum_{g=1}^{G} \gamma_{jg} \lambda_g$ is the posterior-weighted MPC for observation j and F_j is a vector of observables. In contrast to the existing literature, with this approach we first recover latent MPC heterogeneity, and then regress that heterogeneity on observable characteristics (henceforth "direct regression"). Previous papers have instead estimated heterogeneity by observables using interacted regressions with a given household characteristic ("interacted regression"). This method measures how the MPC changes with the chosen observable, as opposed to recovering a true distribution.

Our approach has at least three conceptual advantages: it allows us to recover the full distribution of "latent" MPC heterogeneity, including variation not associated with observables; it does not require us to pre-suppose *ex ante* the determinants of heterogeneity; and it allows us estimate regressions on all observables jointly to better assess which have truly significant relationships. However, our approach may also be desirable in a practical sense due to power concerns.

To illustrate this point, we consider two simple data-generating processes (DGPs) based on (6): one in which the MPC is a continuous function of some scalar observable variable, f_j, and one in which it is a discrete function of f_j. The former is consistent with the posterior-weighted MPCs we use in our regressions on observables, while the latter is consistent with our assumed model with discrete true MPCs. In Supplement A.2, we compute the non-centrality parameters of the asymptotic distributions of the $F-$statistics for the estimates of the association between heterogeneity and f_j under each DGP. In particular, the two estimates are the coefficient on f_j in our direct regression approach, where MPCs (which may or may not contain measurement error) are regressed on f_j, and the interaction coefficient in the interacted regression approach, where ΔC_j is regressed on both the rebate R_j and the rebate R_j interacted with an indicator for whether f_j is above its median value.

In the continuous MPC case, we find that our direct regression approach is always more powerful in the absence of measurement error. In the presence of measurement error, it is more powerful if the observable f_j explains a small share of variation in MPCs, a condition which is consistent with our findings for the R^2 in Section 4.3.

In the discrete MPC case, we find the direct regression approach has a "maximin" power property in the absence of measurement error, with constant power independent of parameter values. In other words, it maximizes the worst case power, and will be more powerful when the correlation of MPCs and observables is small. In the presence of measurement error, it remains the case that direct regression will be more powerful when the effect size is small, or when the measurement error is small relative to the noise-to-signal ratio in the consumption equation. Overall, these two cases present an additional argument in favor of our direct regression approach, whether measurement error is present or not.

The precision of estimates in these regressions is important in our setting for two reasons. First, in the conventional approach that uses interactions observables to measure, imprecise estimates suggest a lack of significant variation in the MPC distribution, which, as we discuss in Section 4.1, our full unconditional distribution belies. Second, if heterogeneity is deemed to exist, it is hard to determine which observable characteristics are truly associated with it, since often none are found to be statistically significant. This means that no clear guidance can be provided either to inform policy, or to discipline and distinguish between consumption models based on their implied correlations.

4 Results

We apply our clustering approach to the 2008 tax rebate. As our baseline specification, we adopt the IV specification, as previously discussed. Similarly, we also drop households that never get the rebate, who may have different characteristics (such as higher income) and thus bias the results. In Supplement C, we show results for a battery of additional specifications, including OLS, alternative numbers of groups, alternative first stage specifications, alternative samples, and more flexible restrictions on coefficients for the controls ω_j.

We find a considerable degree of MPC heterogeneity, the extent of which varies depending on the consumption category considered. We first show the distribution of the MPC for total expenditures and use bootstrapping to show its stability. We then investigate how the MPC distribution changes as we consider nondurable and durable goods as the dependent variables. Importantly, our approach also allows us to directly test whether households display similar propensities for different consumption goods, or instead display differential responses across expenditure types when they receive a transitory income shock such as a tax rebate. Finally, we explore which observable household characteristics are correlated with the estimated marginal propensities to consume, both individually and jointly, and analyze the longer-run spending effects of the 2008 tax rebates.

4.1 The distribution of marginal propensities to consume

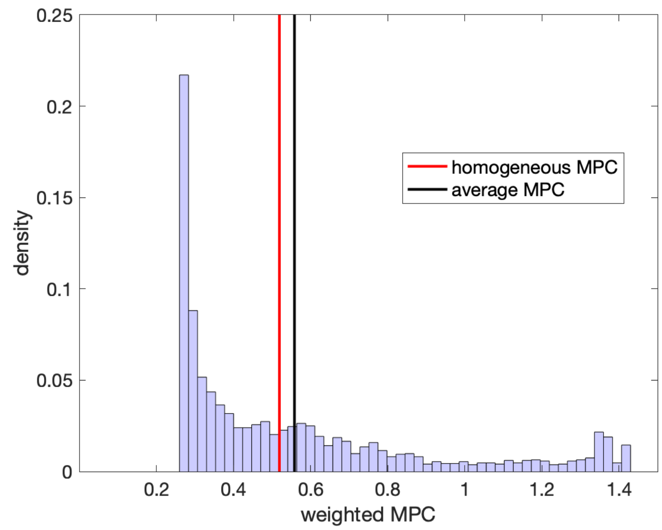
We start by considering total expenditures, defined as in Parker et al. (2013). Following Kaplan and Violante (2014), who show that properly accounting for outliers reduces the homogeneous rebate coefficient, while increasing precision, we drop the top and bottom 1.5% of consumption changes.[17]

In order to choose the number of groups for GMIVR, G, we use the Bayesian Information Criterion (BIC) as discussed in Section 3.1. We find that it flattens between 2 and 5 groups and increases thereafter. Moreover, a 10-fold cross-validation follows a similar pattern. For these reasons, we choose $G = 5$ as our baseline, but show in Supplement C.3 that the unconditional distribution of MPCs, as well as the correlation of household MPCs with household observable characteristics, barely change when considering any alternative number of groups G between 2 and 5. For each household that receives the rebate, we compute the posterior weighted MPC, using the household-specific weights γ_{jg} and the group-specific MPCs λ_g estimated by the GMIVR algorithm. Figure 1 shows the distribution of this object for the $j = (i, t)$ pairs receiving the rebate. The ability to plot this distribution is in itself novel, since previous approaches measure heterogeneity simply as a set of interaction coefficients, rather than a true distribution. Quantile regression also does not permit the recovery of the distribution of any single coefficient, rather recovering coefficients at different quantiles of the conditional distribution of ΔC_j.

We find that the vast majority of households display a relatively low (but certainly non-negligible) MPC of about 27 cents on the dollar, and the share of households with a given MPC slowly decays as the MPC increases. 14% of households consume the rebate in its entirety or even have an MPC above one. Fagereng et al. (2016) and Olafsson and Pagel (2018) find sizable spending responses even for households with high liquid wealth, in Norwegian administrative data and Icelandic application user data, respectively. We

[17]This is the only way in which our sample departs from the sample used in Panel B of Table 3 in Parker et al. (2013), and explains why the homogeneous MPC we estimate for total consumption differs from theirs.

Figure 1: Estimated distribution of MPCs out of the tax rebate



Notes: The histogram (light blue bars) plots the GMIVR-estimated distribution of MPCs for total expenditures among households that received the rebate, defined as in Parker et al. (2013). The sample is defined as in the text. The BIC chooses $G = 5$. For each household we compute the weighted MPC across groups. The black vertical line shows the average weighted MPC in our sample. The homogeneous MPC (red vertical line) is estimated assuming a homogeneous response to the tax rebate as in Parker et al. (2013), following the IV implementation of Equation (1).

likewise find evidence that even the smallest MPCs are substantially larger than zero, even when estimating the full unconditional MPC distribution, in standard U.S. survey data. In contrast, in this same data, Misra and Surico (2014) use quantile regression to estimate a substantial share of MPCs at or *below* zero. We discuss in Supplement B how our approach differs from theirs. A lower bound of the MPC distribution above zero can be explained by bounded rationality. Ilut and Valchev (2020), for instance, develop a model in which MPCs can be high for all households, even those with slack liquidity constraints. Due to limited cognitive perception, households can find themselves in the midst of a "learning trap", "which makes the high MPC behavior the norm, rather than exception."

Aggregating the individual-level responses, we find that the average marginal propensity to consume is similar to the homogeneous specification, as shown by the black and red vertical lines, respectively. Moreover, the former is within the 1-standard deviation confidence bands of the homogeneous MPC. However, this need not be the case, as we discuss in Supplement D. In general, estimates from a homogeneous specification like Equation (1) will equal the average of estimates from a heterogeneous specification only if all regressors x_j are exogenous and have distributions that are invariant across groups, or if the regressor of interest (R_j) is exogenous with constant distribution across groups and is uncorrelated with any other regressor in ω_j whose distribution varies by group.

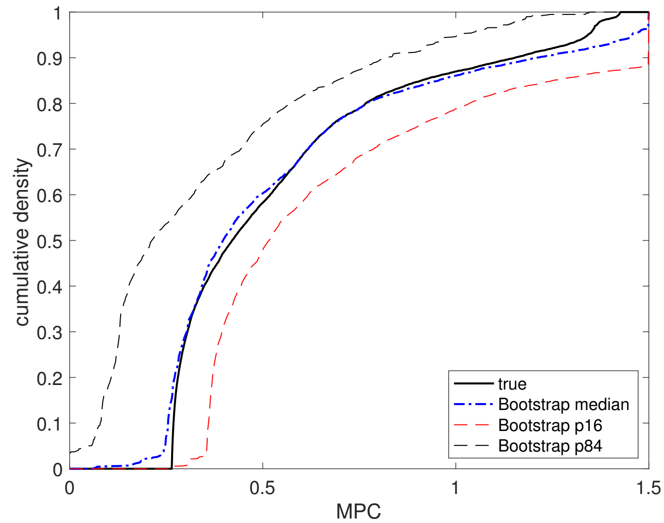
In our setting, while our instrument, a rebate timing indicator, is independent of group membership and household characteristics, it is correlated with the month dummies we include in ω_j, since the rate of rebate disbursement varied over time. These time dummies may be related to group membership due to changes in the aggregate state of the economy. This means that, even in population, there may be some difference between our average heterogeneous MPC and the homogeneous MPC from an equivalent specification.

We now turn to discussing the statistical uncertainty around the estimated distribution. As previously discussed, we implement an observed information approach to compute analytical standard errors for the estimated group-specific MPCs from GMIVR that account for both uncertain group membership and estimation error in the first stage. We present our results in Supplement C.3 and discuss how statistical significance is affected by taking estimation error of the weights into account. We find that the lowest group-specific MPC point estimate, equal to 0.08, is not statistically significant from zero.[18]This is important to bear in mind in light of the previously discussed lower bound. However, the main goal of our analysis is to evaluate the full distribution of individual MPCs. While we cannot formally conduct inference on individual weights – as they are a function of a single realization of a random error, not a parameter – we assess the stability of our findings by bootstrapping. In particular, we repeat the GMIVR estimation of the distribution of MPCs for total expenditures, with 5 groups, over 250 samples obtained by bootstrap with replacement. Figure 2 plots the cumulative density functions. Specifically, the dash-dotted blue line shows the median across bootstraps, which reassuringly tracks the CDF-equivalent of the distribution shown in Figure 1, here depicted in solid black. Moreover, there is reasonably little variation across bootstraps, as evidenced by the centered 68% confidence interval. In particular, more than half of the bootstraps predict that the lowest individual MPC will be above 0, but a good share of the remaining estimated distributions does not. This makes us conclude that, while quantitatively large, the lower bound of MPCs for total expenditures may be subject to uncertainty. We return to this point when discussing the MPC distribution for specific goods in the next section.

These results show that there is indeed considerable variation in the MPC. From a policy perspective, this implies that there is potentially significant benefit to targeting transfers to certain households. For a given dollar value of transfer, those households with a higher MPC will spend more and save less, leading to a greater increase in consumption and stimulatory effect on aggregate demand. We return to the question of whether such targeting is feasible in practice in Section (4.3).

[18]Formally, the lowest $\lambda_g = 0.08$, but $\min_i \sum_{g=1}^{G} \gamma_{i,g}\lambda_g = 0.27$, as shown in Figure 1.

Figure 2: Bootstrapped distribution of MPCs out of the tax rebate: GMIVR



Notes: The black solid line plots CDF of the estimated distribution of MPCs for total expenditures, shown in Figure 1. The blue dash-dotted line shows the median CDF of the estimated distribution of MPCs, across 250 bootstraps with replacement. The dashed red and black lines correspond to the centered 68% confidence interval.

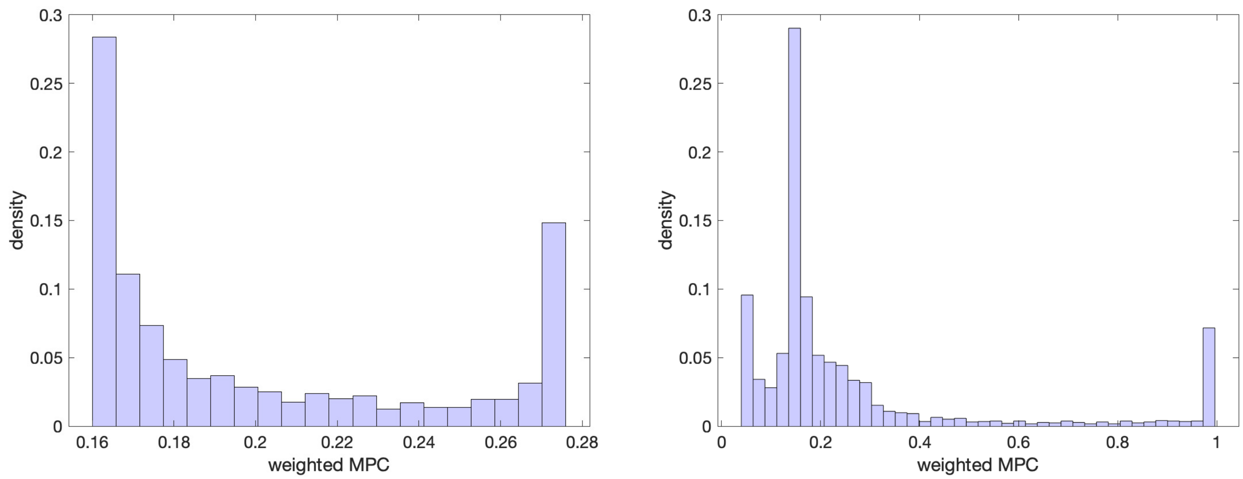
4.2 The MPC distribution for different consumption goods

We have shown how households differ with respect to their propensity to consume the rebate. How does the distribution of these propensities change across consumption goods? The granularity of the CEX data allows us to tackle this question, while our approach allows us to explore how good-specific MPCs vary at the household level.

First, in the left panel of Figure 3, we report the weighted MPC distribution for nondurable goods. As expected, the distribution is shifted to the left with respect to the distribution corresponding to total expenditures in Figure 1, as nondurable goods account for, on average, only 57% of household total expenditures.

An important share of households consumes a small value of nondurables, although at least 16 cents for each dollar of rebate. Strictly speaking, therefore, no household behaves following the Permanent Income Hypothesis (Friedman (1957)). Looking at the other end of the distribution, households consume at most one third of the rebate. While more limited than for total expenditures, the heterogeneity in nondurable MPCs is economically meaningful and statistically significant, as we show in Supplement C.3. Importantly, the lowest group-specific MPC is statistically different from zero at the 5% level. Moreover, the minimum individual weighted MPC is higher than 10 cents in nearly 90% of the bootstraps, and always above 4 cents. Hence, while quantitatively lower than for total expenditures, nondurable MPCs exhibit a lower bound on MPCs which we can confidently place above zero.

Figure 3: MPCs out of the tax rebate: nondurables and durables



(a) Nondurables (b) Durables

Notes**:** The histograms (light blue bars) plot the GMIVR-estimated distribution of MPCs for nondurable and durable expenditures respectively among households that received the rebate, defined as in Parker et al. (2013). The sample is defined as in the text. The homogeneous MPCs estimated with the IV implementation of Equation (1) are 0.19 for nondurables and 0.26 for durables. For each household we compute the weighted MPC across groups. For nondurables the BIC selects $G = 2$ and for durables $G = 4$. Nondurable goods are defined, following Parker et al. (2013), as strictly nondurables (Lusardi (1996)) plus apparel goods and services, health care expenditures (excluding payments by employers or insurers), and reading material (excluding education). As in Parker et al. (2013), we define durable expenditures as the difference between total and nondurable expenditures.

The right panel of Figure 3 shows the estimated MPC distribution for durable goods. About 7% of households do not change their durable expenditures in response to the rebate; their weighted MPC is less than 5 cents to the dollar. Moreover, the lowest group-specific MPC is slightly below zero and bootstrapping confirms that an important share of households does not use the rebate to consume durables, as shown in Supplement C.3. This finding helps reconcile the uncertainty around the lower bound on the MPC for total expenditures, which we discussed previously. The vast majority of households consume around 15 cents to the dollar in durable goods. A non-negligible fraction of households, however, has a durable MPC close to one; 6% of households are approximately hand-to-mouth when it comes to durables. The dichotomy of this MPC distribution is in line with the discrete nature of durable goods purchases.

Finally, we assess whether households with high propensities to consume nondurable goods are also more likely to consume durable goods after receiving the rebate. While we can rule out substitution between goods, the estimated complementarity at the margin is, however, quantitatively small. The correlation between household-level weighted MPCs for nondurable goods with those for durables is 0.12 (significant at the 1% level). Albeit small, the complementarity might signal the presence of heterogeneous preferences or a small share of “spender” types, who are more prone to adjust *any* type of consumption in response to transitory income shocks. While the structure of our data does not allow us to draw conclusions regarding whether the heterogeneity we measure is permanent or transitory, we can investigate what observable characteristics explain the estimated MPC distributions that we recover. We tackle this issue in the next section.

4.3 What drives MPC heterogeneity?

Our approach uncovers the distribution of marginal propensities to consume without taking a stance, *ex ante*, on its observable determinants. Consequently, we can use the estimated distribution to understand how MPCs correlate, *ex post,* with observable characteristics. We start by examining how observables are individually correlated with MPCs. We then turn to investigate the joint relationship between the estimated MPCs and various household characteristics. As such, we contribute to the literature in three ways. First, we show that, with our approach, a large number of statistically significant individual correlations between MPCs and observable drivers emerge. This is true despite the fact that we use a dataset and an identification strategy that previously failed to find statistically significant relationships (e.g., Parker et al. (2013)). Second, we show how the distribution of MPCs is *jointly* correlated with observable characteristics, and can be confident that any lack of significant correlations is not due to loss of statistical power introduced by progressive interactions. In Supplement A.2 we formally assess the power properties of our approach to recovering observable determinants of MPC heterogeneity. Third, we can quantitatively assess the share of MPC heterogeneity that can be explained by observables. This metric is important for assessing the distributional effects of fiscal policy, gauging the potential the government has for targeting payments explicitly, and for disciplining heterogeneous agent models of consumption and savings.

Table 1 reports individual correlations. Our estimated weighted MPCs for total expenditures (column (1)) are positively correlated with salary and non-salary income, the mortgage interest-to-income ratio, the average propensity to consume (APC),[19] and liquid wealth; however, they are negatively correlated with age.[20] Similar relationships hold

[19]Empirically, we define the APC as average lagged consumption divided by average lagged total income. We lag expenditures to avoid the possibility of a mechanical positive correlation with the MPC. To ensure stability of APCs, we average expenditures over all the available lagged quarters at the household-level, but the results are virtually unchanged if we only consider the first lag. We consider income as measured in the first interview for each household, which refers to the previous 12 months. We winsorize the APC upwards at 3, which is about 1.5% of the observations.

[20]Additional relationships hold unconditionally. For instance, we find that households that put money into a tax-deferred or tax-free educational savings plan have a significantly higher MPC. Moreover, MPCs increase with education. All these relationships, however, are insignificant when tested jointly with other observables as in Table 2.

Table 1: Individual correlations with the MPC for total expenditures

	Total	Nondurables	Durables
log salary income	0.09***	0.12***	0.09***
log non-salary income	0.19***	0.20***	0.17***
mortgage interest/income	0.07***	0.05**	0.06***
APC	0.12***	0.04*	0.12***
age	-0.03*	-0.04*	-0.03
log liquid assets	0.14***	0.09***	0.14***

Notes: Table 1 shows the correlations between MPC estimates listed in columns and observables listed in rows. We report results for total expenditures, nondurables and durables. All logged variables take a value of log(0.001) when the raw value is 0 or negative. *, ** and *** denote significance of the correlation at 10, 5 and 1% levels, respectively.

for nondurable (column 2) and durable (column 3) MPCs.

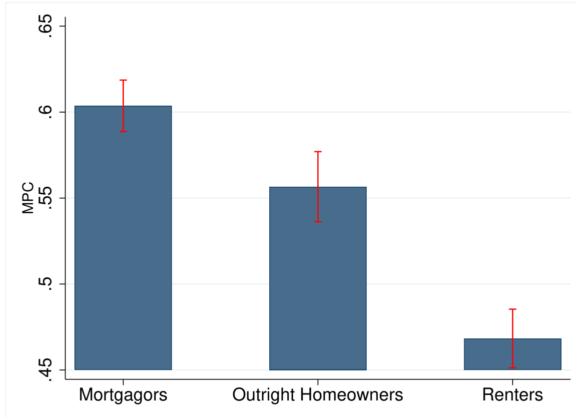
We also find that homeowners have larger MPCs for total expenditures, a result that echoes findings in Parker et al. (2013). Moreover, having a mortgage is associated with an even higher propensity to consume, as shown in Figure 4.

In Table 2 we regress our estimated weighted MPCs on an array of household observable characteristics. Aware of the low number of respondents for liquid wealth, and the potential non-response bias associated with it, we do not include it as an explanatory variable here, but report the associated findings in Supplement C.4. Our results are robust across specifications and even when considering the MPC distribution estimated with a different number of groups, as shown in Supplement C.4. Importantly, only two explanatory variables remain statistically significant after the inclusion of additional covariates: non-salary income and the average propensity to consume, both of which are positively correlated with the marginal propensity to consume. We expand on these two drivers in the remainder of the section.

While higher income households have higher MPCs, it is mainly the non-salary component of income that drives this relationship.[21] This effect is partly the result of a particular category of households, such as entrepreneurs or investors (for example, those with a positive business or financial income), who have a significantly higher MPC. The intensive margin, however, seems to play the most prominent role. The other components of

[21] Income in the CEX is measured in the first interview and relates to income over the prior 12 months. Non-salary income consists of farm and business income, financial income (e.g., income from interest, dividends, pensions and annuities) and all other income except foodstamps (e.g., retirement, supplemental security, unemployment compensation), following the categorization in Coibion et al. (2017).

Figure 4: MPCs by housing status



Notes: The bars display the average weighted MPC for total expenditures by housing status of households receiving a rebate for the baseline specification estimated with GMIVR. Error bars show 90% confidence intervals, computed suing the standard deviation of weighted MPC within each group. A similar relationship holds for both nondurables and durables.

nonsalary income, such as unemployment compensation, retirement, and transfers, are also positively associated with the MPC. We split non-salary income into its business-financial and transfer components and find that both sources of income are positively and significantly correlated with the MPC, even after controlling for all the observables in Table 2. Putting the estimates together, we find that a 100% increase in non-salary income is associated with an increase in the MPC of 19 cents for each dollar of the rebate for total expenditures. Put differently, a 170% increase in non-salary income predicts a 1 standard deviation increase in the MPC. Finally, the positive correlation between income and MPCs does not only hold for total expenditures, but also for nondurable and durable expenditures, as shown in the last two columns of Table 2. In general, the relationship between MPC and observable characteristics, and their statistical significance, are remarkably similar when considering different goods.

Some studies find that low-income households have a higher marginal propensity to spend: see, for instance, Johnson et al. (2006) for the 2001 tax rebate and Jappelli and Pistaferri (2014), with respect to cash on hand, for Italian data on self-reported MPCs. Other studies, however, find mixed results or even the opposite relationship, as we do. While Broda and Parker (2014) find that low-income households had larger propensities to spend in the month of the 2008 rebate receipt than households in the top income tercile, this difference "becomes indistinguishable by the end of the quarter". Misra and Surico (2014) also find that median income is higher at the top of the conditional distribution of consumption changes, which they find to be associated with higher propensities to consume, although the overall relationship is U-shaped. We instead find a monotonic relationship with income, as in Kueng (2018), who studies consumption responses to regular and predetermined payments from the Alaska Permanent Fund. Boutros (2020) finds that households whose 2008 rebate was a smaller fraction of their income – typically higher-income households – had a higher MPC. He explains this finding with a model of bounded intertemporal rationality, in which the smaller the relative size of the payment, the more planning costs dominate the benefits of consumption smoothing. The theory of limited cognitive perception developed by Ilut and Valchev (2020) also delivers rich agents with high MPCs. Shapiro and Slemrod (2009) use data on self-reported propensities to spend the 2008 rebate and show that low-income individuals were more likely to pay off debt. They also find that 21% of households making more than $75,000 of total annual income reported to spend most of the rebate, compared to 18% for households with total income below $20,000. Miranda-Pinto, Murphy, Walsh, and Young (2020) develop a model that can rationalize these findings via time-varying consumption thresholds.

Our findings put non-salary income in the spotlight. The importance of business and financial income for the MPC might suggest the presence of wealthy hand-to-mouth households, as first posited by Kaplan and Violante (2014). However, the importance of the other components of non-salary income, such as retirement income and transfers, coupled with the significance of the APCs discussed below, suggests other mechanisms may also be at play.

Marginal propensities to consume also increase with the average propensity to consume (APC). Households that spent 1 percentage point more of their income before receiving the rebate spent 29 additional cents out of each rebate dollar. This effect is significant also for nondurable and durable MPCs: households that typically spend more relative to their income have a greater MPC.[22]

In Figure 5 we show how the MPC varies jointly with the APC and total income. We separately compute quintiles of the APC and total income, and calculate the average weighted MPC for each quintile pair. The MPC increases with income, conditional on the APC, and vice versa. As the figure shows, our analysis uncovers three main groups. First, households with low total income and a low APC display the lowest marginal propensity to consume. We label these households "poor savers". Second, households with a high APC and low total income, and vice versa, display intermediate MPCs. Third, the greatest marginal propensity to consume is found among households with a high APC and high total income. We label this group "rich spenders".[23]

[22] A 1 percentage point increase in the APC for nondurables predicts 3 additional cents per rebate dollar spent on nondurables. This effect goes up to 7 when considering the APC for nondurable expenditures only.

[23] We find similar relationships for MPC for nondurables and durables, especially the presence of "rich

Table 2: Explanatory variables of weighted MPCs

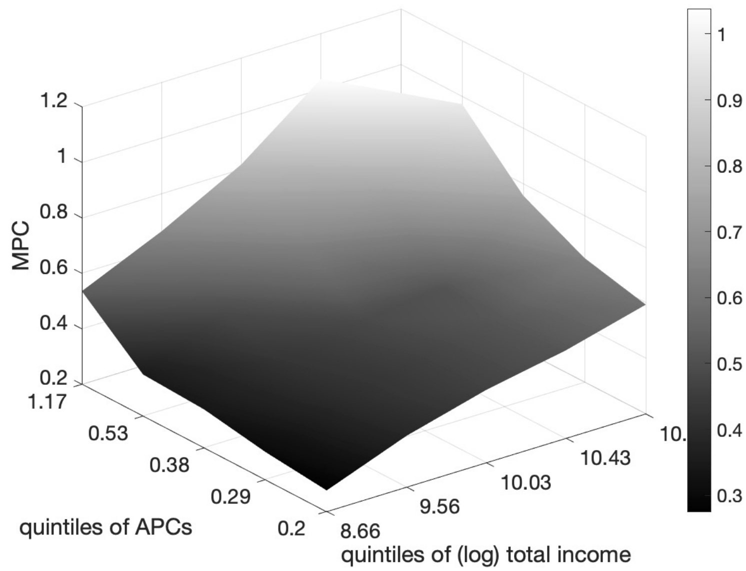
	(1) Total	(2) Nondurables	(3) Durables
dummy for no salary	-0.371	-0.041	-0.080
	(0.233)	(0.029)	(0.176)
log salary income	-0.018	-0.002	-0.000
	(0.015)	(0.002)	(0.011)
log non-salary income	0.192***	0.024***	0.129***
	(0.022)	(0.003)	(0.019)
mortgage interest/income	0.129	0.004	0.105
	(0.113)	(0.015)	(0.100)
APC	0.293***	0.033***	0.210***
	(0.033)	(0.004)	(0.031)
homeowner dummy	0.042	0.007	0.007
	(0.029)	(0.004)	(0.021)
dummy for mortgage	-0.049	-0.003	-0.040
	(0.030)	(0.004)	(0.022)
N	1079	1058	1078
adj. R^2	0.146	0.126	0.112

Notes: All logged variables take a value of log(0.001) when the raw value is 0 or negative. Non-salary income is positive for all observations. Standard errors are robust to heteroskedasticity and reported in parentheses. We control for marriage dummies, education dummies, number of children, age and age squared; those coefficients are not reported. *, **, and *** denote significance at the 10, 5, and 1% levels respectively.

We regard the results presented in this section as particularly relevant for disciplining macro models of household consumption. For example, the relationship between MPC, APC and income can be directly tested in even the simplest of consumption/savings models. Existing models predict very different relationships between MPCs and APCs. Hand-to-mouth, constrained agents, will typically have large MPCs and APCs. As they save towards their target level of wealth, both propensities fall. If agents are infinitely-lived, they eventually reach the target level of wealth, at which they stop saving (i.e. APC = 1) and have an MPC equal to the annuity value of the transitory income shock. A life-cycle model can, in contrast, generate the empirically observed positive relationship between MPC and APC, as older households dissave, but also have a high MPC

spenders", as we show in Supplement C.4.

Figure 5: The relationship between MPCs, APCs, and income



Notes: The plot surface shows the average weighted MPC for total expenditures for pairs of quintiles of APC and log total income. The colorbar on the right represents the MPC.

due to a low effective discount factor. This standard model, however, generates a clear relationship between the MPC and age, which our results do not bear out.[24]

All these characterizations are conditional on homogeneous preferences. Preference heterogeneity, in contrast, can break these relationships and rationalize some of our findings. Aguiar et al. (2019), for instance, highlight the importance of heterogeneity in the intertemporal elasticity of substitution in order to generate heterogeneous target levels of wealth; high-IES households have high MPCs and high APCs. Consistent with this, Parker (2017) finds that the majority of consumption responsiveness to the tax rebate in the Nielsen data is driven by a measure of impatience, defined as households reporting to be "the sort of people who would rather spend their money and enjoy today rather than save more for the future."

An additional finding underscores the importance of unobserved heterogeneity. All the observable drivers mentioned in this section – as well as other household characteristics that do not strongly correlate with the MPC – explain a relatively small portion of the variance of the weighted MPC distribution. Indeed, our linear regression framework of weighted MPCs on observable characteristics delivers an adjusted R^2 of 15%. Such ex-

[24]Moreover, most incomplete markets models typically fail to generate savings rates (APCs) that increase (decrease) with wealth and permanent income, at odds with what is observed in the data and documented by Dynan, Skinner, and Zeldes (2004) and Straub (2017).

planatory power is even lower for nondurables and especially durables. Unlike previous studies, we obtain a statistical measure of the portion of the variance in the MPC distribution explained by observable characteristics through the R^2. Technically, the reported R^2 is a lower bound on the true R^2 due to measurement error in the estimated MPCs. We discuss this issue in detail in Supplement C.6 and propose a back-of-the-envelope adjustment to the R^2 for to account for measurement error in recovering the MPCs in our clustering approach. Such a correction increases the R^2 for total expenditures to 26%, which still indicates that only a quarter of the MPC heterogeneity can be explained by observables.

A low R^2 could also be partly explained by non-linear relationships that are either difficult to parametrize or simply not captured by variables in our dataset.[25] For example, the CEX contains only sparsely populated information on wealth. In Supplement C.4, we show the relationship between the MPC and liquid wealth, aware of the potential nonresponse bias highlighted by Parker et al. (2013), but we refrain from showing any relationship with total wealth, given the lack of reliable data. While such unmeasured characteristics could potentially explain some of the variation in MPCs, our results strongly suggest the presence of latent drivers, since some of those unobserved characteristics may give rise to the observables we analyze in the first place, such as the APC. This finding is not only useful for disciplining heterogeneous agent models, but is also informative about the degree to which fiscal policy can target high MPC households.

As a final exercise, we directly compare our approach to that typically taken in the literature. To do so, we take our estimated posterior weights as a form of (probabilistic) sample splitting and use them to estimate 5 group-specific MPCs. We then compare these results with regression estimates in which we instead split the sample using quintiles of commonly-studied observable characteristics.[26] Table 3 shows the results. In the first column we report our GMIVR MPCs, ordered from low to high. We then report the estimated MPCs across quintiles of age, non-salary income and APC, ordered from the lowest to the highest quintile. The heterogeneity by age is unclear, in line with findings in Table 1. Moreover, if a researcher used only age to characterize the extent of MPC heterogeneity, she would obtain estimates between 28 and 79 cents, much narrower than the range we uncover. Splitting by either non-salary income or the APC, which we show above to be the most robust drivers of MPC heterogeneity, would allow a researcher to

[25]Our results are robust to different sets of explanatory variables. For instance, we consider quantiles of the observables in order to gauge nonlinear effects. We also run a linear Lasso for the selection of the array of predictors. Regressing the MPC on the selected right hand side delivers the same R^2.

[26]Formally, we estimate $\Delta C_j = \beta' \omega_j + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \left(\lambda_g R_j + \alpha_g\right) + \epsilon_j$, in which $\mathbf{1}\left[j \in g\right]$ is defined by quintiles of a certain characteristic such as age.

Table 3: MPC heterogeneity: full vs observable distribution

	GMIVR	age	non-salary income	APC
$g = 1$	0.08	0.38	0.45	0.15
	(0.02)	(0.23)	(0.14)	(0.14)
$g = 2$	0.35	0.65	0.48	0.45
	(0.00)	(0.03)	(0.39)	(0.19)
$g = 3$	0.68	0.28	1.06	0.21
	(0.11)	(0.42)	(0.01)	(0.59)
$g = 4$	1.37	0.79	0.36	1.10
	(0.00)	(0.01)	(0.38)	(0.00)
$g = 5$	1.44	0.53	1.31	0.45
	(0.00)	(0.06)	(0.00)	(0.32)

Notes: Table 3 reports estimated MPCs for different groups, using total expenditures. In parentheses, we report $p-$values from a test of equality to zero. In the first column, we report the results of a weighted least squares taking our estimated GMIVR weights as given, as in panel b of Table 7 in Supplement C.3. In the other columns, we report MPCs obtained by estimating Equation 2, using quintiles of age, non-salary income and APC as $1[j \in g]$ respectively. Groups are ordered from the lowest to the highest MPCs in the first column and by quintile in the other columns. Standard errors in these columns are not adjusted for estimation error in the first stage, since we found that such error had negligible impact on the standard errors in the first column.

uncover some MPCs above 1 but still delivers a lower degree of heterogeneity than using our approach. Therefore, the existing literature, by splitting on observables that are likely noisy in practice, and correlated with only a portion of MPC heterogeneity, would under-estimate the true extent of MPC heterogeneity. Moreover, nearly all the MPCs estimated with our approach are statistically different from zero, while only few are when interacting by observables. In our approach many pairs of MPCs are also statistically different from each other, as we show in Supplement C.3, while virtually no pairs are when interacting by household characteristics. Therefore, these results corroborate earlier statements that our approach may deliver improvements in statistical power.

From a policy perspective, the results in this section have two implications. First, we find that only two observable household observable traits are robustly correlated with the MPC in a statistically significant manner. Among these, our results suggest that fiscal authorities might consider targeting relatively higher-income households as recipients of lump-sum transfers, in the attempt to maximize the effect on aggregate consumption. While we cannot speak to the MPCs of the highest earners absent from our natural experiment, we find that, among the subset of middle-income rebate recipients, higher MPCs are more likely to be found towards the upper end of the (non-salary) income distribution. This implication poses a potential trade-off between the stimulus and relief/insurance effects of lump-sum transfers. This tension is consistent with the empirical finding that low-income households are more likely to use stimulus checks to pay down debt, both in 2008 (Shapiro and Slemrod (2009)) and in 2020 (Armantier et al. (2020)). Second, however, the finding of a small R^2 in the regression of the MPC on observable characteristics suggests that attempts to target transfers based on factors observable by policymakers will ultimately exploit only a small fraction at best of the variation in households' MPCs. This means that feasible targeted transfers can harness only a small share of the gains in terms of consumption response available if policymakers could observe the identity of high MPC households directly or if such MPCs were more strongly associated with observable characteristics.

4.4 Robustness to spurious heterogeneity

In this section we show that the results shown in the previous section are unlikely to be driven by spurious heterogeneity. For this exercise, we generate data using estimates from the homogeneous regression, with errors drawn from a Gaussian distribution with the empirical variance. We then obtain GMIVR estimates under the faulty assumption that more than 1 group is present, and repeat the same analysis for 250 Monte Carlo samples. First, our BIC approach correctly selects $G = 1$ in all but one samples, so that it is very unlikely one would choose a description of the data that allows for heterogeneity. The BIC steadily increases as more groups are added; stronger departures from true homogeneity are penalized more harshly.

Nevertheless, we show in Table 4 what happens if we impose the incorrect degree of heterogeneity on a homogeneous distribution. For small departures ($G = 2$), very limited spurious heterogeneity arises. When fitting a homogeneous distribution with many more groups ($G = 5$), spurious heterogeneity is unsurprisingly more pronounced. However, the average MPC remains still close to the truth.

These spuriously estimated MPCs do not invalidate our headline results regarding observable correlations. To see this, we regress the estimated weighted MPCs for each Monte Carlo sample above on the array of observable predictors used in Table 2. We adopt a conservative approach and show here the results for $G = 5$, in line with our baseline specification, but the results are confirmed when considering $G = 2$, as shown in Supplement C.7.

On average across samples, all the estimated correlations are small and, most importantly, statistically insignificant. Moreover, the adjusted R^2 is 0.1% on average across samples, and never higher than 2%. For illustrative purposes, Figure 6 displays the dis-

Table 4: Over-fitting G: median quantiles of the MPC distribution across simulated samples

	Average	p25	p50	p75
truth:	0.52	0.52	0.52	0.52
$G = 1$	0.50	–	–	–
	(0.18)			
$G = 2$	0.52	0.36	0.50	0.66
	(0.18)	(0.45)	(0.20)	(0.47)
$G = 5$	0.52	-0.85	0.52	1.88
	(0.18)	(1.09)	(0.48)	(1.09)

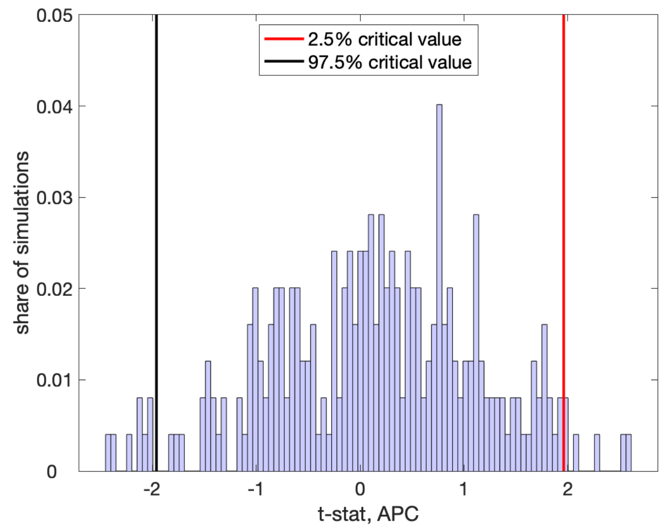
Notes: Each row reports the median of various summary statistics of the distribution of (weighted) MPCs across Monte Carlo samples for models estimated imposing different numbers of groups on data generated from a homogeneous DGP with Gaussian errors. The first row reports the truth, which is 0.52 for all statistics, since the distribution is homogeneous. The second row corresponds to a correctly-specified homogeneous regression in repeated samples (with the standard deviation across samples below in parentheses) and the third and fourth to GMIVR incorrectly assuming the presence of two and five distinct groups, respectively. pxx denotes the xx^{th} percentile. The values in parentheses report the standard deviation of each moment for the $G = 2$ and $G = 5$ specifications across simulated samples.

tribution of the $t-$statistic for the coefficient on the APC, across samples. In only 5.6% of the samples is there significant evidence of a relationship between MPC heterogeneity and APC at the 5% level, a size distortion within the scope of Monte Carlo error. The same is also true for coefficients on all other observables, with even lower shares of significant coefficients. This finding that tests for the significance of correlations with observables achieve close to nominal size when no heterogeneity is present increases our confidence that the rejections we obtain in the data are not due to spurious heterogeneity. Further, a common concern when using the CEX data is the role of measurement error. These exercises also serve to show that grouping on noise alone – like measurement error – does not dictate a distribution like that we recover or the correlations with observable characteristics that we estimate.

4.5 The longer-run effects of the 2008 ESA

In this section, we estimate household-level longer-run spending effects of the 2008 tax rebates, considering a lagged specification that takes the possible persistent effects of rebate receipt into account, as in Parker et al. (2013). In particular, we estimate the following model:

Figure 6: Minimal correlation of spurious heterogeneity with observables



Notes: For each of 250 simulated samples, we regress the weighted MPCs from our baseline specification for total expenditure estimated imposing spurious heterogeneity ($G = 5$) on the set of observables used in Table 2. The histogram (light blue bars) plots the $t-$statistics for the coefficient on the APC across samples. The red and black lines represent the critical values for a 5% test of equality with zero.

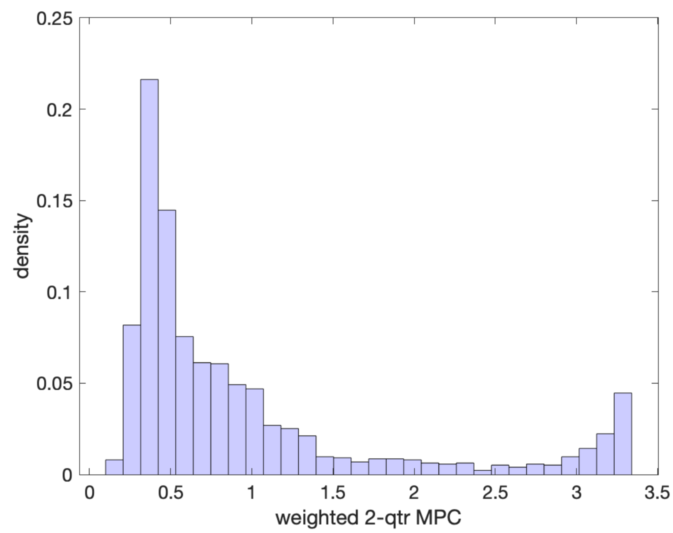
$$\Delta C_j = \beta' \omega_j + \sum_{g \in G} \left(\lambda_g \mathbf{1}\left[j \in g\right] R_j + \lambda_g^{\text{lag}} \mathbf{1}\left[j \in g\right] R_j^{\text{lag}} + \alpha_g \mathbf{1}\left[j \in g\right] \right) + \epsilon_j, j = 1, \ldots, N, \quad (7)$$

where the coefficient λ_g^{lag} represents the lagged effect of the rebate for group g.[27] In line with our baseline specification, we instrument both R_j and R_j^{lag} by an indicator for whether the rebate was received by household $j = (i, t)$, and another indicator for the receipt of the rebate in the previous period. We do not force group membership for household i to be fixed across t, since we want to preserve flexibility; even if individuals' preferences are constant, the MPC may be time-varying, due, for instance, to changes in state variables such as income and wealth.[28] To correctly estimate the cumulative consumption response to the rebate, we therefore track individual weights over the two quarters following the rebate. We use these to construct the individual 2-quarter total effect of the rebate, by adding twice the weighted contemporaneous rebate coefficient to the weighted

[27] Kaplan and Violante (2014) suggest that the rebate coefficient might differ from the marginal propensity to consume because some households in the control group have already received the rebate, and some households might anticipate receiving the rebate in the future. Adding lagged rebate partially address this concern. See Supplement C.5 for further discussion of this specification.

[28] Even in the homogenous case, λ can be different from λ^{lag} because they measure two different objects; the coefficient on the lagged rebate value is an inter-temporal MPC which can be different from the contemporaneous MPC. See Auclert et al. (2018) for a theoretical discussion of intertemporal MPCs.

Figure 7: Estimated distribution of total 2-quarter effect of the tax rebate



Notes: The histogram (light blue bars) plots the estimated distribution of the total effect of the 2008 ESA for total expenditures, defined as in Parker et al. (2013), using the lagged specification in Equation (7). For each household we compute the weighted contemporaneous and lagged MPCs across groups and plot the total response as discussed in the text. The BIC selects 3 groups.

lagged coefficient.[29]

Figure 7 plots a histogram of this object among those who received the rebate. Relative to the baseline results depicted in Figure 1, the distribution spreads out, with some households having a total effect near zero but with most cumulated effects being larger than responses within the quarter. Similarly to the findings for a homogeneous specification in Parker et al. (2013), our λ_g, and especially the corresponding contemporaneous weighted MPCs, are barely affected by controlling for lagged rebate receipt. Moreover, all individual lagged responses ($\sum_{g=1}^{G} \gamma_{i,g}\lambda_g^{\text{lag}}$) are negative, suggesting that in the second period households consume a smaller fraction of the rebate than in the first (since a value of zero indicates a constant consumption response). 95% of the households, however, still displays a positive net effect in the second period. Therefore, as documented by Parker et al. (2013), we show that spending does not only increase upon receipt of the rebate, but also remains high but lower in the subsequent 3 months. We complement this finding by showing that such behavior is qualitatively widespread across households, but is quantitatively quite heterogenous.

Finally, we show in Table 5 that our previous analysis regarding the drivers of MPC

[29]For example, a household may be categorized to be in some group a in the period in which they receive the rebate, and then in some group b the period after they receive the rebate. For such an individual, we construct the individual 2-quarter total effect of the rebate by adding twice the contemporaneous rebate coefficient for group a to the lagged rebate coefficient of group b, since λ_g^{lag} captures the change in consumption *relative* to consumption in the period of rebate receipt.

Table 5: Explanatory variables of 2-quarter MPCs

	(1) 1-qtr MPC	(2) 2-qtr MPC
log non-salary income	0.192***	0.503***
	(0.022)	(0.083)
APC	0.293***	0.857***
	(0.033)	(0.138)
N	1079	535
adj. R^2	0.146	0.148

Notes: See notes for Table 2. MPCs for total expenditures. Column 1 repeats the same analysis of Table 2, whereas column 2 uses as dependent variable the 2-quarter MPC computed as described in the text. Regression results for column 1 are unaffected if restricting to the subsample of column 2, in which we consider households observed for at least one full quarter after rebate receipt. Standard errors are robust to heteroskedasticity and reported in parentheses. *, **, and *** denote significance at the 10, 5, and 1% levels respectively.

heterogeneity is confirmed when looking at longer-run spending responses. Non-salary income and the APC remain the only two variables that are significantly correlated with the longer-run marginal propensities to consume. These results are in line with the fact that the spending effects of the rebate are persistent for most households. Moreover, it confirms that the relationship between MPC and its drivers is not the result of short-lived effects that could be erased by inter-temporal substitution. In addition, the R^2 remains low, suggesting that unobserved heterogeneity is important even at this longer horizon.

5 Conclusion

We exploit a flexible clustering method to uncover the unconditional distribution of the marginal propensity to consume. Our strategy improves on existing approaches by recovering the full distribution of MPCs and not simply estimating how the MPC varies with observable characteristics. Applying this methodology to consumption data following the 2008 Economic Stimulus Payments, households display a considerable degree of heterogeneity in their MPCs. A non-negligible share of households spent the checks in their entirety, and all households spent at least one fourth of the rebate within one quarter, although this lower bound appears subject to statistical uncertainty. Nondurable consumption is also characterized by a lower bound that is significantly larger than zero, while durable consumption features two distinct groups with MPCs close to zero and one.

We then examine which observables – individually and jointly – best predict different portions of the MPC distribution. We obtain various statistically significant relationships with the MPC, but only those associated with non-salary income and the APC survive the inclusion of additional drivers. These results hinge on the fact that our approach proves statistically more powerful than existing methodologies. Moreover, we estimate a formal metric for the share of the unconditional MPC heterogeneity that can be explained by observables. Since observable characteristics explain a minor portion of the estimated MPC heterogeneity, we posit that other latent factors, such as preference heterogeneity, might be important in determining marginal propensities to consume. Taken together our results provide a range of facts useful to discipline an emerging literature of macroeconomic models as well as significant policy implications, particularly for the targeting of transfers.

Finally, two caveats help to highlight possible avenues for future work. Importantly, we measure the distribution of MPCs out of the 2008 tax rebate. This means our estimated distribution uses a single cross-section of data during a recession; if an individual's MPC is a function of the aggregate state, extrapolating our estimates requires caution. Second, because our empirical setting is one in which individuals only experience positive transitory shocks, we cannot speak to income windfalls, to which households may respond differently (Fuster et al. (2018)). However, clustering approaches like the one we use can easily be applied to other datasets with suitably identified transitory income shocks, making comparisons straightforward. We leave such exercises for future work.

References

Aguiar, M., C. Boar, and M. Bils (2019): "Who Are the Hand-to-Mouth?" in *2019 Meeting Papers*, Society for Economic Dynamics, 525.

Alan, S., M. Browning, and M. Ejrnaes (2018): "Income and Consumption: A Micro Semistructural Analysis with Pervasive Heterogeneity," *Journal of Political Economy*, 126, 1827–1864.

Armantier, O., L. Goldman, G. Koşar, J. Lu, R. Pomerantz, W. Van der Klaauw, et al. (2020): "How Have Households Used Their Stimulus Payments and How Would They Spend the Next?" Tech. rep., Federal Reserve Bank of New York.

Auclert, A. (2019): "Monetary Policy and the Redistribution Channel," *American Economic Review*, 109, 2333–67.

Auclert, A., M. Rognlie, and L. Straub (2018): "The Intertemporal Keynesian Cross," Working Paper 25020, National Bureau of Economic Research.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *American Economic Review*, 98, 1887–1921.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–739.

——— (2021): "Discretizing Unobserved Heterogeneity," .

Bonhomme, S. and E. Manresa (2015): "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 83, 1147–1184.

Boutros, M. (2020): "bounded Intertemporal rationality and the Marginal Propensity to Consume," Tech. rep.

Broda, C. and J. A. Parker (2014): "The economic stimulus payments of 2008 and the aggregate demand for consumption," *Journal of Monetary Economics*, 68, S20–S36.

Carroll, C., J. Slacalek, K. Tokuoka, and M. N. White (2017): "The Distribution of Wealth and the Marginal Propensity to Consume," *Quantitative Economics*, 8, 977–1020.

Carroll, C. D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 23, 61–156.

Celeux, G., S. Fruewirth-Schnatter, and C. Robert (2018): "Model Selection for Mixture Models - Perspectives and Strategies," *arXiv: Methodology*.

Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia (2017): "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 88, 70–89.

Crawley, E. and A. Kuchler (2018): "Consumption Heterogeneity: Micro Drivers and Macro Implications," *Danish National Bank Working Paper 129*.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977): "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, 39, 1–38.

Desarbo, W. and W. Cron (1988): "A Conditional Mixture Maximum Likelihood Methodology for Clusterwise Linear Regression," *Journal of Classification*, 5, 249–282.

Dynan, K. E., J. Skinner, and S. P. Zeldes (2004): "Do the Rich Save More?" *Journal of Political Economy*, 112, 397–444.

Fagereng, A., M. B. Holm, and G. J. Natvik (2016): "MPC Heterogeneity and Household Balance Sheets," Discussion Papers 852, Statistics Norway, Research Department.

FRIEDMAN, M. (1957): *A Theory of the Consumption Function*, Princeton University Press.

FUSTER, A., G. KAPLAN, AND B. ZAFAR (2018): "What Would You Do With $500? Spending Responses to Gains, Losses, News, and Loans," Staff Reports 843, Federal Reserve Bank of New York.

GELMAN, M. (2019): "The Self-Constrained Hand to Mouth," .

ILUT, C. L. AND R. VALCHEV (2020): "Economic agents as imperfect problem solvers," Tech. rep., National Bureau of Economic Research.

JAPPELLI, T. AND L. PISTAFERRI (2014): "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, 6, 107–36.

JOHNSON, D. S., J. A. PARKER, AND N. S. SOULELES (2006): "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96, 1589–1610.

JONES, P. AND G. J. MCLACHLAN (1992): "Fitting Finite Mixture Models in a Regression Context," *Australian Journal of Statistics*, 34, 233–240.

KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

KAPLAN, G. AND G. L. VIOLANTE (2014): "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82, 1199–1239.

KAPLAN, G., G. L. VIOLANTE, AND J. WEIDNER (2014): "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 45, 77–153.

KUENG, L. (2018): "Excess sensitivity of high-income consumers," *The Quarterly Journal of Economics*, 133, 1693–1751.

LUSARDI, A. (1996): "Permanent Income, Current Income, and Consumption: Evidence from Two Panel Data Sets," *Journal of Business & Economic Statistics*, 14, 81–90.

MCLACHLAN, G. J. AND K. E. BASFORD (1988): *Mixture models. Inference and applications to clustering*.

MIRANDA-PINTO, J., D. MURPHY, K. J. WALSH, AND E. R. YOUNG (2020): "A model of expenditure shocks," .

MISRA, K. AND P. SURICO (2014): "Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs," *American Economic Journal: Macroeconomics*, 6, 84–106.

OLAFSSON, A. AND M. PAGEL (2018): "The liquid hand-to-mouth: Evidence from personal finance management software," *The Review of Financial Studies*, 31, 4398–4446.

PARKER, J. A. (2017): "Why Don't Households Smooth Consumption? Evidence from a $25 Million Experiment," *American Economic Journal: Macroeconomics*, 9, 153–83.

PARKER, J. A., N. S. SOULELES, D. S. JOHNSON, AND R. MCCLELLAND (2013): "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review*, 103, 2530–53.

QUANDT, R. E. (1972): "A New Approach to Estimating Switching Regressions," *Journal of the American Statistical Association*, 67, 306–310.

SAHM, C. R., M. D. SHAPIRO, AND J. SLEMROD (2010): "Household Response to the 2008 Tax Rebate: Survey Evidence and Aggregate Implications," in *Tax Policy and the Economy, Volume 24*, National Bureau of Economic Research, Inc, NBER Chapters, 69–110.

SHAPIRO, M. D. AND J. SLEMROD (2009): "Did the 2008 Tax Rebates Stimulate Spending?" *American Economic Review*, 99, 374–79.

STRAUB, L. (2017): "Consumption, Savings, and the Distribution of Permanent Income," Working Paper.

Supplemental Appendix to "Latent Heterogeneity in the Marginal Propensity to Consume"

Daniel Lewis Davide Melcangi Laura Pilossoph

Abstract

This supplement contains additional material for the paper "Latent Heterogeneity in the Marginal Propensity to Consume". In Section A, we discuss our specification choice, and relate it to possible alternatives. In particular, we detail the homogeneous first stage regression and provide a power comparison of the direct regression and interacted regression approaches. Section B presents an illustrative example outlining potential pitfalls of quantile regression. Section C contains additional empirical results, exploring the role of heterogeneous coefficients on controls and alternative first stage specifications, presenting additional results on the MPC distribution and the correlation of MPCs and observables, examining the effect of controlling for lags, describing an adjustment to the R^2 to counter measurement error, and discussing further results discounting the possibility of spurious heterogeneity. Section D explains why homogeneous and average heterogeneous effects may differ in our setting, even though our instrument for treatment is independent of household characteristics.

A Specification details

A.1 Homogeneous first stage

In this section, we show that even if there is heterogeneity in the intercept in the first stage (5), second stage estimates of λ_g will be unbiased when controls, ω_j, are omitted from the first stage, as in our baseline specification. This follows from the fact that the instrument z_j is independent from individual characteristics, including latent group structure and characteristics that determine rebate value.

Suppose that the true first stage takes the form

$$R_j = \sum_{g=1}^{G} \mathbf{1}\left[i \in g\right] a_g + \Pi z_j + \kappa' \omega_j + u_j^h, j = 1, \ldots, N, \tag{8}$$

where intercepts a_g are group-specific. Note that we do not consider heterogeneity in κ, the first stage coefficients on ω_j, since our baseline model does not allow for such heterogeneity in the second stage. For the majority of the controls we include in ω_j, which are time dummies and dummies for changes in household size, this makes sense, since they are not related to the previous year's tax return, which determines rebate value. Π, the coefficient on the instrument, cannot be heterogeneous, since it cannot be correlated with group membership, since that would violate random assignment (in fact, we check *ex post* if the distribution of z_j varies with group membership and find that it does not). Our vector of controls, ω_j, contains two types of variables: potentially endogenous household characteristics (which may be correlated with unobserved first-stage heterogeneity) and exogenous time dummies for the month in which the rebate is received (which are uncorrelated with unobserved first-stage heterogeneity, by construction). The first-stage coefficients on the time dummies should be zero since rebate timing is exogenous, and thus not correlated with the determinants of rebate value. Nevertheless, the time dummies are correlated with the rebate receipt instrument, z_j, in the data because the rate at which rebates were issued varied over time. Conventionally, if a control is correlated with the instrument and appears in the second stage, it must also be included in the first stage, or else the second stage coefficient of interest will be biased. However, because in population the first stage coefficients on the time dummies should be zero, we impose that restriction in the first stage, omitting them, since in this special case doing so does not introduce bias.[30] For this reason, we limit our consideration below to the subvector of potentially endogenous controls, denoted ω_j^e, which are uncorrelated with z_j, and corresponding coefficients, κ^e, abstracting from the presence of additional exogenous controls. We first argue that estimating a homogenous first stage, with ω_j^e omitted, does not bias λ_g. For tractability, our argument focuses on the case where group assignment is known *ex ante*. Consider a homogenous first stage,

$$R_j = a + \Pi z_j + u_j, j = 1, \ldots, N, \tag{9}$$

[30] As a robustness check (see Supplement C.2), we include time dummies in the first stage and find that the majority of first-stage coefficients are quite small and not statistically different from zero, while the MPC distribution is largely unchanged.

where $u_j = u_j^h + \left(\sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] a_g - a\right) + \kappa^{e\prime}\omega_j^e$, which is independent of z_j by construction. Thus, the coefficient on z_j recovered from (9) is identical to Π in (8). Then the predicted rebate is given by

$$\tilde{R}_j = \Pi z_j.$$

In the second stage, we have

$$\begin{aligned}
\Delta C_j &= \beta^{e\prime}\omega_j^e + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right]\left(\alpha_g + \lambda_g R_i\right) + \epsilon_j \\
&= \beta^{e\prime}\omega_j^e + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right]\left(\alpha_g + \lambda_g\left(\tilde{R}_j + u_j\right)\right) + \epsilon_j \\
&= \left(\beta^{e\prime} + \kappa^{e\prime}\sum_{g=1}^{G} \mathbf{1}\left[j \in g\right]\lambda_g\right)\omega_j^e + \sum_{g=1}^{G}\left[\alpha_g + \lambda_g a_g\right]\mathbf{1}\left[j \in g\right] && (10) \\
&+ \sum_{g=1}^{G} \lambda_g \mathbf{1}\left[j \in g\right]\tilde{R}_i + \left(\epsilon_j + \sum_{g=1}^{G}\lambda_g \mathbf{1}\left[j \in g\right] u_j^h\right). && (11)
\end{aligned}$$

Since $\mathbf{1}\left[j \in g\right]\tilde{R}_j = \mathbf{1}\left[j \in g\right]\Pi z_j$ is independent of the other regressors as well as the errors, λ_g is recovered via two-stage least squares, where the second stage includes group dummies, predicted rebate interacted with group dummies, and controls. If ω_j^e is correlated with group membership, then the coefficients on ω_j^e and group dummies will be altered relative to those in (10), but λ_g will be unaffected. While the preceding development assumes that group structure in the first and second stages aligns, the argument holds even if the group structure differs; under such a structure, the form of both coefficients on group dummies and the errors change, but the coefficients on $\mathbf{1}\left[j \in g\right]\tilde{R}_j$ – and its independence of other regressors and errors – are unchanged.

Now suppose instead that the econometrician estimates a homogeneous first stage, but includes controls ω_j^e, as is conventional:

$$R_j = a + \Pi z_j + \tau^{e\prime}\omega_j^e + u_j. \tag{12}$$

First, note that estimating (9) still recovers Π in population, since z_j is independent of ω_j^e and any unmodeled group heterogeneity in the intercept. τ^e will be distinct from κ^e if group membership, and thus a_g, is correlated with ω_j^e. Thus, the predicted values from the first stage, $\tilde{R}_j$, are given by

$$\tilde{R}_j = a + \Pi z_j + \tau^{e\prime}\omega_j^e.$$

In this way, the first-stage coefficient on ω_j^e is biased (from κ^e to τ^e) by the omission of group heterogeneity (in the intercept) in (12). Turning to the second stage, we apply Frisch-Waugh-Lovell to more compactly characterize coefficients on ω_j^e and $\tilde{R}_j$. In particular, the residuals from regressing ω_j^e on group dummies are

$$\hat{\omega}_j^e = \omega_j^e - \sum_{g=1}^{G} E\left[\omega_j^e \mid j \in g\right]$$

and those from regressing $\mathbf{1}\left[j \in g\right] \tilde{R}_j$ on group dummies separately for each g are given by

$$\hat{\tilde{R}}_{jg} = \mathbf{1}\left[j \in g\right]\left[\Pi \hat{z}_j + \tau^{e\prime} \hat{\omega}_j^e\right],$$

where $\hat{z}_j = z_j - \sum_{g=1}^{G} E\left[z_j \mid j \in g\right] \equiv z_j - E\left[z_j\right]$. Define $\hat{\tilde{R}}_j = \begin{pmatrix} \hat{\tilde{R}}_{j1} & \dots & \hat{\tilde{R}}_{jG} \end{pmatrix}'$. Regressing $\Delta C_j = \beta^{e\prime}\omega_j^e + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right]\left(\alpha_g + \lambda_g R_j\right) + \epsilon_j$ on group dummies yields the residuals

$$\begin{aligned}
\widehat{\Delta C}_j &= \beta^{e\prime}\hat{\omega}_j^e + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \lambda_g \hat{R}_j + \epsilon_j \\
&= \beta^{e\prime}\hat{\omega}_j^e + \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \lambda_g \left(\Pi \hat{z}_j + \kappa^{e\prime}\hat{\omega}_j^e + u_j^h\right) + \epsilon_j, \\
&= \left(\beta^{e\prime} + \left(\kappa^e - \tau^e\right)' \sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \lambda_g\right) \hat{\omega}_j^e + \sum_{g=1}^{G} \lambda_g \hat{\tilde{R}}_{jg} + \left(\sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \lambda_g u_j^h + \epsilon_j\right),
\end{aligned}$$

using (8) and rearranging. However, the coefficient above on $\hat{\omega}_j^e$ varies by group provided $\kappa^e \neq \tau^e$, in other words, as long as omitting heterogeneity in the first stage biases the first stage coefficient on ω_j^e. Recovering a homogeneous coefficient, $\tilde{\beta}^e$, in the second stage will result in a residual term of the form

$$\sum_{g=1}^{G} \mathbf{1}\left[j \in g\right] \left(\tilde{\beta}^e - \beta_g^e\right)' \hat{\omega}_j^e,$$

which is correlated with each $\hat{\tilde{R}}_{jg}$, since $\mathbf{1}\left[j \in g\right] \tau^{e\prime} \hat{\omega}_j^e$ appears in $\hat{\tilde{R}}_{jg}$. Thus, the heterogeneity induced in the second stage coefficient on $\hat{\omega}_j^e$ due to misspecification of the first stage biases estimates of λ_g, the parameters of interest, even if z_j is a valid instrument. Note that when the group structure in first and second stages aligns this bias could be eliminated by allowing for group-specific coefficients on ω_j^e in the second stage; however, if the group structure in first and second stages differs, this need not be the case.

For the reasons outlined above, we estimate a first stage that omits controls ω_j^e, since in our setting this does not induce bias in the second stage, given the independence of z_j from ω_j^e (as well as any group structure). Including the controls is not innocuous, and would likely lead to second-stage bias. We opt not to model heterogeneity in the first stage out of concern that such heterogeneity need not coincide with the group structure in the second, and could thus confound our clustering approach. Moreover, we opt not to pursue a model with a homogeneous first stage and full heterogeneity in the second stage (including coefficients on controls) for reasons of computational tractability and because such a modification need not address the possible bias if unobserved heterogeneity is not aligned across first and second stages. However, we show robustness of our results to such modifications in Supplement C.1.

A.2 Power comparison of direct regression and interaction approaches

In this section, we assess the power properties of both the direct regression and interacted regression approaches to recovering observable determinants of MPC heterogeneity. We describe two simple DGPs and calculate the non-centrality parameters of the asymptotic distributions of the $F-$statistics testing the association of heterogeneity with an observable for each approach. We find that, in the absence of measurement error in the MPCs, our direct regression approach has superior power, or a "maximin" power property. In the presence of measurement error, we show that the direct approach will be more powerful when the observable variable explains a small amount of overall MPC heterogeneity (as we find empirically) or the size of the effect is quite small (as has often been found in the literature).

Case 1: MPCs as a continuous function of observables

We first consider a model where the MPCs are a continuous function of some scalar observable characteristic, f_j. In particular,

$$\lambda_j = a + \delta f_j + v_j.$$

Note that while our empirical approach is based on a model of G discrete MPCs, the individual-specific weighted MPCs we use in the regression on observables are indeed continuous.

Under the direct regression approach, regressing λ_j on f_j (and a constant) recovers $\delta_c^{dir} = \delta$, with $\text{avar}\left(\sqrt{N}\hat{\delta}_c^{dir}\right) = \frac{\sigma_v^2}{\sigma_f^2}$. In the presence of classical measurement error, meaning that $\hat{\lambda}_j = \lambda_j + e_j$ is used in the regression on observables, we have instead

$\text{avar}\left(\sqrt{N}\hat{\delta}_{c,e}^{dir}\right) = \frac{\sigma_v^2+\sigma_e^2}{\sigma_f^2}$.

We next consider an interacted regression using a dummy variable, I_j, indicating whether f_j is above or below its median value. We assume that f_j follows a normal distribution with mean zero and variance σ_f^2. We assume that the consumption DGP has the form

$$\Delta C_j = \lambda_j R_j + \eta_j,$$

and the estimated interacted regression is

$$\Delta C_j = \Lambda_c R_j + \xi_c R_j I_j + \tilde{\eta}_j,$$

where $R_j \sim N\left(0, \sigma_R^2\right)$ for simplicity. It follows that

$$\begin{aligned} \xi_c &= 2\delta E\left[f_j \mid I_j = 1\right] \\ &= \delta\sqrt{\frac{8\sigma_f^2}{\pi}}, \end{aligned}$$

and

$$\text{avar}\left(\sqrt{N}\hat{\xi}_c\right) = 4\left(3\sigma_v^2 + \sigma_\eta^2/\sigma_R^2\right),$$

which imply that

$$\text{avar}\left(\sqrt{N}\hat{\delta}_c^{int}\right) = \frac{3\sigma_v^2 + \sigma_\eta^2/\sigma_R^2}{2\sigma_f^2/\pi},$$

since an estimate of δ can be obtained by rescaling ξ_c.

These results imply that the non-centrality parameters of the asymptotic distributions of the $F-$statistics corresponding to each regression (interacted, direct without measurement error, and direct with measurement error) are:

$$\rho_c^{2,int} = N\frac{\delta^2}{\frac{3\sigma_v^2+\sigma_\eta^2/\sigma_R^2}{2\sigma_f^2/\pi}}, \quad \rho_c^{2,dir} = N\frac{\delta^2}{\sigma_v^2/\sigma_f^2}, \quad \rho_{c,e}^{2,dir} = N\frac{\delta^2}{\left(\sigma_v^2+\sigma_e^2\right)/\sigma_f^2}.$$

It follows immediately that $\rho_c^{dir} > \rho_c^{int}$, so, in the absence of measurement error, the direct regression approach has greater power. In the presence of measurement error, $\rho_{c,e}^{dir} > \rho_c^{int}$ provided

$$\sigma_v^2 > \frac{2/\pi\sigma_e^2 - \sigma_\eta^2/\sigma_R^2}{3-2/\pi},$$

the variance of the MPC unexplained by f_j is large enough. This condition will always be satisfied if $2/\pi\sigma_e^2 < \sigma_\eta^2/\sigma_R^2$, so measurement error is small relative to residuals in the consumption regression and and/or the variance of the rebate is small. While this example is highly stylized, the finding that observables explain very little of the variation in MPCs is suggestive that our direct regression approach may have desirable power properties, even in the presence of measurement error.

Case 2: MPCs as a discrete function of observables

We now consider a model where the MPCs follow a discrete function of some scalar observable characteristic, f_j, consistent with the discrete MPCs in our empirical model. For simplicity, we assume the MPCs take only two values. In particular,

$$\lambda_j = a + \delta I_j.$$

Otherwise, we maintain the assumptions and notation of Case 1. Our direct approach estimates the same regression as before,

$$\lambda_j = \iota_d + \zeta_d f_j + v_j.$$

It follows that

$$\iota_d = a + \delta/2, \quad \zeta_d = \quad \delta/\sqrt{2\pi\sigma_f^2} \quad v_j = \delta\left[\left(2I_j - 1\right)/2 - \frac{f_j}{\sqrt{2\pi\sigma_f^2}}\right].$$

The variance of v_j is $\delta^2\left(\frac{\pi-2}{4\pi}\right)$, from which it follows that

$$\text{avar}\left(\sqrt{N}\hat{\zeta}_d\right) = \delta^2\frac{\pi-2}{4\pi\sigma_f^2}$$

and thus

$$\text{avar}\left(\sqrt{N}\hat{\delta}_d^{dir}\right) = \delta^2\frac{\pi-2}{2}.$$

In the presence of measurement error in the individual MPCs (which can either represent simple measurement error in binary estimates, or also account for the fact that we use weighted MPCs which consist of a weighted average of the individual's true MPC and additional MPCs), this final expression becomes

$$\text{avar}\left(\sqrt{N}\hat{\delta}_{d,e}^{dir}\right) = \delta^2\frac{\pi-2}{2} + 2\pi\sigma_e^2$$

Turning to the interaction approach, we have

$$\Lambda_d = a, \quad \xi_d = \delta \ ,$$

so

$$\text{avar}\left(\sqrt{N}\hat{\delta}_d^{int}\right) = \frac{4\sigma_\eta^2}{\sigma_R^2}.$$

The corresponding non-centrality parameters are:

$$\rho_d^{2,int} = N\frac{\delta^2}{4\sigma_\eta^2/\sigma_R^2}, \rho_d^{2,dir} = N\frac{2}{\pi-2}, \delta^2 > 0, \rho_{d,e}^{2,dir} = N\frac{\delta^2}{\delta^2\frac{\pi-2}{2} + 2\pi\sigma_e^2}.$$

Note that the non-centrality parameter for the direct case without measurement error is not well-defined for $\delta^2 = 0$, since the regression residuals will be identically zero. A striking result is that, in the absence of measurement error, the power for direct regression is a constant for a given N, and does not depend on any parameter values. This means the direct regression approach has a "maximin" power property, maximizing worst-case power; the power is constant, no matter how small δ may be. The direct regression approach will be more powerful provided

$$\delta^2 < \frac{8}{\pi-2}\frac{\sigma_\eta^2}{\sigma_R^2},$$

which expresses the size of the squared effect relative to the scaled noise-to-signal ratio in the consumption equation. For small effects, direct regression will have better power properties. With measurement error, the condition instead depends on the noise-to-signal ratio minus $2\pi\sigma_e^2$, so for a given noise-to-signal ratio, the size of the effect for which direct regression is preferable is decreasing in the size of measurement error.

Overall, these two cases illustrate that, regardless of the precise theoretical setting, there may be power advantages to using the direct regression approach. The argument is naturally strongest in the absence of measurement error in recovering the MPCs, but is by no means limited to such settings.

B The role of correlated heterogeneity in quantile regression

In this section, we illustrate a challenge faced when using quantile regression to recover the MPC distribution in an intuitive example that shows how the recovered MPC distribution is impacted by heterogeneity in additional parameters.

We consider a simple setting, where there are two possible fixed effect values, $\alpha_j \in \{-10,000, 10,000\}$ (the order of magnitude of our estimated fixed effects), and two MPCs, $\lambda_j = \{0.20, 0.70\}$. We draw non-zero rebate values $R_j \sim N\left(900, 100^2\right)$, centered at the median in our data. We then generate data according to

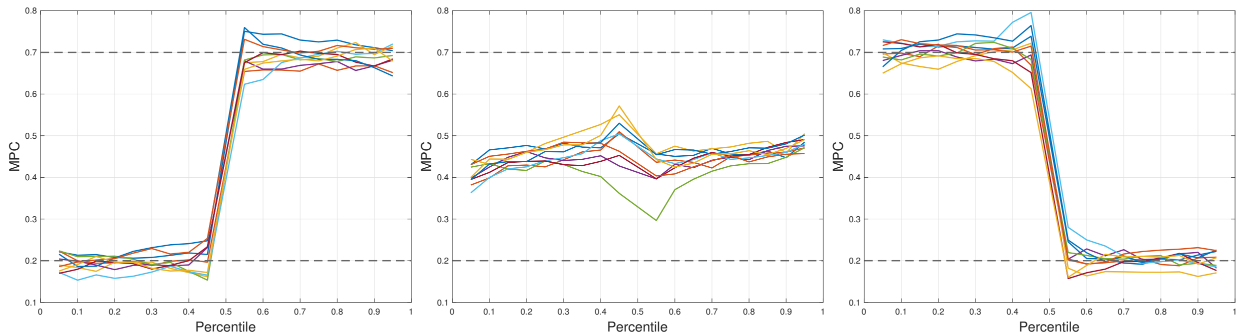
$$\Delta C_j = \alpha_j + \lambda_j R_j + \epsilon_j, j = 1, \ldots, N,$$

where $\epsilon_j \sim N\left(0, 1000^2\right)$, somewhat lower than the estimated noise in the data. We set $N = 100,000$, with 17.5% of observations receiving a rebate, as in our data, with $R_j = 0$ for the others.

We assume α_j and λ_j take each value with 50% probability. We consider three possible relationships between α_j and λ_j. First, we assume that they are perfectly positively correlated, so $\left(\alpha_j, \lambda_j\right) \in \{(-10,000, 0.20), (10,000, 0.70)\}$, with equal probabilities. Next, we assume that fixed effects and MPCs have zero correlation. Thus, $\left(\alpha_j, \lambda_j\right) \in \{(-10,000, 0.20), (10,000, 0.70), (-10,000, 0.70), (10,000, 0.20)\}$, with equal probabilities. Finally, we assume that MPCs and fixed effects are perfectly negatively correlated, so $\left(\alpha_j, \lambda_j\right) \in \{(10,000, 0.20), (-10,000, 0.70)\}$, with equal probabilities. For each specification, we draw 10 samples, estimate the model using quantile regression for every fifth percentile, and plot the estimated MPC distributions in Figure 8. The first panel shows that when the fixed effects and MPCs are positively correlated, the MPCs are well estimated; half of the distribution is associated with an MPC around 0.20, and half with an MPC around 0.70. Because the fixed effects dominate the conditional distribution, and the MPCs are correlated with the fixed effects, the lower MPC aligns with the lower half of the distribution. In the second panel, there is zero correlation between fixed effects and MPCs. Since the percentile of the distribution to which each observation corresponds is driven largely by the fixed effect, the two MPCs occur with approximately equal frequency at each percentile, so a value near the average MPC is estimated *at each percentile*. Finally, the third panel shows that when fixed effects and MPCs are negatively correlated, the MPCs are again well-estimated, as in the first panel. However, this time the high MPC corresponds to the lower half of the distribution, since it aligns with the lower fixed effect.

Figure 8: The role of correlated heterogeneity in quantile regression

(a) Positive correlation (b) Zero correlation (c) Negative correlation



Notes: Figure 8 plots the estimated MPCs from quantile regression for every fifth percentile for 10 samples of simulated data for three specifications. In each specification, both fixed effects and MPCs take two possible values. In the first panel, fixed effects and MPCs are perfectly positive correlated, in the second they have zero correlation, and in the third they are perfectly negatively correlated. The dashed lines represent the two true MPC values.

These results show that if there is heterogeneity in other parameters besides the MPC, the relationship between such heterogeneity and the MPC will impact the econometrician's ability to recover the distribution of MPCs using quantile regression.

We find in our empirical results that the MPC heterogeneity estimated by Misra and Surico (2014) is in fact exaggerated relative to ours, as opposed to the compressed distribution we observe in this highly simplified example. Alternative patterns, like that one, are entirely possible depending on the precise DGP as groups and controls are added and the correlations between group-specific parameters change.

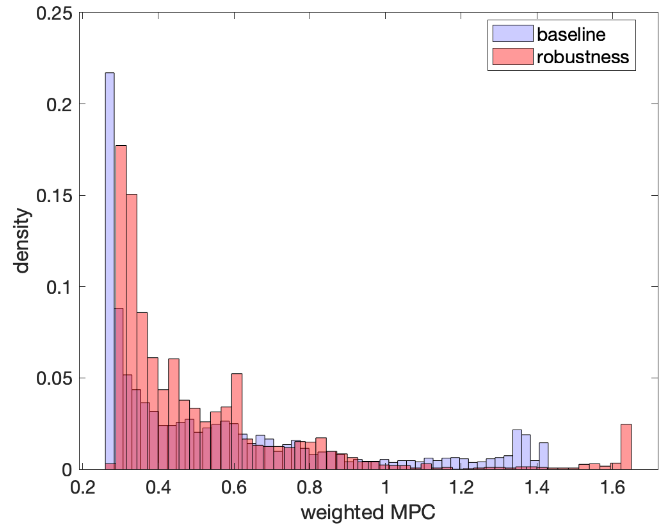
C Supplemental empirical results

In this section we report additional empirical results. First, we allow for heterogeneity in all coefficients in Equation (2), which leaves the estimated distribution of MPCs largely unchanged. Second, we explore robustness to different first stage specifications. Third, we report additional details on the MPC distribution for several specifications. Fourth, we document further results on the relationship between MPC heterogeneity and observable characteristics. Fifth, we provide details on the interpretation of the MPC in this setting. Finally, we describe how we adjust the R^2 in Table 2 to account for measurement error.

C.1 Heterogeneous coefficients on controls

Our baseline specification assumes common coefficients on time dummies and household-level controls. It is natural to wonder if there is also a role for heterogeneity with respect to

Figure 9: The MPC distribution with heterogeneous coefficients on controls



Notes: The light blue histogram plots the estimated MPC distribution for the baseline total expenditures specification as in Figure 1. The red histogram plots the weighted MPC distribution estimated in a specification that allows coefficients on all controls ω_j to be group specific. In this specification, the BIC chooses $G = 7$.

those covariates. In Equation (2), this amounts to interacting ω_j with the group dummies and allowing β to vary across group. Figure 9 plots the estimated MPC distributions for total expenditures allowing for this heterogeneity against our baseline distribution; the distribution is very similar across the two specifications. The distribution still displays a clear lower bound of about 26 cents to the dollar. While allowing for full heterogeneity stretches out the distribution slightly to the right, the average MPC differs only by 4 cents in the two specifications. Very similar findings are obtained if we allow heterogeneity in all household characteristics, but maintain homogeneous coefficients on time dummies.

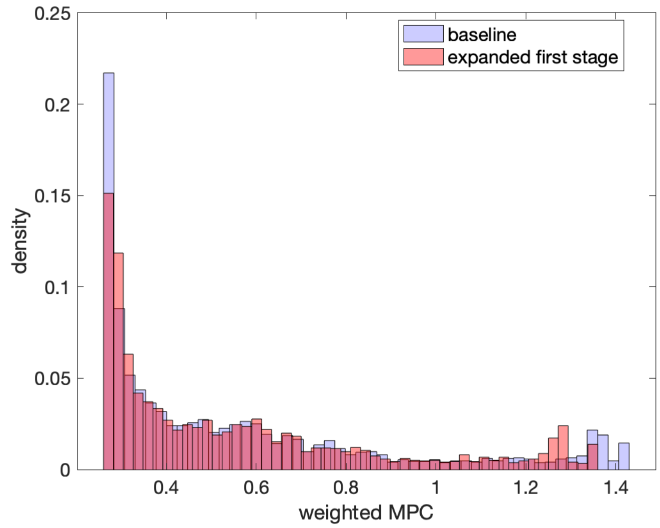
C.2 Alternative first stage

As discussed in 3.1, we estimate a homogeneous first stage that omits demographic controls and time dummies. As we argue in Supplement A.1, this approach allows us to obtain unbiased estimates. In this section, we report empirical results supporting this claim and the robustness of our recovered distribution to the specification of the first stage.

First, we report the estimated distribution of weighted MPCs when the first stage includes all controls. As discussed in Supplement A.1, such an approach may lead to biased MPC estimates if there is unmodeled heterogeneity in the first stage. Figure 10 shows that any such bias appears small in practice.

Second, we show in Figure 11 that results are largely unchanged if we only add time

Figure 10: The MPC distribution with expanded first stage



Notes: The light blue histogram plots the estimated MPC distribution for the baseline total expenditures specification as in Figure 1. The red histogram plots the weighted MPC distributions estimated with a first stage regression that also includes all controls (demographics and time dummies). In this alternative model, the BIC chooses $G = 5$.

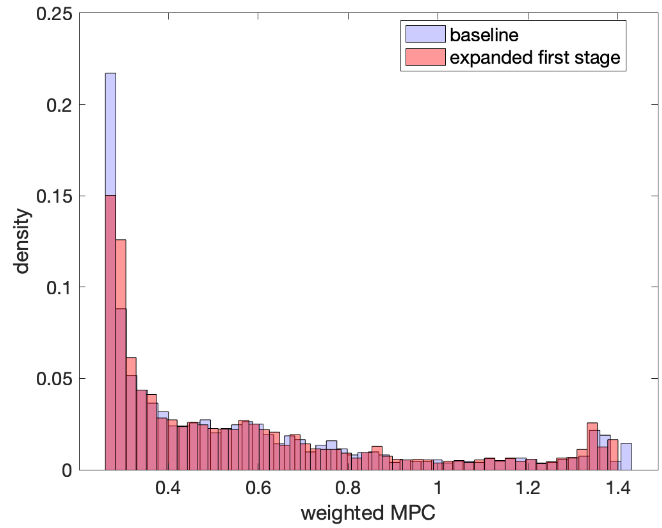
dummies, the only controls correlated with our instrument, to our baseline first stage. In our baseline, we impose that the first stage coefficients on these dummies are zero since the month in which a household receives a rebate is random and unrelated to its rebate value. In practice, the estimated first stage coefficients are generally small and not statistically significant, so this modification has little effect on the distribution recovered in the second stage.

As discussed in Supplement A.1, if heterogeneity in the first stage and second stage are perfectly aligned, our homogeneous first stage without controls should recover the same second stage distribution as when estimating a first stage including controls with heterogeneity aligned to that of the second stage. Both will be unbiased. We consider an iterated estimation approach where the group structure estimated in the second stage is used to dictate a group-specific intercepts in the first stage, before re-estimating the second stage (see Section (A.1) for a discussion of why other coefficients remain fixed). Figure 12 reports the results, with the distributions essentially unchanged, suggesting that in our case first and second stage heterogeneity are well-aligned.

C.3 The MPC distribution: additional results

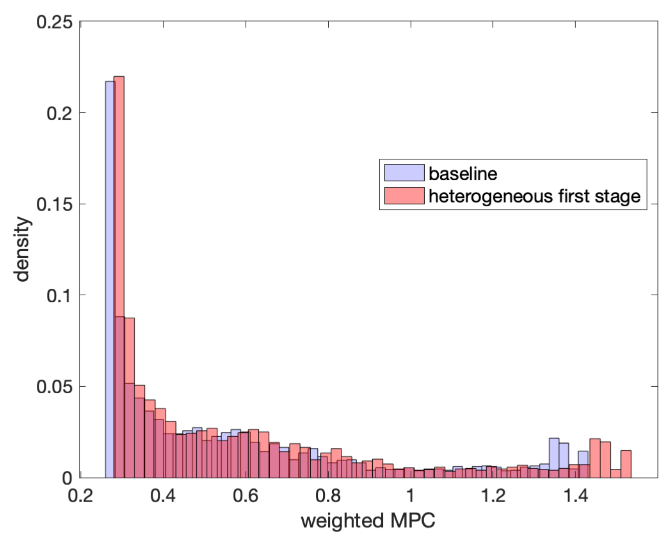
We show how the distribution of marginal propensities to consume is robust to the choice of G. In Table 6 we report relevant moments of the MPC distribution when estimated

Figure 11: The MPC distribution with expanded first stage - time dummies



Notes: The light blue histogram plots the estimated MPC distribution for the baseline total expenditures specification as in Figure 1. The red histogram plots the weighted MPC distributions estimated with a first stage regression that also includes time dummies. In this alternative model, the BIC and CV choose $G = 5$.

Figure 12: The MPC distribution with heterogeneous first stage



Notes: ***Notes***: The light blue histogram plots the estimated MPC distribution for the baseline total expenditures specification as in Figure 1. The red histogram plots the weighted MPC distributions estimated in a specification that allows coefficients on first-stage intercepts to be group specific. In this alternative model, the BIC chooses $G = 5$.

with different numbers of groups. As discussed in the text, the BIC meaningfully flattens between $G = 2$ and $G = 5$, and hence we show results for these groups. The distribution is barely changed. As expected, we see a more pronounced bimodality with 2 groups. This is visually represented in Figure 13. Moreover, the average of weighted MPCs is quite similar across groups, ranging between 0.54 and 0.62. In all distributions, about 25% of households consume between one quarter and one third of the rebate.

Table 6: MPCs out of the tax rebate under GMIVR: different G

	Mean	Std. Dev.	p25	p50	p75
$G = 1$	0.52	-	0.52	0.52	0.52
$G = 2$	0.62	0.44	0.32	0.37	0.77
$G = 3$	0.62	0.42	0.31	0.46	0.74
$G = 4$	0.53	0.31	0.28	0.41	0.70
$G = 5$	0.56	0.33	0.29	0.43	0.69

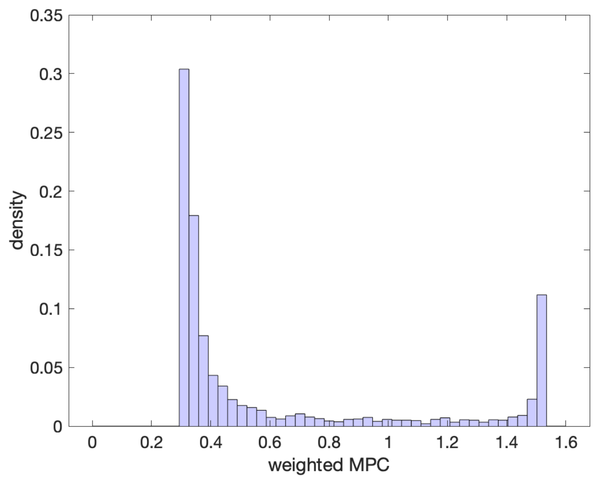
Notes: Each row reports various statistics for the distribution of weighted MPCs, estimated via GMIVR with a different number of groups G.

Table 7 shows the statistical significance of the point estimates for the MPCs, in the baseline specification under GMIVR. In the left panel, we make use of the analytical formulas outlined in Section 3 to compute Wald tests of pairwise equality across MPCs (accounting for both uncertainty in individual weights and first-stage estimation error). The right panel shows the same tests, taking the estimated weights as given; we report these results to parallel tests typically conducted in the literature, where group membership is taken as known (based on assumed observable relationships). In Table 8, we repeat the same analysis for nondurables and durables. For compactness, we only show tests accounting for uncertainty in the weights.

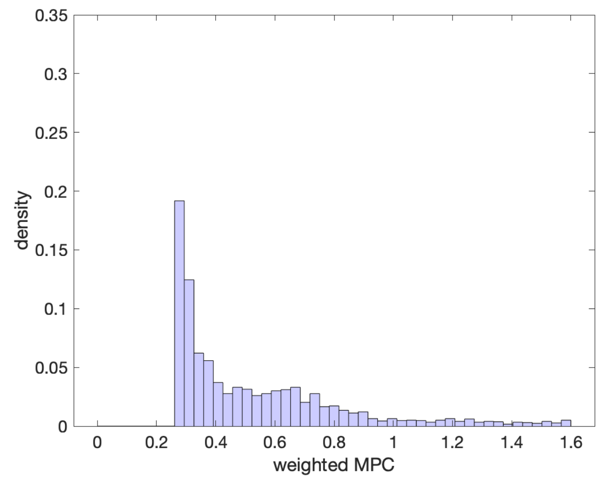
We also show that nondurable MPCs are meaningfully above zero when repeating the bootstrap exercise of Figure 2. Results are shown below in Figure 14a. In contrast, the same bootstrap exercise confirms there is a non-negligible portion of households who do not spend at all on durables, reported in Figure 14b.

The results shown in Section 4 use an IV strategy and omit households that never receive the rebate from the sample. Such an approach focuses on the most credible identification strategy. In this section we show how our results change for alternative specifications. In Figure 15, we show the IV results from Figure 1, but retaining the households that never get the rebate. We note that the shape of the distribution is broadly unchanged, but it shifts to the left. Similarly, the homogeneous MPC is substantially lower when including all households, in line with findings by Parker et al. (2013). Such behavior suggests

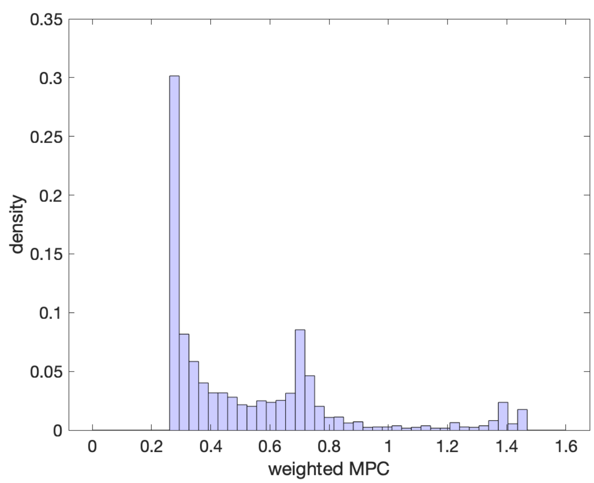
Figure 13: MPCs out of the tax rebate under GMIVR: different G



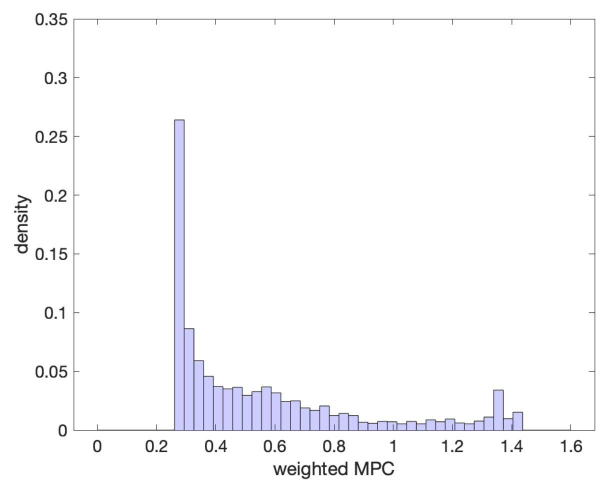
(a) G=2



(b) G=3



(c) G=4



(d) G=5

***Notes*:** As in Figure 1, each panel plots a histogram (light blue bars) of the GMIVR-estimated distribution of MPCs for total expenditures among households that received the rebate, defined as in Parker et al. (2013), for a different number of groups, *G*. The sample is defined as in the text.

Table 7: Statistical tests of MPCs: total expenditures

(a) Analytical standard errors

	MPC				
	0.08	0.35	0.68	1.37	1.44
0.08	0.61				
	(0.44)				
0.35	2.29	5.61			
	(0.13)	(0.02)			
0.68	0.61	2.29	0.62		
	(0.43)	(0.67)	(0.37)		
1.37	10.08	4.86	0.72	11.13	
	(0.00)	(0.03)	(0.39)	(0.00)	
1.44	3.62	2.22	0.55	0.01	4.06
	(0.06)	(0.14)	(0.46)	(0.93)	(0.04)

(b) Conditional on weights

	MPC				
	0.08	0.35	0.68	1.37	1.44
0.08	5.23				
	(0.02)				
0.35	16.89	26.48			
	(0.00)	(0.00)			
0.68	1.95	0.57	2.50		
	(0.16)	(0.45)	(0.11)		
1.37	23.81	14.14	1.90	26.73	
	(0.00)	(0.00)	(0.17)	(0.00)	
1.44	17.99	11.14	2.03	0.03	20.12
	(0.00)	(0.00)	(0.15)	(0.87)	(0.00)

Notes: The table shows $F-$statistics from pairwise two-sided Wald tests of equality of MPCs (the diagonals show tests of equality with zero) for the baseline total expenditures specification estimated under GMIVR. In the left panel, standard errors account for uncertainty in the weights and first stage estimation error. In the right panel, GMIVR weights are taken as given, to parallel the way that group assignment is taken as known in the existing literature. $p-$values are reported in parentheses.

that including households that never get the rebate leads to attenuation bias, as they are likely to have different characteristics than rebate recipients. Such bias is present across the whole distribution and affects the average MPC as well.

Estimating the MPC distribution with a standard OLS approach, using the entire sample, has similar implications, as we show in Figure 16. This approach introduces additional potential bias, since the size of the rebate is potentially endogenous. Even in this imperfect setting, however, our main results are confirmed. All households consume at least one fifth of the rebate, the share of households with a given MPC slowly decays as the MPC increases, and a non-negligible portion of households consume the rebate in its entirety.

C.4 What drives MPC heterogeneity: additional results

In this section we show additional results on MPC heterogeneity, for the baseline specification estimated with GMIVR. Figure 17 graphically displays the correlation of MPCs for durable and nondurable goods.

We have shown in Section 4.3 that some household characteristics individually correlate with the MPC distribution. Here, we analyze whether the linear correlation with age and liquid wealth hides some non-linear pattern. Figure 18a suggests a positive and convex relationship between the weighted MPC and log liquid wealth. The relationship

Table 8: Test for MPC equality: nondurables and durables

(a) Nondurables

	MPC	
	0.12	0.27
0.12	5.03	
	(0.02)	
0.27	2.81	11.36
	(0.09)	(0.00)

(b) Durables

	MPC			
	−0.05	0.11	0.14	0.99
−0.05	5.30			
	(0.02)			
0.11	14.68	7.10		
	(0.00)	(0.01)		
0.14	5.30	14.68	3.19	
	(0.07)	(0.83)	(0.19)	
0.99	4.02	2.88	2.47	3.65
	(0.04)	(0.09)	(0.12)	(0.06)

Notes: The table shows $F-$statistics from pairwise two-sided Wald tests of equality across MPCs (the diagonals show tests of equality with zero). Standard errors account for uncertainty in the weights and first stage estimation error. $p-$values are reported in parentheses. Standard errors account for uncertainty in the weights and first stage estimation error, as outlined in Section 3.

looks instead concave with respect to age, as shown in Figure 18b.

Neither relationship is robust to the inclusion of a set of controls, as we further show in Table 9. In the same table, we also confirm that the findings shown in Section 4.3 are robust to controlling for liquid wealth.

In Figure 19, we show that "rich-spenders" (i.e. households with high APC and high total income) have high MPCs for nondurable and durable expenditures too.

Finally, we show that the relationship between MPCs and observable characteristics is robust to the selection of the number of groups used in the GMIVR estimation. In Table 10 we repeat the analysis of Table 2 for the groups associated with the flattening in the BIC. We report the results for total expenditure MPCs. In all instances, non-salary income and the APC are the only two observables that are statistically significantly associated with the MPC. The relationship is also quantitatively stable, especially when looking at 4 and 5 groups. Moreover, the adjusted R^2 is low in all cases, never exceeding 16%.

Table 9: Explanatory variables of MPCs: including liquid assets

	(1) Total	(2) Nondurables	(3) Durables
dummy for no salary	-0.416	-0.009	-0.310
	(0.278)	(0.035)	(0.212)
log salary income	-0.018	0.001	-0.015
	(0.018)	(0.002)	(0.014)
log non-salary income	0.179***	0.022***	0.137***
	(0.031)	(0.004)	(0.026)
mortgage interest/income	0.255	0.015	0.189
	(0.137)	(0.019)	(0.121)
APC	0.262***	0.032***	0.191***
	(0.040)	(0.006)	(0.037)
homeowner dummy	0.041	0.014**	-0.018
	(0.035)	(0.005)	(0.026)
dummy for mortgage	-0.059	-0.006	-0.036
	(0.039)	(0.005)	(0.027)
age	-0.004	0.000	-0.006
	(0.005)	(0.001)	(0.004)
log liquid assets	0.002	-0.000	-0.000
	(0.007)	(0.001)	(0.005)
dummy for liquid assets <= 0	0.039	0.005	-0.043
	(0.090)	(0.012)	(0.068)
N	723	712	722
adj. R^2	0.145	0.136	0.115

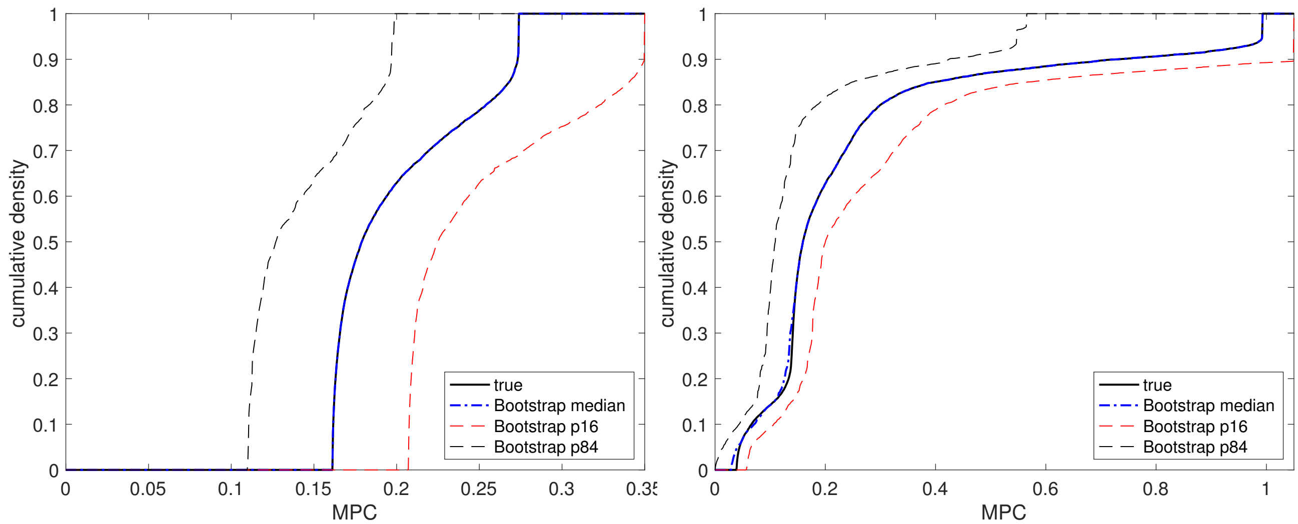
Notes: All logged variables take a value of log(0.001) when the raw value is 0 or negative. Non-salary income is positive for all observations. Standard errors are robust to heteroskedasticity and reported in parentheses. We control for marriage dummies, education dummies, number of children and age squared; those coefficients are not reported. *, **, and *** denote significance at the 10, 5, and 1% levels respectively.

Table 10: Explanatory variables of MPCs: robustness to G

	(1) G = 2	(2) G = 3	(3) G = 4	(4) G = 5
dummy for no salary	-0.500	-0.384	-0.225	-0.371
	(0.300)	(0.282)	(0.209)	(0.233)
log salary income	-0.025	-0.017	-0.008	-0.018
	(0.020)	(0.018)	(0.014)	(0.015)
log non-salary income	0.246***	0.244***	0.184***	0.192***
	(0.029)	(0.027)	(0.020)	(0.022)
mortgage interest/income	0.204	0.106	-0.020	0.129
	(0.149)	(0.135)	(0.090)	(0.113)
APC	0.376***	0.394***	0.309***	0.293***
	(0.045)	(0.046)	(0.035)	(0.033)
homeowner dummy	0.053	0.045	0.024	0.042
	(0.038)	(0.035)	(0.026)	(0.029)
dummy for mortgage	-0.081*	-0.061	-0.027	-0.049
	(0.039)	(0.037)	(0.027)	(0.030)
N	1079	1079	1079	1079
adj. R^2	0.136	0.152	0.159	0.146

Notes: All logged variables take a value of log(0.001) when the raw value is 0 or negative. Non-salary income is positive for all observations. Standard errors are robust to heteroskedasticity and reported in parentheses. We control for marriage dummies, education dummies, number of children, age and age squared; those coefficients are not reported. *, **, and *** denote significance at the 10, 5, and 1% levels respectively.

Figure 14: Bootstrapped distribution of MPCs out of the tax rebate



(a) Nondurable MPCs (b) Durable MPCs

Notes: The black solid line plots the CDF of the estimated distribution of MPCs for nondurables (left panel) and durables (right panel), shown in Figure 3. The blue dash-dotted line shows the median CDF of the estimated distribution of MPCs across 250 bootstraps. The dashed black and red lines denote the centered 68% confidence interval.

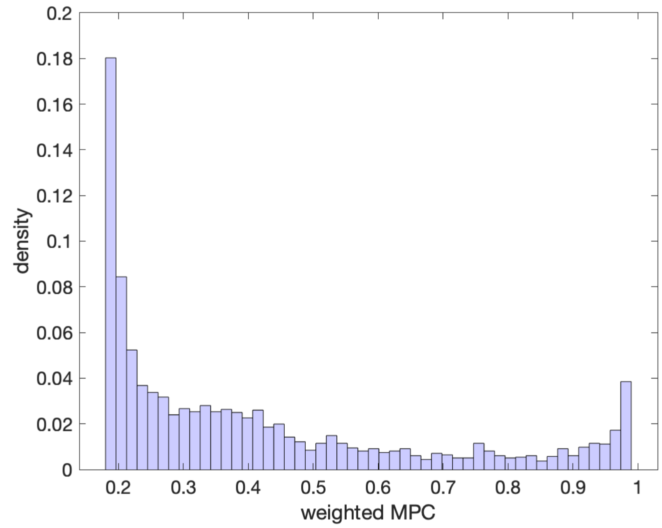
C.5 Rebate coefficient versus MPC

As discussed by Kaplan and Violante (2014), λ may not correctly measure the marginal propensity to consume out of a transitory income shock, but is instead better thought of as a "rebate coefficient". This is because the control group of non-recipients in period t is made of three groups: (i) households that never receive the rebate, (ii) households that have not yet received a rebate, but may anticipate receiving the rebate in the future, and (iii) households that have already received the rebate. Our baseline specification drops group (i). The second group might display a positive MPC out of news of the rebate, biasing the estimated rebate coefficient λ downward. Similarly, the third group might also have a positive lagged MPC out of the rebate, further contributing to a downward bias. Following Kaplan and Violante (2014), we modify the specification in Equation (1) by introducing the lag of the rebate variable R_j^{lag} so that the estimated rebate coefficient can be interpreted as an MPC:

$$\Delta C_j = \beta' W_j + \lambda R_j + \lambda^{\text{lag}} R_j^{\text{lag}} + \alpha + \epsilon_j \tag{13}$$

By absorbing the lagged consumption response, this modification accounts for the fact that, in the baseline specification, the control group includes households that received

Figure 15: Estimated distribution of MPCs out of the tax rebate: full sample



Notes: The histogram (light blue bars) plots the estimated distribution of MPCs for total expenditures among households that received the rebate, defined as in Parker et al. (2013). Differently from Figure 1, we do not drop from the sample the households that never get the rebate. As in that framework, the BIC selects 5 groups.

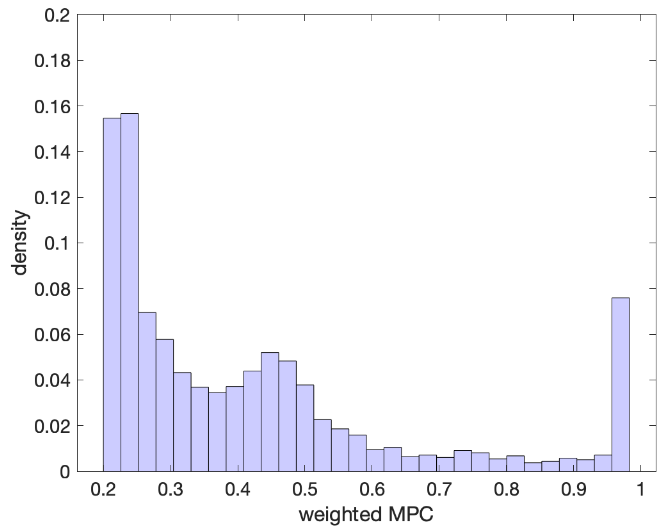
the rebate in the past, and whose consumption response might be persistent.[31] We then interact the rebate, its lagged value, and the constant with the group indicators $\mathbf{1}[j \in g]$, and use GMIVR algorithm to estimate the vector of coefficients $\left\{\lambda_g, \lambda_g^{\text{lag}}, \alpha_g, \beta\right\}$. In our implementation, we consider a two-stage least squares specification as discussed in the text. In this specification, however, we have two instruments - an indicator for whether the rebate was received today and an indicator for whether the rebate was received in the previous period - for two instrumented variables, R_j and R_j^{lag}. In Figure 20, we show that the distribution of weighted contemporaneous MPCs, λ_g, is very similar to the one estimated in the baseline specification.

C.6 Measuring the explanatory power of observables for heterogeneity

An advantage of our two-stage approach is that we can regress the full heterogeneity of the MPC distribution on observables. Not only does this allow us to characterize which observable variables remain significant predictors of the MPC in these joint regressions, but it also enables us to compute the share of heterogeneity that is predicted based on observables. This exercise gives a measure of what share of heterogeneity is truly latent

[31] This is true as long as the persistent effect of the rebate lasts strictly less than three quarters. Moreover, we assume that the policy is fully anticipated by all households. In an intermediate information case in which, for instance, the policy enters the agents' information set after the receipt of the first rebate, this specification cannot fully account for anticipatory effects often labelled as the MPC out of news.

Figure 16: Estimated distribution of MPCs out of the tax rebate: OLS



Notes: The histogram (light blue bars) plots the estimated distribution of MPCs for total expenditures using GMLR (OLS) among households that received the rebate, defined as in Parker et al. (2013). With this specification, the BIC selects 3 groups.

– driven by fundamentally unobservable factors such as preference heterogeneity, or correlated with variables that simply are not included in our dataset. One complication is that the R^2 computed from these regressions provides a lower bound on the true R^2 due to measurement error in the estimated MPCs. This is particularly true since we use individuals' weighted MPCs, which, unless posterior weights are binary, will always contain measurement error (even if $\hat{\lambda} = \lambda$). In this section, we describe a simple exercise to adjust the R^2 for measurement error in recovering the MPCs in the first stage.

Denote the estimated posterior-weighted MPC of individual i at time t as $\hat{\lambda}_j = \lambda_{g(j)} + \hat{e}_j$. Suppose that regressing the true MPCs, $\lambda_{g(j)}$, we have

$$\lambda_{g(j)} = c + \mu' F_j + v_j. \tag{14}$$

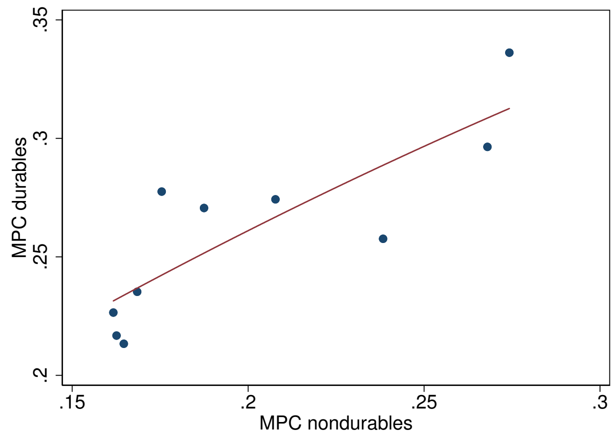
Then the regression of estimated MPCs on observables takes the form

$$\hat{\lambda}_j = c + \mu' F_j + v_j + \hat{e}_j. \tag{15}$$

Based on the infeasible (14),

$$R^2_{true} = 1 - \frac{E\left[v_j^2\right]}{\text{var}\left(\lambda_{g(j)}\right)},$$

Figure 17: The correlation of MPCs across consumption goods



Notes: The blue dots display a binscatter of household MPC estimates for durables against those for nondurables. Each dot shows the average weighted MPC for durable goods for each decile of the distribution of weighted MPCs for nondurable goods. The red line shows the quadratic fit.

while the value computed based on (15) is

$$R^2_{raw} = 1 - \frac{E\left[\left(v_j + \hat{e}_j\right)^2\right]}{\text{var}\left(\hat{\lambda}_j\right)}.$$

Under classical measurement error assumptions, so that $\hat{e}_j$ is orthogonal to $\lambda_{g(j)}$, the formula simplifies to

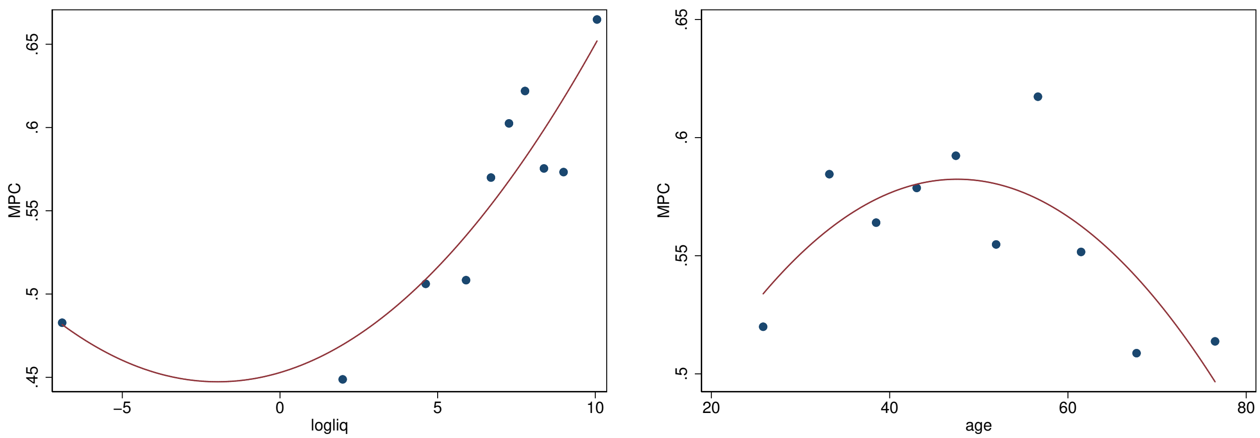
$$R^2_{raw} = 1 - \frac{E\left[v_j^2\right] + E\left[\hat{e}_j^2\right]}{\text{var}\left(\hat{\lambda}_j\right)}.$$

This value is biased towards zero (since $\text{var}\left(\hat{\lambda}_j\right) = \text{var}\left(\lambda_{g(j)}\right) + E\left[\hat{e}_j^2\right]$), potentially leading us to conclude that too small a share of MPC heterogeneity can be explained by observables. As in Majeske et al. (2010), these expressions can be rearranged to show that

$$R^2_{true} = \frac{R^2_{raw}}{1 - E\left[\hat{e}_j^2\right] / \text{var}\left(\hat{\lambda}_i\right)}. \tag{16}$$

To apply this formula, some measure of $E\left[\hat{e}_j^2\right]$ must be computed. The methods proposed in Majeske et al. (2010) – based on taking repeated measurements in experimental settings – are infeasible. A bootstrap procedure cannot quantify measurement error in household MPCs (which is fixed for a given observation across samples). Indeed, it can only measure the variation in parameter estimates as the sample composition changes. For this reason,

Figure 18: Marginal propensities to consume: liquid wealth and age



(a) Liquid wealth

(b) Age

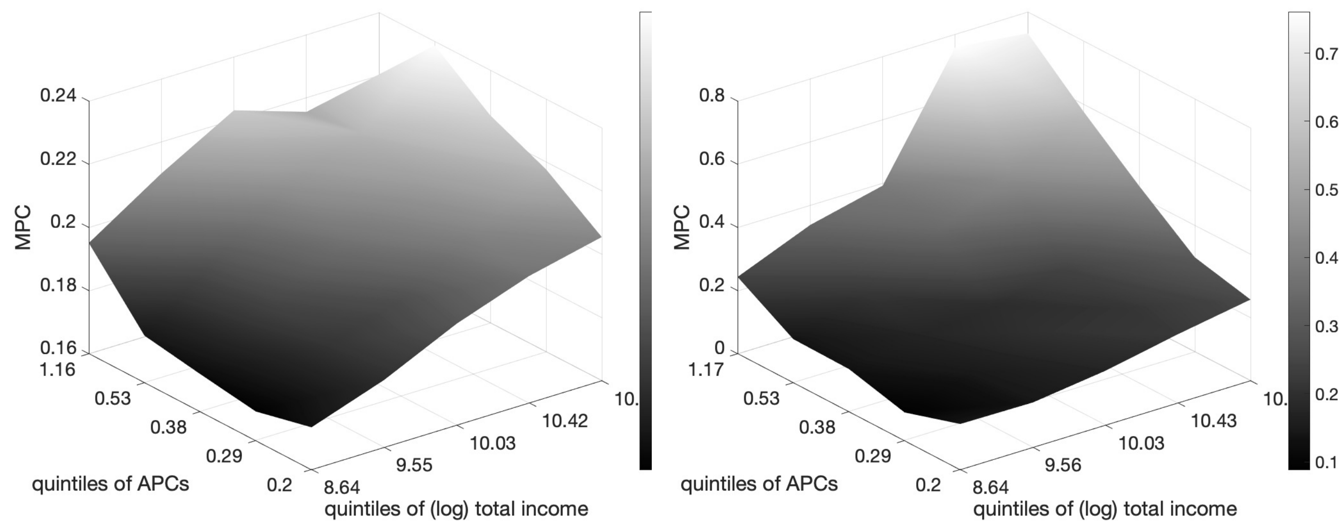
Notes: The blue dots display binscatters of the weighted MPC for total expenditures for each decile of the distribution of lagged log liquid wealth (left panel) and for each decile of the distribution of age of the reference person in the household (right panel). Log liquid wealth takes log(0.001) when liquid wealth is zero or negative. The red line shows the quadratic fit.

we use a proxy for measurement error motivated by our setting. In the true model, each individual has one of the G discrete MPCs, which we consistently estimate. However, the individual MPCs we measure are weighted averages of the G MPCs, where noise in each observation dictates non-binary weights. In the absence of measurement error, the distribution of weighted MPCs would collapse to G point masses. Thus, we use the deviation of an individual's weighted MPC from the discrete MPC with the highest posterior probability as a proxy for measurement error, and estimate $E\left[\hat{e}_j^2\right]$ as the sample average of these deviations squared (taking the estimated point MPCs as the truth). With this proxy in hand, we can implement (16) to obtain a back-of-the-envelope estimate of R^2_{true}. When we do so, the baseline R^2 rises from 15% to 26%, still indicating that the majority of heterogeneity remains unexplained by observables.

C.7 Uncorrelated spurious heterogeneity

In Figure 21, we repeat the analysis shown in Figure 6, but with spurious heterogeneity arising from imposing $G = 2$ on samples generated assuming homogeneity of coefficients. Results are analogous. In only 7.2% of the samples is there significant evidence of a relationship between MPC heterogeneity and APC at the 5% level, a size distortion within the scope of Monte Carlo error. Moreover, the adjusted R^2 is 0.1% on average across samples, and never higher than 2%.

Figure 19: The relationship between MPCs, APCs, and income



(a) Nondurable MPCs

(b) Durable MPCs

The surfaces plot average weighted MPCs for pairs of quintiles of APC and log total income. The left panel considers nondurable expenditures and the right durable expenditures. The color bar on the right represents the MPC.

D Homogeneous effects and average heterogeneous effects

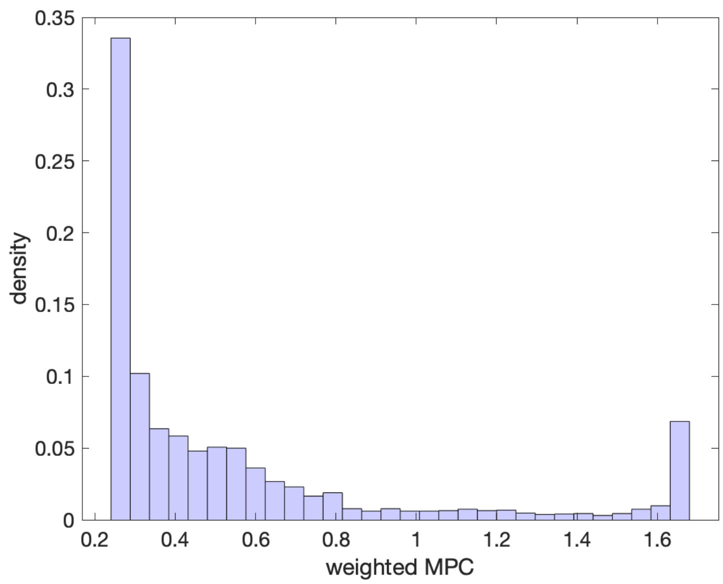
Across specifications, we find some deviation of the average MPC in our heterogeneous models from the MPC estimated in a homogeneous regression on the same data. If all regressors are independent of group membership, then the two coincide in population. However, if the distributions of even a subset of the regressors vary by group, the coefficients will generally differ. In this section, we present a simple example to illustrate why the two estimands may deviate and why such a discrepancy arises in our setting, even though the regressor of interest is constructed to be independent of household characteristics and group membership.

Consider a simple univariate regression with known discrete assignment $g(j)$, and two groups,

$$y_j = \alpha_{g(j)} + \delta_{g(j)} W_j + u_j, j = 1, \ldots, N,$$

and $E\left[W_j \mid g(j) = h\right] = 0$, $h = 1, 2$ (without loss of generality). Note that allowing for group-specific coefficients on W_j is without loss of generality; the results below go through with $\delta_1 = \delta_2 \equiv \delta$, since the expression for the average heterogenous coefficient simplifies and the homogeneous estimator is unchanged. Denote $X_j = \begin{bmatrix} 1 & W_j \end{bmatrix}'$. The average heterogeneous coefficients are simply $\sum_{h=1}^{2} \pi_h E\left[X_j X_j' \mid g(j) = h\right]^{-1} E\left[X_j y_j \mid g(j) = h\right]$.

Figure 20: Estimated distribution of MPCs out of the tax rebate: lagged responses



Notes: The histogram (light blue bars) plots the estimated distribution of the contemporaneous MPC for total expenditures, defined as in Parker et al. (2013), controlling for whether households received a rebate in the preceding quarter. The BIC selects 3 groups.

The homogeneous coefficients are given instead by

$$\left(\sum_{h=1}^{2} \pi_h E\left[X_j X_j' \mid g(j) = h\right]\right)^{-1} \begin{bmatrix} \sum_{h=1}^{2} \pi_h \alpha_h \\ \sum_{h=1}^{2} \pi_h E\left[W_j^2 \mid g(j) = h\right] \delta_h \end{bmatrix}.$$

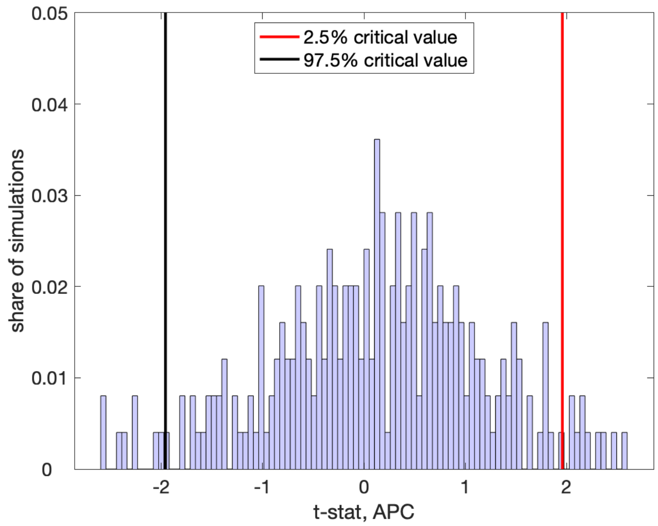
If the distribution of W_j does not vary over group, then both expressions simplify to

$$E\left[X_j X_j'\right]^{-1} \begin{bmatrix} \sum_{h=1}^{2} \pi_h \alpha_h \\ E\left[W_j^2\right] \sum_{h=1}^{2} \pi_h \delta_h \end{bmatrix},$$

since the expectations can be factored outside the summations. Otherwise, the estimated homogeneous coefficient on W_j, $\bar{\delta}$, will be tilted towards whichever group has a higher variance for W_j. This is because the "weight" each heterogeneous coefficient receives in the numerator is a function of $E\left[W_j^2 \mid g(j) = h\right]$. To see this explicitly, consider an example where $\pi_1 = \pi_2 = 1/2$, $E\left[W_j \mid g(j) = 1\right] = E\left[W_j \mid g(j) = 2\right] = 0$, $E\left[W_j^2 \mid g(j) = 1\right] = 1$, and $E\left[W_j^2 \mid g(j) = 2\right] = 99$. In this setting, the denominator for $\bar{\delta}$ is

$$\begin{bmatrix} 1 & 0 \\ 0 & \frac{1+99}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 50 \end{bmatrix},$$

Figure 21: Minimal correlation of spurious heterogeneity with observables ($G = 2$)



Notes: For each of 250 simulated samples, generated from the IV-estimated homogeneous DGP for total expenditures, we regress the weighted MPCs for total expenditure estimated imposing spurious heterogeneity ($G = 2$) on the set of observables used in Table 2. The histogram (light blue bars) plots the $t-$statistics for the coefficient on the APC. The red and black lines represent the critical values for a 5% test of equality with zero.

and the numerator is

$$\begin{bmatrix} \frac{\alpha_1+\alpha_2}{2} \\ \frac{\delta_1+99\delta_2}{2} \end{bmatrix},$$

so $\bar{\delta} = \frac{\delta_1+99\delta_2}{100} = 0.01\delta_1 + 0.99\delta_2$. While this example is particularly stark, it is clear that $\bar{\delta}$ is far from the π_h-weighted average of δ_h across groups. This example illustrates how differences may arise between homogeneous and average heterogeneous coefficients when the distribution of the regressor of interest covaries with the heterogenous parameter (or, equivalently in this setting, group membership). This discrepancy is analogous to the well-understood breakdown of identification of average causal effects by homogeneous regression in panel data when an entity's causal effect fails to be mean independent of regressors (see e.g., Arellano (2003), page 11).

However, in our model our regressor of interest is exogenous, with a constant distribution across groups by construction. Augmenting the model above to include such an exogenous variable, R_j, it can be shown that, if R_j is additionally independent of W_j, the distribution of which may vary by group, the homogeneous coefficient on R_j, $\bar{\lambda}$ will be identical to the average heterogeneous coefficient. Assuming without loss of generality that $E\left[R_j\right]=0$, it is easy to see why: the denominator is given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \sum_{h=1}^{2}\pi_h E\left[W_j^2 \mid g\left(j\right)=h\right] & 0 \\ 0 & 0 & E\left[R_j^2\right] \end{bmatrix},$$

and the numerator by

$$\begin{bmatrix} \sum_{h=1}^{2}\pi_h\alpha_h \\ \sum_{h=1}^{2}\pi_h E\left[W_j^2 \mid g\left(j\right)=h\right]\delta_h \\ \sum_{h=1}^{2}\pi_h E\left[R_j^2\right]\lambda_h \end{bmatrix},$$

so $\bar{\lambda}=E\left[R_j^2\right]^{-1}\sum_{h=1}^{2}\pi_h E\left[R_j^2\right]\lambda_h=\sum_{h=1}^{2}\pi_h\lambda_h$.

While in our setting the distribution of our instrument, the rebate receipt indicator, does not vary by group, and the instrument is not correlated with household-specific controls, it may be correlated with the exogenous time dummies for month of rebate receipt, since the rate of rebate disbursement was not constant. In the example above, this means that the additional assumption of independence from W_j does not hold, so $\sum_{h=1}^{2}\pi_h E\left[R_j W_j \mid g\left(j\right)=h\right]$ is non-zero and potentially varies by group, appearing off-diagonal in the denominator and in the latter two entries of the numerator. As in the simple bivariate case, $\bar{\lambda}$ will be tilted towards groups with higher-variance W_j, via the $E\left[R_j W_j \mid g\left(j\right)=h\right]$ "weights" in the numerator. In our empirical results, we nevertheless find that the homogeneous and average heterogeneous estimates are quite similar for the majority of our specifications. However, when a discrepancy does arise, these results suggest that it should not be surprising.

References

ARELLANO, M. (2003): *Panel Data Econometrics*, Advanced texts in econometrics, Oxford University Press.

KAPLAN, G. AND G. L. VIOLANTE (2014): "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82, 1199–1239.

MAJESKE, K. D., T. LYNCH-CARIS, AND J. BRELIN-FORNARI (2010): "Quantifying R 2 bias in the presence of measurement error," *Journal of Applied Statistics*, 37, 667–677.

MISRA, K. AND P. SURICO (2014): "Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs," *American Economic Journal: Macroeconomics*, 6, 84–106.

PARKER, J. A., N. S. SOULELES, D. S. JOHNSON, AND R. MCCLELLAND (2013): "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review*, 103, 2530–53.