CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 803

Hypothetical Shocks as Designed Data: What Survey Experiments Can Teach Us about Consumption Behavior

Tullio Jappelli

November 2026



University of Naples Federico II



University of Salerno



Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 803

Hypothetical Shocks as Designed Data: What Survey Experiments Can Teach Us About Consumption Behavior

Tullio Jappelli*

Abstract

Survey questions based on hypothetical income and wealth shocks are increasingly used to measure marginal propensities to consume (MPCs) and other behavioral responses. Their main advantage is not simply that they provide additional observations on household behavior, but that alternative survey designs generate different sources of variation and therefore identify different features of the consumption function. This paper distinguishes five designs: cross-sectional surveys in which all respondents face the same hypothetical shock; panel surveys that repeat the same shock; panels in which respondents face different shocks over time; cross-sectional surveys in which the same respondent answers several hypothetical scenarios; and randomized survey experiments in which different respondents are assigned different shocks. We examine what can and cannot be learned from each design, the assumptions required for identification, and the role of classical and non-classical measurement error. Applications to income shocks, intertemporal MPCs, house-price shocks, uncertain transfers, and social interactions illustrate the trade-offs involved. The comparison across designs shows that survey experiments are especially informative when treatments are closely linked to theoretical predictions and when their implications can be assessed across alternative identification strategies and against quantitative models.

JEL Codes: C83, C90, D12, E21.

Keywords: Marginal propensity to consume, Hypothetical income shocks, Survey experiments, Experimental design, Identification, Measurement error.

Acknowledgments: This paper grew out of a keynote lecture delivered at the Fourth Badolato Labour & Organization Workshop on 8 July 2026. I thank the participants for their helpful comments and discussion.

* University of Naples Federico II, CSEF, and CEPR. E-mail: tullio.jappelli@unina.it

1. Introduction

The marginal propensity to consume is a central empirical object in household economics. It summarizes how consumption responds to changes in household resources and plays an important role in models of consumption and saving, heterogeneous-agent macroeconomics, and fiscal policy. Empirical evidence on MPCs comes from three broad approaches. The first infers the MPC from observed expenditure responses to realized shocks, including tax rebates, stimulus payments, lotteries, unemployment shocks, and asset-price movements (Johnson, Parker and Souleles, 2006; Parker et al., 2013; Broda and Parker, 2014; Fagereng, Holm and Natvik, 2021; Golosov et al., 2024). A second approach uses stated responses to actual or anticipated policy payments, as in Shapiro and Slemrod (2003) and Sahm, Shapiro and Slemrod (2012). A third approach, which is the focus of this paper, relies on researcher-designed hypothetical scenarios.[1]

This third approach is part of a broader methodological shift in which surveys are increasingly used as experimental tools rather than simply as devices for measuring beliefs or intended behavior. Haaland et al. (2023) show how survey experiments can generate exogenous variation by manipulating respondents' information sets and discuss the design choices required for credible inference, including treatment construction, measurement, framing, and experimenter-demand effects. They also emphasize the usefulness of hypothetical vignettes when relevant features of the economic environment cannot be manipulated directly. In a macroeconomic context, Coibion and Gorodnichenko (2026) similarly stress that randomized survey treatments can create exogenous variation in otherwise endogenous variables and thereby identify their effects on household decisions. They also view hypothetical scenarios as a complementary and often more tractable approach when full randomized interventions, follow-up waves, or linked behavioral data are unavailable. These contributions motivate the perspective adopted here: hypothetical questions

[1] The MPC has attracted rapidly growing attention in the economics literature over the past two decades. Between 2000 and 2026, approximately 530 articles mentioning the MPC in their title or abstract were indexed in Scopus. Annual publications increased from fewer than 10 in the early 2000s to around 20–25 in the late 2010s, and 46 in 2025. The literature has also gained prominence in leading economics journals, including the *AER*, *QJE*, *ReStud*, *JME* and *AEJ-Macro*. This growth reflects the increasing importance of MPC heterogeneity for understanding household consumption behavior, the transmission of economic shocks, and the effectiveness of fiscal policy.

are valuable not only because they elicit stated responses, but because they can be designed to generate economically meaningful variation.

This is precisely their main advantage in the study of MPCs. By varying the size, sign, timing, uncertainty, or social environment of a hypothetical shock, either within individuals or across randomized groups, researchers can create sources of identifying variation that are often unavailable in naturally occurring data. Hypothetical scenarios are therefore particularly useful for testing comparative statics and distinguishing among theoretical mechanisms. At the same time, they are not a substitute for evidence from realized shocks. Rather, the two approaches are complementary: hypothetical experiments offer control over the treatment environment, while realized-shock studies provide evidence on actual behavior.

The central argument of this paper is that what economists learn from hypothetical MPC questions depends critically on how the survey is designed. Consider a generic question: *Suppose you unexpectedly receive (Y) euros. How much would you spend over the next 12 months?* The researcher can ask the same question to everyone, repeat it for the same respondents over time, change (Y) across waves, ask each respondent about several values of (Y), or randomly assign different scenarios across respondents. Although the underlying question is similar, these designs generate very different identifying variation. We organize the analysis around five survey architectures:

- cross-sectional data in which everybody faces the same hypothetical shock;
- panel data in which the same shock is repeated over time;
- panel data in which the hypothetical shock changes across periods;
- cross-sectional data in which each individual answers several different shock scenarios;
- randomized cross-sectional experiments in which each respondent receives only one randomly assigned scenario.

The five designs differ in the sources of identifying variation they exploit and the assumptions required for interpretation. Cross-sectional comparisons document MPC heterogeneity but cannot establish the causal effect of household resources. Panel data eliminate permanent individual heterogeneity, while repeated scenarios permit within-person comparisons

of alternative shocks. Randomized experiments provide exogenous variation in treatment characteristics but require larger samples to estimate heterogeneous responses. These distinctions matter for interpreting empirical findings: for example, a negative relationship between the MPC and cash-on-hand has different implications when estimated from cross-sectional differences rather than within-household changes.

A second theme concerns measurement error, which is especially important when MPCs are elicited through hypothetical scenarios. Classical response error in the reported MPC mainly increases residual variance and reduces statistical precision, provided it is unrelated to the variables used for identification. More problematic is non-classical reporting error, and responses influenced by anchoring, question order, framing, or heuristic reporting rules, potentially generating patterns that resemble theoretical predictions. We discuss how alternative elicitation methods and complementary survey designs can help detect these problems, and how comparisons with realized consumption responses and independently calibrated quantitative models can strengthen the interpretation of survey evidence.

The paper proceeds by first linking a simple consumption function to predictions for MPC heterogeneity, and then comparing the five survey designs. We subsequently discuss measurement and validation, and show how the same experimental logic has been extended beyond consumption to uncertainty, social interactions, labor supply, and portfolio choice.

2. A common framework

Consider a household with cash-on-hand X that receives an unanticipated transitory income shock Y.[2] Let consumption be determined by the consumption function $C = f(\mathrm{X} + Y)$; the marginal propensity to consume out of a shock of size Y can be defined as:

$$MPC(\mathrm{X}, Y) = \frac{\Delta C}{\Delta Y}$$

[2] Some studies define the hypothetical shock as an amount proportional to the respondent's income, for example, one or three months of household earnings, rather than as a fixed euro amount. This introduces additional complications because the size of the treatment then varies systematically across households and may be correlated with other determinants of the MPC. For ease of exposition, we abstract from this issue and present the discussion as if all respondents faced a fixed monetary shock (Y).

Thus, the MPC depends both on the household's initial cash-on-hand *X* and on the size and sign of the shock *Y*.[3] A simple reduced-form representation of this dependence is

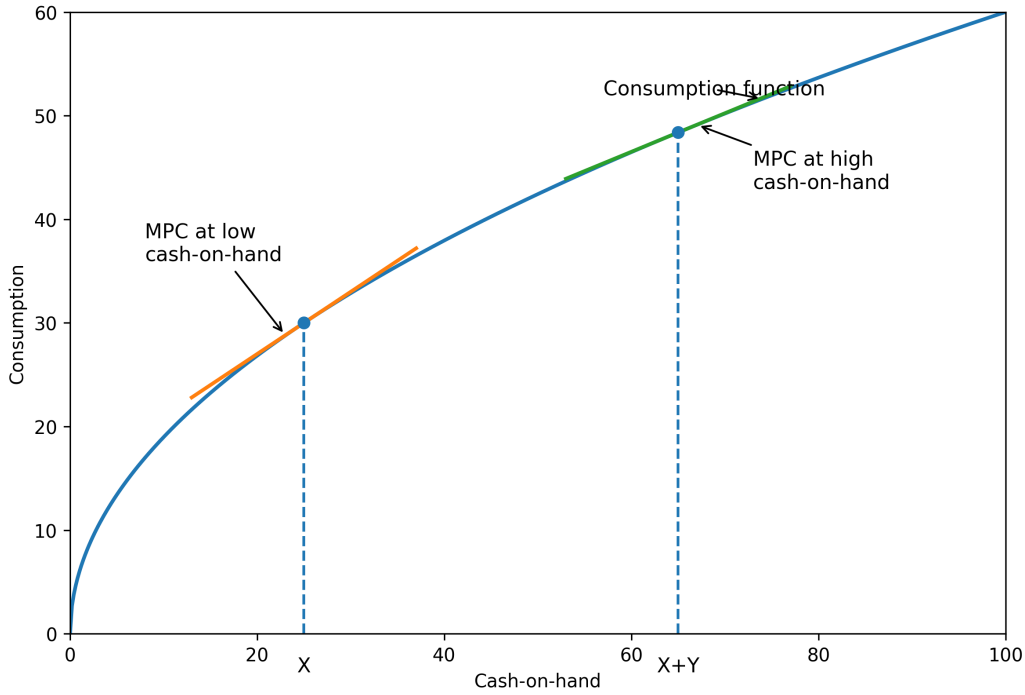
$$MPC_{it} = \alpha + \beta(X_{it} + Y_{it}) + f_i + d_t + \nu_{it} \quad (1)$$

where f_i captures persistent individual heterogeneity in preferences or other determinants of consumption behavior, d_t the effect of the business cycle and other macroeconomic variables, and ν_{it} captures reporting error and other determinants of the observed MPC.

This simple equation is useful for discussing identification across alternative models. Under a linear consumption rule, the MPC is constant. This implies that MPC does not depend on initial cash-on-hand, and therefore $\beta = 0$. It also implies that MPC does not depend on the size of the shock, or on whether the shock is positive or negative.

Figure 1. A concave consumption function and the MPC

[3] For a finite shock, this measure is the average consumption response per unit of the change in resources. The theoretical marginal propensity to consume is instead the response to an infinitesimal change in resources. Following common usage in the survey literature, we refer to the finite-shock measure simply as the MPC. This distinction matters for interpreting shock-size effects: under concavity, the finite-shock MPC declines as the size of a positive shock increases.



Note: The figure illustrates a concave consumption function, where X denotes cash-on-hand and Y the income shock. Concavity implies a higher MPC at low levels of cash-on-hand and a lower MPC at high levels of cash-on-hand.

With precautionary saving and borrowing constraints, consumption can be represented as a nonlinear function of cash-on-hand, as shown by Deaton (1991), Carroll (1997) and Carroll and Kimball (1996). As illustrated in Figure 1, concavity of the consumption function implies that the MPC is higher at low levels of cash-on-hand and lower after resources increase to $(\mathrm{X}+Y)$, or $\beta < 0$ in equation (1). Second, for a given positive shock, households with lower cash-on-hand have larger consumption propensities, $MPC(X_{low}, Y) > MPC(X_{high}, Y)$. Third, holding initial resources fixed, the average slope over a large positive shock is smaller than the average slope over a small positive shock, $MPC(X, Y_{small}) > MPC(X, Y_{large})$.

Fourth, for positive and negative shocks of equal absolute size $Y > 0$, define the respective MPCs as $MPC^{+}(X,\mathrm{Y}) = \frac{C(X+Y)-C(X)}{Y}$ and $MPC^{-}(X,\mathrm{Y}) = \frac{C(X-Y)-C(X)}{-Y}$ .Concavity implies $MPC^{-}(X,\mathrm{Y}) \geq MPC^{+}(X,\mathrm{Y})$ provided both shocks are feasible. Thus, an income loss induces a consumption reduction at least as large in absolute value as the consumption increase induced by an equally sized gain.

These predictions provide a natural way of using survey questions to learn about the shape of the consumption function. By varying cash-on-hand, shock size, or the sign of the shock across individuals, over time, or experimentally, alternative survey designs generate different sources of variation with which to test whether consumption is indeed concave. The same logic can be extended to the timing of consumption responses: if the effect of a shock is distributed over several periods, the relevant object becomes the intertemporal MPC, and its profile may itself vary with the size and nature of the shock. The empirical problem associated with equation (1) is to obtain exogenous variation in $Z_{it} = X_{it} + Y_{it}$, while dealing with persistent heterogeneity and reporting error. The different survey designs provide different solutions to this problem.

2.1. Cross-sectional data: the same shock for everyone

The simplest survey design asks every individual the same question. Suppose that in a cross-section everyone is asked how much they would consume if they unexpectedly received Y, so that the empirical MPC is simply the ratio between the reported spending and the shock Y. Then

$$MPC_i = \alpha + \beta(X_i + Y) + f_i + \nu_i.$$

Because Y is common across individuals, all identifying variation comes from X_i. This design therefore allows us to ask whether households with different initial resources report different MPCs.

The central problem of this approach is that X_i is not randomly assigned. If $\mathrm{Cov}(X_i, f_i) \neq 0$, then OLS does not identify the slope of the consumption function. Instead, the estimated and true parameters are linked by the following relation:

$$\operatorname{plim} \hat{\beta}_{OLS} = \beta + \frac{\mathrm{Cov}(X_i, f_i)}{\mathrm{Var}(X_i)}.$$

Households with high cash-on-hand may differ from households with low cash-on-hand in patience, risk aversion, income expectations, financial sophistication, family composition, or reporting styles.[4] Any of these differences may independently affect their reported MPC. Thus a

[4] As an example, suppose that f_i represents unobserved differences in rates of time preference, implying that people with high values of f_i have high tastes for current consumption. Since individuals with high rates of time preference have a tendency to report high MPC and may be more likely to have low cash-on-hand, we expect $cov(X_{it}, f_i) < 0$.

cross-sectional relationship between cash-on-hand and the MPC is informative about heterogeneity, but it cannot by itself establish that moving a household to a higher level of cash-on-hand would lower its MPC.

An early application by Jappelli and Pistaferri (2014) uses the Italian SHIW (Survey of Household Income and Wealth).[5] The cross-sectional relation is negative: the MPC falls by approximately 30 percentage points between the lowest and highest percentiles of the cash-on-hand distribution. This is suggestive of a concave consumption function: households with few liquid resources report substantially larger MPCs than households with large buffers. But the design does not allow us to distinguish genuine curvature from persistent differences across households. This application illustrates both the attraction and the limitation of simple survey MPC questions. They reveal economically important heterogeneity with relatively little survey space, but identification relies entirely on naturally occurring differences between respondents.

2.2. Panel data: the same shock in different periods

A first way to improve identification is to repeat the same hypothetical question for the same respondents at different dates. Suppose

$$MPC_{it} = \alpha + \beta(X_{it} + Y) + f_i + d_t + v_{it}.$$

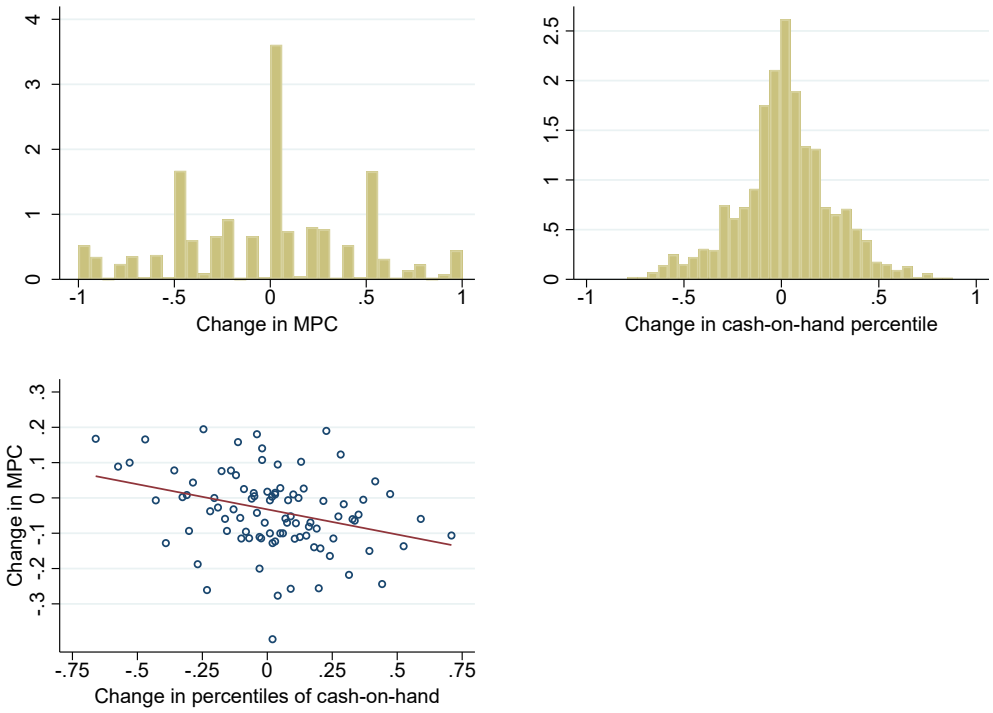
Because Y is unchanged across waves, taking first differences eliminates the individual effect f_i:

$$\Delta MPC_{it} = \beta \Delta X_{it} + \Delta d_t + \Delta v_{it}.$$

[5] Drescher et al. (2020) use harmonized data from the Household Finance and Consumption Survey to elicit MPCs out of a hypothetical windfall across European countries. The common question makes it possible to compare reported consumption responses across countries and household characteristics. Albuquerque and Green (2023) show that for a common hypothetical £500 transfer, reported MPCs in the U.K. vary systematically with households' expectations about their future financial situation, suggesting that subjective financial concerns contain information beyond standard balance-sheet measures.

Identification now relies on within-person changes in cash-on-hand rather than comparisons across households. Differencing eliminates permanent differences in preferences and reporting styles, but changes in household resources may still be correlated with evolving income expectations, precautionary motives, or other determinants of consumption. Fixed effects therefore improve identification without necessarily establishing a causal effect of cash-on-hand on the MPC. Moreover, because the hypothetical shock remains unchanged across waves, this design cannot identify shock-size effects.

Figure 2. Within-household changes in MPC and cash-on-hand



Note: The top panels show the distributions of changes in reported MPCs and changes in the percentile of the cash-on-hand distribution between survey waves. The bottom panel plots the change in the reported MPC against the change in the household's cash-on-hand percentile, together with a fitted regression line. The negative relationship illustrates that MPCs tend to decline when the same household moves to a higher position in the cash-on-hand distribution. Source: Jappelli and Pistaferri (2020).

This point was made in a subsequent paper by Jappelli and Pistaferri (2020), using panel data from the 2010 and 2016 SHIW. In the cross-section, the estimated coefficient is -0.266 , whereas in the fixed-effects estimate is -0.153. As illustrated in Figure 2, within-household increases in the cash-on-hand distribution are associated with declines in reported MPCs. The attenuation of the coefficient in the fixed-effects estimates suggests that part of the cross-sectional gradient reflects persistent differences across households. At the same time, the negative within-person relationship remains substantial, strengthening the interpretation that the consumption function itself is concave.

2.3. Panel data with different shocks in different periods

A natural extension is to change the hypothetical shock across waves. Suppose there are only two periods and individual *i* faces hypothetical shock Y_1 in period 1 and Y_2 in period 2:

$$MPC_{i1} = \alpha + \beta(X_{i1} + Y_1) + f_i + d_1 + \nu_{i1}$$
$$MPC_{i2} = \alpha + \beta(X_{i2} + Y_2) + f_i + d_2 + \nu_{i2}$$

where d_t is a time effect common to all respondents in period t. Differencing gives

$$\Delta MPC_i = \beta\Delta X_i + \beta(Y_2 - Y_1) + (d_2 - d_1) + \Delta\nu_i$$

The individual fixed effect disappears. In principle, this design appears to provide two sources of identifying variation: changes in cash-on-hand, ΔX_i, and changes in the hypothetical shock, $(Y_2 - Y_1)$.

There is, however, an important design issue. If every respondent faces Y_1 in period 1 and Y_2 in period 2, then the change in the hypothetical shock is common to all respondents and occurs at exactly the same time as the change in the period effect. Hence $\beta(Y_2 - Y_1)$ is perfectly collinear with $(d_2 - d_1)$. Once time effects are included, the effect of the change in Y cannot therefore be separately identified. Identification of β again comes only from within-person variation in cash-on-hand, ΔX_i, just as in a panel in which the same hypothetical shock is repeated over time.[6]

[6] Although the average shock-size effect cannot be separated from a common time effect, differences in shock-size effects across household groups may still be identified through interactions with household characteristics. Their interpretation, however, requires additional assumptions ruling out differential time trends correlated with those characteristics

This point has important implications for survey design. Changing the hypothetical question across waves does not automatically generate usable experimental variation. To distinguish the effect of shock size from common time effects, the treatment must vary across respondents within the same wave (for example through random assignment) or respondents must be assigned different sequences of shocks over time. More generally, survey variation contributes to identification only when it varies independently of the other dimensions of the data.

Andreolli and Surico (2026) provide a useful application of this design using hypothetical questions from the SHIW. Respondents in the 2010 wave are asked about a relatively small unexpected gain, while in 2012 they face a larger hypothetical gain. They show that the relation between MPC and shock size is heterogeneous: households with low cash-on-hand have higher MPCs out of small gains, whereas more affluent households respond more strongly to large gains. The former pattern is consistent with borrowing constraints and income risk, while the latter can be rationalized by non-homothetic demand for non-essential goods. However, because shock size changes between survey waves, its estimated effect cannot be separated from common time effects without additional assumptions. The application therefore illustrates both the potential insights and the identification limitations of varying hypothetical shocks over time.

2.4. Repeated questions in the same cross-section

Instead of changing Y over time, one can ask each individual several hypothetical questions during the same interview. Drop the time subscript, and suppose each respondent reports an MPC for Y_1 and another for Y_2:

$$MPC_i(Y_1) = \alpha + \beta(X_i + Y_1) + f_i + \nu_i(Y_1),$$

$$MPC_i(Y_2) = \alpha + \beta(X_i + Y_2) + f_i + \nu_i(Y_2).$$

Taking the within-person difference gives

$$MPC_i(Y_2) - MPC_i(Y_1) = \beta(Y_2 - Y_1) + \nu_i(Y_2) - \nu_i(Y_1).$$

Both X_i and f_i disappear. Consequently, β identifies the difference in the average MPC of the same individual:

$$\beta = \frac{E[MPC_i(Y_2) - MPC_i(Y_1)]}{Y_2 - Y_1},$$

provided that the difference in reporting errors has zero mean. The design is efficient because every respondent serves as his or her own control. Persistent preferences, household characteristics, and initial cash-on-hand are automatically held constant.

The strength of the design is also its main weakness. Respondents know what they answered to the first question when answering the second. The first scenario may therefore anchor the response to the second. Respondents may also infer that the researcher expects the two answers to differ, or they may mechanically preserve a fixed euro amount rather than reconsider their consumption decision. Randomizing the order in which Y_1 and Y_2 are presented provides an important diagnostic.

The repeated-question design is particularly attractive for comparing positive and negative shocks because naturally occurring gains and losses of precisely the same size are difficult to observe. Bunn et al. (2018), using the Bank of England/NMG Survey, vary the size of hypothetical unexpected windfalls and bills across respondents. They find larger spending responses to unexpected negative income shocks than to positive shocks of comparable size, consistent with a nonlinear consumption response and the presence of liquidity constraints or other forms of state dependence. Fuster et al. (2021) use **s**pecial modules added to the Federal Reserve Bank of New York's Survey of Consumer Expectations (SCE) with repeated hypothetical scenarios involving gains, losses, news about future income changes, and an interest-free loan to test competing models of consumption behavior. Their design illustrates how hypothetical questions can be used not merely to elicit an MPC, but to distinguish among mechanisms such as liquidity constraints and adjustment costs.

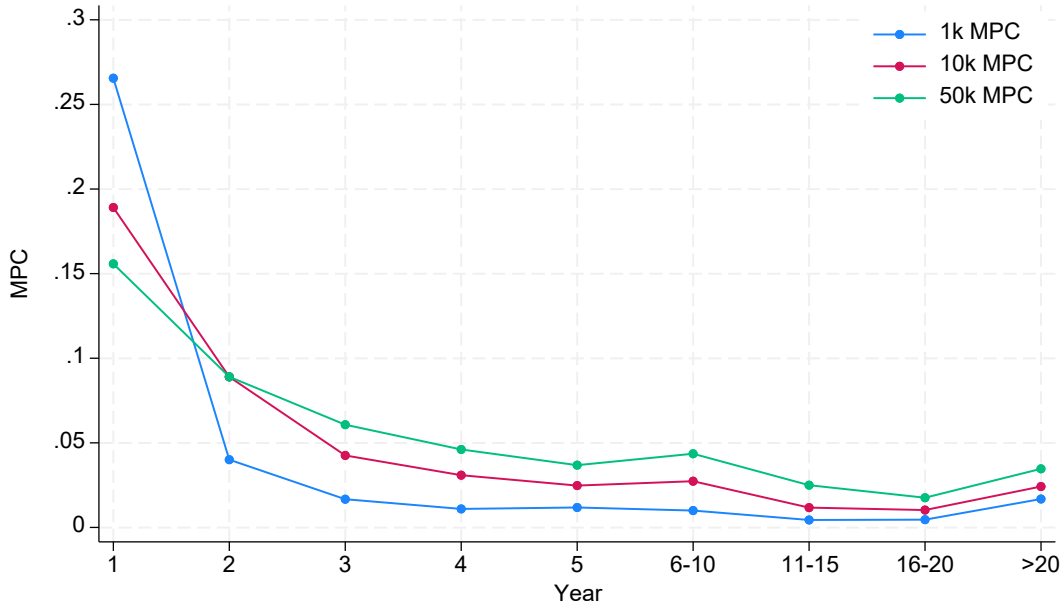
In Dutch survey data, respondents are asked how they would react to an unexpected one-time bonus, and how they would respond to an unexpected one-time tax (Christelis et al., 2019). The questions are then repeated in the same survey for different shock sizes. The paper shows that transitory shocks generate sizeable consumption responses, of roughly 15–25%. Second,

consumption responds more strongly to negative shocks than to positive shocks. Third, smaller shocks produce larger MPCs than larger shocks. These results correspond to the predictions of a concave consumption function, as shown by simulations of a model with precautionary saving and liquidity constraints.

The same logic can be extended beyond two shock sizes. The Italian Survey of Consumer Expectations asks respondents how consumption would respond over different horizons to transfers €1,000, €10,000 and €50,000. The design therefore generates a response surface indexed by both shock size and horizon.

$$iMPC_{i,j,t} = \frac{\Delta C_{i,j,t}}{\partial Y_j}.$$

Figure 3. Intertemporal MPCs by shock size



Note: The figure reports additional spending attributable to hypothetical transfers of €1,000, €10,000, and €50,000 in each period, expressed as a share of the transfer amount. Smaller transfers generate higher initial MPCs that decline more rapidly, while larger transfers generate lower initial MPCs but more persistent responses. The horizontal axis shows years since receipt of the transfer; later horizons are grouped into intervals. Source: Jappelli, Savoia and Sciacchetano (2026).

The resulting intertemporal MPCs show that small shocks have especially large short-run effects (Jappelli et al., 2026). As illustrated in Figure 3, for the €1,000 shock, the first-year response is approximately 25% and then declines sharply. Large shocks have smaller initial MPCs but more persistent effects; for the largest shock, the initial response is around 15%, while the longer-horizon MPC remains higher than for the small shock. These statistics can be compared directly with quantitative consumption models. The paper shows that one-asset models with income risk and borrowing constraints reproduce much of the empirical pattern: small shocks generate large but rapidly declining responses, whereas large shocks have smaller but more persistent effects.

Repeated scenarios are also useful when actual shocks are particularly difficult to isolate. For house prices, respondents can be asked how they would respond to a positive 10% change in home value and to a negative 10% change (Christelis et al., 2021). The average MPC out of housing wealth is relatively small, around 2–4%, and more than 90% of homeowners report no adjustment. Yet responses are heterogeneous: the MPC is negatively associated with cash-on-hand, and homeowners with loan-to-value ratios above 100% respond more strongly to negative house-price shocks. The key benefit of the survey design is that positive and negative wealth shocks can be made identical in absolute size while all other conditions are held fixed.

2.5. Randomized cross-sectional survey experiments

The final design assigns different hypothetical scenarios randomly across respondents. Consider a cross-sectional design in which one random group receives a hypothetical shock Y_1 and the other a shock Y_2:

$$MPC_i = \alpha + \beta\left(X_i + Y_j\right) + f_i + \nu_i, \qquad j = 1,2.$$

Because Y_j is randomly assigned, $E[X_i \mid Y_1] = E[X_i \mid Y_2]$ and $E[f_i \mid Y_1] = E[f_i \mid Y_2]$, the treatment effect can therefore be recovered by comparing mean responses:

$$\beta = \frac{E[MPC_i \mid Y_2] - E[MPC_i \mid Y_1]}{Y_2 - Y_1}.$$

Randomization addresses the central identification problem of the simple cross-section by making treatment assignment independent of household characteristics and persistent unobserved

heterogeneity. Comparisons across treatment groups therefore identify the causal effect of the experimentally varied scenario on reported consumption responses. Interpreting these differences as features of the underlying consumption function additionally requires that reporting errors do not vary systematically across treatments. Whether the same comparative statics hold for actual consumption decisions is a separate question of external validity, addressed in Section 3.

Compared with repeated within-person scenarios, randomized cross-sectional experiments avoid anchoring, contrast effects, and attempts to maintain consistency across answers, since each respondent evaluates only one treatment. However, individual heterogeneity is balanced across groups only in expectation rather than eliminated through differencing. Consequently, randomized experiments generally require larger samples, particularly when estimating heterogeneous treatment effects. Combining the two approaches, where feasible, offers a useful check on the robustness of the findings.

In the MPC context, there are several possible applications of this design. For instance, respondents can be randomly assigned different lottery prizes or fiscal transfers. Comparisons across groups then identify shock-size effects without asking any individual to compare multiple amounts, providing a direct test of the prediction that the MPC falls as the size of a positive shock rises. In other applications the expected size of a transfer can be held fixed while its variance is experimentally changed. In Georgarakos et al. (2026), for example, one group can receive €10,000 with certainty, while another faces equal probabilities of €8,000 and €12,000, and still other groups face progressively wider mean-preserving spreads. Because the expected transfer is identical, differences in reported spending identify the effect of transfer risk rather than the effect of expected income.

Randomization can also alter the informational and social environment of the shock, as in Jappelli and Pistaferri (2026). Consider a €1,000 workplace bonus. Respondents can be randomly assigned to scenarios in which no additional information is provided; only the respondent receives the bonus and only the respondent knows; only the respondent receives it but coworkers know; or everybody at the workplace receives the bonus and everyone knows. The amount of the income shock is unchanged. What varies is its observability and whether it is idiosyncratic or common. The experiment therefore isolates social channels that would be extremely difficult to identify from conventional observational data.

The results indicate an MPC of 0.33 in the baseline and private treatments, 0.30 when the individual bonus is publicly observable, and 0.36 when coworkers also receive the bonus. The first comparison suggests social restraint, while the second suggests peer amplification. The broader lesson is that survey randomization changes the question researchers can ask: instead of merely estimating the average response to an income shock, one can estimate the causal effect of individual components of the economic environment surrounding the shock.

A useful extension of the randomized survey approach is to study behavioral margins other than consumption. Christelis et al. (2025) use randomized hypothetical wealth shocks of different sizes to examine how households adjust portfolio choice, including stock-market participation and the share invested in risky assets. Georgarakos et al. (2025) use a similar randomized design to study labor-supply responses to windfall gains, considering both the extensive margin of employment and the intensive margin of hours worked, as well as job-search behavior. These studies do not elicit an MPC directly, but they illustrate the broader usefulness of randomized hypothetical shocks: by varying the size of the windfall across respondents, the researcher can trace how different behavioral responses change with the magnitude of the shock while preserving exogenous treatment variation. In this sense, they extend the same experimental logic used for MPCs to other margins of household adjustment and provide additional evidence on how nonlinear responses to wealth shocks can be identified in survey data.

2.6. Non-classical measurement error

The comparison of survey designs would be incomplete without considering measurement error. In a cross-section, let the true, latent MPC be

$$MPC_i^* = \alpha + \beta Z_i + f_i$$

where Z_i denotes the determinant of the MPC used for identification in a particular design (for example cash-on-hand or shock size), f_i captures unobserved individual heterogeneity, and β is the parameter of interest. The MPC observed in the survey is

$$MPC_i^{\text{obs}} = MPC_i^* + \nu_i$$

where ν_i is reporting or measurement error. Combining the two equations gives

$$MPC_i^{\text{obs}} = \alpha + \beta Z_i + f_i + \nu_i$$

If response error is classical, so that $\mathrm{E}(\nu_i \mid Z_i, f_i) = 0$, it does not introduce by itself an additional bias in the estimated effect of Z_i, conditional on the identifying assumptions of the particular survey design. Because the measurement error affects the dependent variable rather than the regressor, its main consequence is to increase residual variance and hence the standard errors of the estimated coefficients. Classical response error does not, however, eliminate other sources of bias. For example, in a cross-sectional design, correlation between Z_i and f_i may still prevent identification of β.

The more serious concern is non-classical response error, in which reporting errors vary systematically with the treatment or with household characteristics. Suppose, for example,

$$\nu_i = \lambda Z_i + e_i$$

with

$$\mathrm{E}(e_i \mid Z_i, f_i) = 0$$

Substituting into the observed MPC equation gives

$$MPC_i^{\text{obs}} = \alpha + (\beta + \lambda) Z_i + f_i + e_i$$

Thus, abstracting from other sources of endogeneity, the coefficient estimated using the reported MPC is

$$\beta^{\text{obs}} = \beta + \lambda$$

If respondents systematically understate the fraction consumed as the hypothetical shock becomes larger, then $\lambda < 0$. An estimated negative shock-size effect could therefore arise even if the true MPC were independent of shock size, so that $\beta = 0$. In this case, the observed decline in the MPC would reflect reporting behavior rather than genuine curvature of the consumption response.

This issue cuts across all five data structures, although its manifestations differ. In a simple cross-section, reporting errors may be correlated with household characteristics such as education, financial literacy, or income, making it difficult to distinguish genuine MPC heterogeneity from differences in the way respondents understand or answer the question. Panel data eliminate permanent components of reporting behavior through individual fixed effects, but they remain vulnerable to time-varying response errors and to changes in respondents' interpretation of the question across waves. Repeated-question designs face an additional concern: answers to later

scenarios may be influenced by earlier ones through anchoring, consistency motives, learning, or fatigue. Randomized cross-sectional designs largely avoid these within-person contamination effects, but they can still be biased if different treatments systematically affect respondents' understanding of the question or their willingness to report particular answers.

Several checks can help assess the importance of measurement and reporting problems in hypothetical MPC questions. A first set of checks focuses on the distribution of responses. Researchers can examine the sensitivity of the results to excluding focal answers, such as MPCs equal to 0, 0.5, or 1, which may reflect rounding, simplified response rules, or limited engagement with the question. These concerns are particularly relevant in repeated-scenario designs, where respondents may anchor on an earlier answer or adopt a simple rule across treatments (Fuster et al. 2021; Crossley et al. 2025).

A second set of checks concerns survey engagement and comprehension. Interview duration and item-specific response times can be used to identify unusually fast respondents, while very long response times may indicate difficulty in understanding the scenario. When respondents are asked to allocate a windfall across consumption, saving, debt repayment, and transfers to others, the components should also add up approximately to the total shock. One can therefore test whether the main results survive after excluding respondents with large adding-up discrepancies. Similarly, answers to related questions can be checked for logical consistency. For example, a respondent who reports lower euro spending after a larger positive windfall than after a smaller one may warrant separate analysis. Comprehension questions administered immediately after the scenario can provide an additional check that respondents correctly understood the size, timing, sign, and uncertainty of the hypothetical shock.

Question order is another potentially important source of non-classical response error. In repeated-question designs, the answer to one scenario may influence the response to the next through anchoring, learning, fatigue, or a desire to maintain consistency across answers. Randomizing the order in which scenarios are presented therefore provides a direct diagnostic for these effects. If the estimated treatment effects are stable across alternative orderings, concerns that they are driven by anchoring are reduced. Designs that combine within-person comparisons with randomized question order are particularly useful in this respect (Crossley et al., 2025).

A further diagnostic is to compare alternative elicitation methods. For example, respondents can be randomly assigned either a direct question asking how much of a transfer they would spend or a two-step procedure that first asks whether spending would change and only then asks by how much. Crossley et al. (2026) show that alternative ways of eliciting the MPC can generate different response distributions, highlighting that question format is itself an important component of measurement. Robustness across alternative elicitation methods therefore provides useful evidence that estimated treatment effects are not artifacts of a particular wording, response scale, or sequence of questions. Similar exercises can vary the framing of the shock, for example, describing the same payment as a lottery prize, bonus, tax rebate, or government transfer, to assess whether the substantive results depend on context.

Finally, an informative validation strategy is to compare estimates obtained from data structures that are subject to different potential biases. A shock-size gradient that appears both in repeated within-person scenarios and in a randomized cross-section is difficult to attribute solely to anchoring or respondent-specific reporting styles. Panel estimates can similarly be compared with cross-sectional estimates, as in Jappelli and Pistaferri (2020). More broadly, hypothetical responses can be confronted with consumption responses to realized shocks such as lottery winnings or fiscal transfers. Stability across alternative samples, elicitation formats, survey waves, and identification strategies does not eliminate measurement error, but it substantially strengthens the interpretation of the observed MPC patterns as genuine features of household behavior rather than artifacts of survey design.

2.7. Comparing the five designs

Table 1 summarizes the identifying variation, advantages, and limitations of the five survey designs. No single design dominates across all research questions. Their relative usefulness depends on the parameter of interest, the available sources of variation, and the potential for measurement error. The five designs should not be viewed as a simple ranking from weak to strong. Their usefulness depends on the question of interest: cross-sections are well suited to documenting heterogeneity, panels to studying how MPCs change with household resources, repeated scenarios to comparing alternative shocks for the same individual, and randomized cross-sections to identifying causal effects of treatment characteristics. Furthermore, the designs are complementary

and can be combined. Panels can incorporate randomized scenarios, and repeated-question modules can randomize question order. The appropriate design depends on whether the objective is to document heterogeneity, study changes in household resources, or compare responses to alternative shocks. Confidence is strengthened when similar patterns emerge across designs with different identifying assumptions and sources of reporting error.

Table 1. Identification and limitations of alternative survey designs

Design	**Identifying variation**	**Main advantage**	**Main limitation**
Common shock, cross-section	Differences in resources across households	Documents MPC heterogeneity	Resources may be correlated with unobserved determinants of MPCs
Common shock, panel	Changes in resources within households	Removes time-invariant household heterogeneity	Changes in resources may remain endogenous
Changing shock, panel	Changes in shocks across waves and, where available, across respondents	Can identify shock effects when variation is separable from time effects	Common shock changes are confounded with survey-wave effects
Repeated scenarios	Differences across scenarios within respondents	Compares shocks holding household characteristics fixed	Anchoring, order effects, and carryover between questions
Randomized scenarios	Differences across randomly assigned groups	Identifies causal effects on reported responses	Treatment-dependent reporting error may affect behavioral interpretation; generally less precise than within-person comparisons

3. Validating survey-based MPC evidence

The main limitation of survey-based MPCs is that they describe intended responses to hypothetical shocks rather than consumption responses to realized changes in resources. Clean identification within a survey experiment does not by itself establish that respondents would behave in the same way if the shock actually occurred. Validation is therefore an important complement to careful survey design. Two strategies are especially useful: comparison with responses to realized shocks and simulation of quantitative models calibrated using independent evidence. A complementary approach uses experimentally generated survey variation to estimate structural parameters. This provides a structural interpretation of the experimental evidence, but constitutes validation only when assessed against additional testable predictions or independent evidence.

A natural validation exercise compares hypothetical responses with consumption responses following actual, exogenous windfalls. The informative comparison is not only whether survey-based and realized MPCs have similar average levels, but whether they display similar comparative statics. Broda and Parker (2014) exploit the randomized timing of the 2008 Economic Stimulus Payments and high-frequency spending data from the Nielsen Consumer Panel to estimate the consumption response to an actual government transfer. Household spending rose sharply when the payment was received and remained elevated over the following weeks. Their design is useful as a benchmark for survey-based MPCs because the income shock is realized rather than hypothetical, while identification relies on quasi-experimental variation in the timing of receipt. An especially informative form of validation is available when hypothetical and realized-shock evidence are examined within the same study. Bunn et al. (2018) combine hypothetical questions on unexpected windfalls and bills with separate evidence on realized income changes. The similarity of the qualitative asymmetry across the two sources of variation provides a useful check that the patterns elicited by hypothetical scenarios are not purely artifacts of survey design.

Realized lottery data provide a particularly useful benchmark for validating survey-based MPCs because they combine plausibly exogenous income shocks with substantial variation in shock size. Fagereng, Holm and Natvik (2021) use Norwegian administrative panel data on lottery winners to estimate how consumption responds to prizes of different sizes and how these responses vary with household balance sheets. They find that MPCs are higher for small prizes and for households with low liquid assets, while spending responses decline as the size of the prize

increases. The response also has a clear intertemporal pattern: spending peaks in the year of winning and gradually returns toward its pre-winning level over the following years. These findings mirror two patterns obtained from hypothetical survey experiments (the decline of the MPC with liquidity and with shock size) and therefore provide useful external validation of the survey evidence. Golosov et al. (2024), using administrative data on U.S. lottery winners and an event-study design, show that exogenous wealth gains generate sizable consumption responses. Their analysis also highlights that realized windfalls can affect labor earnings, a channel that is typically held fixed by construction in hypothetical survey experiments. Together, these studies provide external evidence that the consumption responses elicited in surveys correspond to economically meaningful responses to actual changes in household wealth.

A second validation strategy asks whether quantitative consumption models can reproduce the patterns observed in survey experiments. This exercise is especially informative when the parameters of the model are calibrated independently of the survey moments being validated. [7] Consider a buffer-stock model with income risk, precautionary saving, and a borrowing constraint. Once preferences, the income process, interest rates, and borrowing limits have been specified, the model generates a consumption function. The simulated households can then be exposed to exactly the same shocks used in the survey. Rather than comparing a generic theoretical MPC with an empirical average, the researcher can compare like with like: positive and negative shocks, small and large shocks, different horizons, etc. [8]

A third approach reverses the simulation exercise. Instead of taking model parameters as given and asking whether the model reproduces survey responses, experimentally generated survey variation can be used to estimate economically meaningful structural parameters. The social-interaction experiment provides a useful example. Varying whether a workplace bonus is private, publicly observable, or received by all coworkers identifies distinct components of the

[7] For validation to be informative, the moments used to assess the model should differ from those used to calibrate it. If parameters are selected specifically to reproduce the survey MPCs, a close fit is largely mechanical. A stronger exercise calibrates the model to external moments such as income dynamics and the wealth distribution and then asks whether the survey response patterns emerge out of sample.

8 For example, Andreolli and Surico (2026) show that borrowing constraints rationalize the high MPCs of liquidity-poor households out of small gains, while non-homothetic preferences help explain the stronger response of affluent households to large gains.

consumption response. These treatment effects map into the strength of peer effects and hence into a social multiplier. The same principle can be applied to experiments varying shock size, timing, or uncertainty: treatment effects can identify parameters governing curvature, consumption smoothing, risk sensitivity, or social interactions.

These approaches provide complementary evidence. Comparisons with realized shocks assess external validity; independently calibrated models test whether the observed responses are consistent with economic mechanisms; and structural estimation helps interpret experimental treatment effects. Together, they can establish whether patterns identified in hypothetical scenarios reflect economically meaningful consumption behavior.

4. Conclusions

What can be learned from survey questions on hypothetical income shocks depends critically on the structure of the data and, in particular, on the source of identifying variation. A cross-section in which all respondents face the same shock is useful for documenting heterogeneity, but correlations between MPCs and cash-on-hand are not causal if resources are correlated with preferences, expectations, or reporting styles. Repeating the same question in panel data replaces between-person comparisons with within-person changes and removes permanent individual heterogeneity, although the hypothetical shock itself remains fixed. Changing the shock across waves adds useful variation only when that variation can be separated from common time effects.

Repeated scenarios within the same interview provide a different source of identification. Because the same respondent faces several shocks, initial cash-on-hand and permanent individual characteristics are held fixed by construction. This makes the design particularly effective for studying the sign and size of shocks and for eliciting intertemporal response profiles. Its main weakness is that answers can be affected by anchoring, question order, and attempts by respondents to make their answers internally consistent.

Randomized cross-sectional survey experiments address these concerns from another direction. Random assignment makes the characteristics of the hypothetical shock independent of observed and unobserved household characteristics, while each respondent answers only one treatment question. The price is that individual fixed effects are not differenced out and larger

samples are required to study heterogeneous treatment effects. Repeated-question and randomized designs should therefore be viewed as complementary rather than as a simple ranking from weaker to stronger methods.

These advantages do not eliminate the central concern with hypothetical survey evidence: stated responses need not coincide with realized behavior. Validation should therefore be part of the research design rather than an afterthought. Survey responses can be compared with realized and plausibly exogenous shocks such as lottery winnings; quantitative models calibrated to independent moments can be subjected to the same shocks used in the survey; and experimental treatment effects can be mapped into structural parameters and compared with estimates from other sources.

The broader methodological lesson is that survey MPC questions should be viewed as designed data. Their usefulness depends on how the scenarios are constructed, how treatments vary across individuals and over time, and which sources of heterogeneity or measurement error the design eliminates. The objective is not simply to ask households for an MPC, but to create variation that discriminates among economic mechanisms. This perspective suggests a natural direction for future work. Panel surveys can combine repeated measurements with randomized modules, allowing researchers to exploit both within-person variation and experimental assignment. Survey experiments can be linked to administrative, transaction, or realized-shock data for external validation. Finally, rather than targeting a single average MPC, future designs can estimate richer consumption response functions indexed by household resources and by the size, sign, timing, risk, and social distribution of shocks. In this role, surveys are not a substitute for observed behavior. They are a complementary empirical laboratory in which theories of consumption can be tested under controlled and economically interpretable variation.

References

Albuquerque, Bruno, and Georgi Green. 2023. "Financial Concerns and the Marginal Propensity to Consume in COVID Times: Evidence from UK Survey Data." *Journal of Macroeconomics* 78: 103563.

Andreolli, Michelle, and Paolo Surico. 2026. "'Less Is More': Consumer Spending and the Size of Economic Stimulus Payments." *American Economic Journal: Macroeconomics* 18 (1): 34–68.

Broda, Christian, and Jonathan A. Parker. 2014. "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption." *Journal of Monetary Economics* 68 (Supplement): S20–S36.

Bunn, Philip, Jeanne Le Roux, Kate Reinold, and Paolo Surico. 2018. "The Consumption Response to Positive and Negative Income Shocks." *Journal of Monetary Economics* 96: 1–15.

Carroll, Christopher D. 1997. "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis." *Quarterly Journal of Economics* 112 (1): 1–55.

Carroll, Christopher D., and Miles S. Kimball. 1996. "On the Concavity of the Consumption Function." *Econometrica* 64 (4): 981–992.

Christelis, Dimitris, Dimitris Georgarakos, Tullio Jappelli, Luigi Pistaferri, and Maarten van Rooij. 2019. "Asymmetric Consumption Effects of Transitory Income Shocks." *Economic Journal* 129 (622): 2322–2341.

Christelis, Dimitris, Dimitris Georgarakos, Tullio Jappelli, Luigi Pistaferri, and Maarten van Rooij. 2021. "Heterogeneous Wealth Effects." *European Economic Review* 137: 103805.

Christelis, Dimitris, Dimitris Georgarakos, Tullio Jappelli, and Geoff Kenny. 2025. "Wealth Shocks and Portfolio Choice." *Journal of Monetary Economics* 149: 103632.

Coibion, Olivier, and Yuriy Gorodnichenko. 2026. "The New Causal Macroeconomics of Surveys and Experiments." *Journal of the European Economic Association* 24 (1): 58–81.

Crossley, Thomas F., Paul Fisher, Peter Levell, and Hamish Low. 2021. "MPCs in an Economic Crisis: Spending, Saving and Private Transfers." *Journal of Public Economics Plus* 2: 100005.

Crossley, Thomas F., Peter Levell, and Sergio Sierra Vásquez. 2025. "What Would You Do with £500? (...in Your Own Words)." IFS Working Paper.

Crossley, Thomas F., Paul Fisher, Peter Levell, and Hamish Low. 2026. "Eliciting the Marginal Propensity to Consume in Surveys." Federal Reserve Bank of Chicago Working Paper 2026-04.

Deaton, Angus. 1991. "Saving and Liquidity Constraints." *Econometrica* 59 (5): 1221–1248.

Drescher, Katharina, Pirmin Fessler, and Peter Lindner. 2020. "Helicopter Money in Europe: New Evidence on the Marginal Propensity to Consume across European Households." *Economics Letters* 195: 109416.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik. 2021. "MPC Heterogeneity and Household Balance Sheets." *American Economic Journal: Macroeconomics* 13 (4): 1–54.

Fuster, Andreas, Greg Kaplan, and Basit Zafar. 2021. "What Would You Do with $500? Spending Responses to Gains, Losses, News, and Loans." *Review of Economic Studies* 88 (4): 1760–1795.

Georgarakos, Dimitris, Tullio Jappelli, Geoff Kenny, and Luigi Pistaferri. 2025. "Labor Supply Response to Windfall Gains." *Journal of Public Economics* 250: 105476.

Georgarakos, Dimitris, Tullio Jappelli, Luigi Pistaferri, and Ettore Savoia. 2026. "Consumption Responses to Uncertain Fiscal Transfers." Work in progress.

Golosov, Mikhail, Michael Graber, Magne Mogstad, and David Novgorodsky. 2024. "How Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income." *Quarterly Journal of Economics* 139 (2): 1321–1395.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. "Designing Information Provision Experiments." *Journal of Economic Literature* 61 (1): 3–40.

Jappelli, Tullio, and Luigi Pistaferri. 2010. "The Consumption Response to Income Changes." *Annual Review of Economics* 2: 479–506.

Jappelli, Tullio, and Luigi Pistaferri. 2014. "Fiscal Policy and MPC Heterogeneity." *American Economic Journal: Macroeconomics* 6 (4): 107–136.

Jappelli, Tullio, and Luigi Pistaferri. 2020. "Reported MPC and Unobserved Heterogeneity." *American Economic Journal: Economic Policy* 12 (4): 275–297.

Jappelli, Tullio, and Luigi Pistaferri. 2026. "Observability and Social Interactions in Consumption." NBER Working Paper No. 35575.

Jappelli, Tullio, Ettore Savoia, and Alessandro Sciacchetano. 2026. "Intertemporal MPC and Shock Size." *European Economic Review* 186: 105303.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles. 2006. "Household Expenditure and the Income Tax Rebates of 2001." *American Economic Review* 96 (5): 1589–1610.

Kaplan, Greg, and Giovanni L. Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82 (4): 1199–1239.

Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson, and Robert McClelland. 2013. "Consumer Spending and the Economic Stimulus Payments of 2008." *American Economic Review* 103 (6): 2530–2553.

Parker, Jonathan A., and Nicholas S. Souleles. 2019. "Reported Effects versus Revealed-Preference Estimates: Evidence from the Propensity to Spend Tax Rebates." *American Economic Review: Insights* 1 (3): 273–290.

Sahm, Claudia R., Matthew D. Shapiro, and Joel Slemrod. 2012. "Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How It Is Delivered?" *American Economic Journal: Economic Policy* 4 (3): 216–250.

Shapiro, Matthew D., and Joel Slemrod. 2003. "Consumer Response to Tax Rebates." *American Economic Review* 93 (1): 381–396.