Using Macro Counterfactuals to Assess Plausibility: An Illustration using the 2001 Rebate MPCs

Jacob Orchard
Federal Reserve Board

Valerie A. Ramey
Hoover Institution, NBER, CEPR

Johannes F. Wieland
UCSD and NBER

October 17, 2023

Abstract

Macroeconomics has increasingly adopted tools from the applied micro "credibility revolution" to estimate micro parameters that can inform macro questions. In this paper, we argue that researchers should take advantage of this confluence of micro and macro to take the credibility revolution one step further. We argue that researchers should assess the *plausibility* of the micro estimates and macro models by constructing macro counterfactuals for historical periods and comparing these counterfactuals with reasonable benchmarks. We illustrate this approach by conducting a case study of the 2001 U.S. tax rebates, as well as briefly summarizing two previous applications of the methodology. In the 2001 rebate case, we calibrate a two-good, two-agent New Keynesian model with the leading estimates of the household marginal propensity to consume (MPC) out of the rebates to construct a counterfactual path for nondurable consumption. The counterfactual path implies that without the tax rebate nondurable consumption spending would have fallen dramatically in the late summer and fall of 2001. Using forecasting regressions and other evidence, we argue that this counterfactual is implausible. When we investigate the source of the discrepancy, we find that the leading MPC estimates are not representative of the response of total consumption.

JEL codes: E21, E27, E62
Keywords: marginal propensity to consume, transfers multipliers, credibility revolution

This work was presented by Valerie Ramey as the Economic Journal Lecture "Do They Add Up? Using Macro Counterfactuals to Assess Micro Estimates and Macro Models" at the 2023 Royal Economic Society conference. We are grateful to participants at the 2023 Royal Economic Society/Scottish Economic Society meetings, the 2023 Society for Economic Dynamics meetings, the Hoover Economic Policy Working Group, and UT Austin. Valerie Ramey gratefully acknowledges financial support from National Science Foundation Grant No. 1658796. The analysis and conclusions in this paper are the work of the authors alone and do not necessarily represent the views of the Federal Reserve Board or the Federal Reserve System.

1 Introduction

In their article "The Credibility Revolution in Empirical Economics," Angrist and Pischke (2010) argued that the widespread adoption of randomized trials and natural experiments had significantly improved the credibility of the estimates in applied microeconomics. However, they reproached the macro and industrial organization fields for being slow to adopt the new methods. Indeed, the dominant macro methods of the 1990s and 2000s were time series models and calibrated or estimated dynamic stochastic general equilibrium (DSGE) models. A few macroeconomists were exploiting natural experiments such as wars, hyperinflations, and the timing of income to estimate parameters of interest to macro, but they were the exception rather than the rule.

During the Great Recession, the applied micro methods quickly diffused into macro, in large part because of the economic questions and natural experiments generated during that period. For example, macro researchers used natural experiments to estimate household consumption responses to the temporary rebates enacted to combat the recession. They also used Bartik-style instruments to estimate fiscal multipliers using variations in spending across states, the effects of house price declines on local economies, and many other questions that could exploit this methodology.

Macroeconomists quickly realized, however, that the parameter estimates obtained with applied micro techniques could not be used by themselves to answer macroeconomic questions. For example, in the case of fiscal estimates it was recognized that the government spending multipliers estimated at the state level are only *relative* multipliers and therefore cannot be used to answer aggregate questions without more assumptions (e.g. Ramey (2011)). When Nakamura and Steinsson (2014) estimated state-level government spending multipliers using data on a panel of states, they also constructed an open economy New Keynesian model of a currency and fiscal union to derive the closed economy aggregate multipliers implied by their estimates. It was for this reason that Ramey (2019b) concluded that there is no "applied micro free lunch" for macroeconomists: to generate estimates that answer macro questions, researchers must either start with macro data or they must use a macro model to translate estimates obtained from household or subregional variation to a macro context.

Even with this caveat, the additional applied micro tools are a valuable addition to the macro toolbox. Sometimes, though, there is a tension between the estimates generated using the different methods. For example, government purchases multipliers

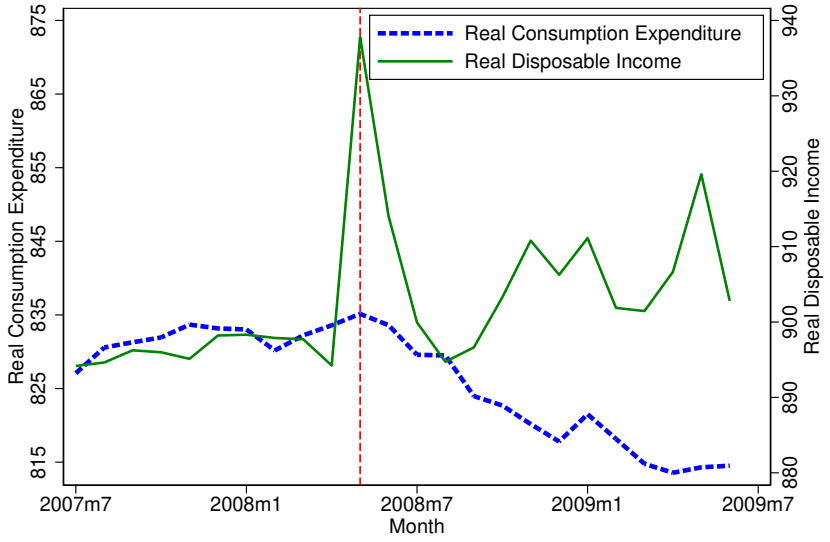
are usually estimated to be unity or below when aggregate times series data or estimated DSGE models are used. In contrast, state-level multiplier estimates fed through popular New Keynesian DSGE models typically imply multipliers of 1.5 or above.

In this paper, we make the case for using the recent micro-macro confluence to take the "credibility revolution" one step further. In particular, we suggest that researchers should assess the micro estimates and macro models by constructing macro counterfactuals for historical periods and analyzing whether they are *plausible*. This type of assessment is useful for several reasons. First, when micro estimates and macro estimates differ significantly, a plausibility test can be used to favor some estimates over others. Too often, policymakers and researchers are faced with a bewildering array of estimates, which might lead to the temptation to choose the most convenient estimate. Second, a plausibility test helps determine which estimates and/or models deserve more scrutiny, which is a useful guide for allocating scarce researcher time. Third, the attempt to reconcile the micro estimates and macro models with the behavior of aggregates is often illuminating, as we demonstrate in this paper, in Orchard et al. (2023), and Ramey (2019a,b).

As an example of discrepancy between estimates, consider the case of the 2008 U.S. tax rebates. Figure 1 shows the behavior of aggregate disposable income and consumption in 2008. Disposable income shows a pronounced spike in late spring and summer 2008 when the rebates were distributed. Consumption, in contrast, shows a tiny bump. Feldstein (2008) and Taylor (2009) analyzed these data using simple graphs and regressions and concluded that the marginal propensity to consume from the 2008 rebate was low.

A few years later, however, Parker et al. (2013) and Broda and Parker (2014) presented evidence of substantial MPCs out of the 2008 rebate. They added rebate questions to two large existing surveys of households, the Consumer Expenditure Survey and the Nielsen survey, and exploited the natural experiment involved in the timing of the distribution of rebates. Using then state-of-the-art difference-in-difference applied micro methods, they estimated very high marginal propensities to consume out of the 2008 rebate, ranging from 50% to 90% for total consumption with expenditures on motor vehicles accounting for a significant part. Policymakers and most academics believed the estimates generated with the applied micro estimates and ignored the previous macro analysis. No one tried to explain how such high MPCs could be consistent with the aggregate data shown in Figure 1.

Figure 1. The Impact of the 2008 Tax Rebates on Disposable Income and Consumption



Sahm et al. (2012) did question the size of Parker et al.'s (2013) MPCs for motor vehicles. In the last table of their paper — Table 14 — they computed the implied motor vehicle spending induced by the rebates, assuming no partial or general equilibrium effects. They noted that the Parker et al. (2013) estimates imply that the rebates accounted for one-third of the total spending on new motor vehicles in the second and third quarters of 2008. However, their statement underestimates the size of the induced expenditure because it averages over several quarters. In Orchard et al. (2023), we used the monthly calculations from their table to create counterfactual paths of new motor vehicle consumption expenditures based on the estimated rebate-induced expenditures and found counterfactuals that implied motor vehicle spending would have declined by 87% in the summer of 2008 had there been no rebate, but then recovered when Lehman Brothers failed. We argued that such an unprecedented decline is implausible.

2 Methodology for Counterfactual Plausibility Analysis

The methodology for assessing the plausibility of micro estimates and macro models combines several key ingredients in order to connect the micro and the macro. The first ingredient is effects estimated at a lower level of aggregation than the desired

level of analysis. In some cases these are micro parameters that govern the behavior of individual households or firms. Examples include the household MPCs discussed above, Frisch elasticities of labor supply for workers, and firm-level supply elasticities. In other cases, they might be local equilibrium effects estimated using variation across regions, such as cities or states. Examples of this type include fiscal multipliers at the city or state level and effects of regional housing price fluctuations on city-level consumer spending.

The second necessary ingredient is an event that is big enough to see in the aggregate data. This part is important because of the need to assess whether the aggregate counterfactual is plausible. The policy event need not coincide with the variation in the data used to generate the micro estimates. For example, one could use estimates of MPCs from experiments in one time period or dataset to construct macro counterfactuals for another period as long as the characteristics of the experiments are similar enough from a micro perspective. For example, in the case of MPCs the income shock should have similar duration (transitory versus persistent), the characteristics of the households affected should be broadly similar, and the level of financial development of the economy should be similar. Although we focus mostly on macro counterfactuals, the methodology can also be used at the industry level if the event is big relative to the industry and the desired level of analysis is at the industry level. In fact, our analysis of the effects of the 2008 rebates in Orchard et al. (2023) included a counterfactual for motor vehicle sales, so we incorporated an industry level analysis in addition to the macro level analysis.

The third necessary ingredient is a model that translates the micro estimates to dynamic aggregate effects. If one is interested in macro effects, then a full dynamic general equilibrium model is required. If only industry or local geographic effects are of interest, an industry equilibrium or local area equilibrium model is sufficient. Careful calibration of the model is important because of the role it plays in generating the counterfactual. The calibration might lead to general equilibrium amplification or dampening, which changes the implied counterfactual relative to the one that holds those forces constant.

The fourth necessary ingredient is a narrative analysis of the time period surrounding the policy or event to determine the plausibility of the constructed counterfactual. Part of this narrative is a search for other possible forces that could have been important at the time and could justify the implied counterfactual path. The assessment of plausibility often involves auxiliary evidence and regressions, such as forecasting equations.

The required narrative analysis typically varies significantly across case studies because the historical context is different.

With those ingredients in hand, the methodology proceeds as follows. The first step uses the micro estimates to calculate the induced changes in aggregates, not taking into account partial or general equilibrium effects. This step is important for determining the pure micro effects without any general equilibrium amplification or dampening. These induced changes are subtracted from the actual aggregate data in order to construct the *micro counterfactual*. The second step uses an appropriate model to account for the full partial or general equilibrium effects of the policy or event on the aggregate quantities. An event or policy with the same characteristics as the actual one is fed through the model to simulate the general equilibrium effects on the key variables. For example, a path of stimulus checks of the same size relative to GDP and the same timing is fed through the model and the impulse responses of the endogenous variables are calculated. These simulated deviations from steady state are then subtracted from the actual data to construct the *macro counterfactual*. The third step uses narrative analysis and any necessary auxiliary evidence to assess whether the counterfactual derived in the second step is plausible.

We apply our method to a case study of the 2001 tax rebate in the U.S. to illustrate our method. We also briefly summarize two previous applications, the 2008 rebate as analyzed in Orchard et al. (2023) and the cross-state fiscal multipliers analyzed in Ramey (2019a). As the three examples show, the procedures, reconciliation, and lessons learned can vary widely across applications.

3 Case Study: The 2001 Tax Rebates

We illustrate how researchers can use macro counterfactuals to assess micro estimates by using the 2001 tax rebates as a "case study." The most cited micro estimates of the effect of the 2001 tax rebate imply that 40 to 66 percent of the tax rebate was spent on nondurable goods and services (henceforth "nondurables") within three to six months of rebate receipt (Johnson et al. (2006)). These estimates played a prominent role in a number of government reports (e.g. Congressional Budget Office (2009), Council of Economic Advisers (2010)) and they inspired important new two-asset heterogeneous agent models of consumer behavior (e.g. Kaplan and Violante (2014), Huntley and Michelangeli (2014)).

We reassess the plausibility of these MPC estimates using our new method. First, we take the existing MPC estimates at face value and use them to construct counterfactual paths for nondurable spending had there been no tax rebate during the fall of 2001. We construct both a micro counterfactual, which does not account for general equilibrium effects, and a macro counterfactual, which allows for general equilibrium effects. We then use narratives and forecasts to argue that neither counterfactual path is plausible. Finally, we determine how to reconcile the implausible macro counterfactuals with the high micro MPCs. In the case of the 2001 tax rebates, we find that the estimated MPCs on nondurables in Johnson et al. (2006) are not robust to alternate definitions of the nondurables category. When we use the BEA's standard definitions of nondurable goods and services, the estimated MPC on nondurables falls to 0.09 and is not statistically significant. Because the new estimates lead to very plausible counterfactuals, we do not explore further whether the macro model requires adjustment.

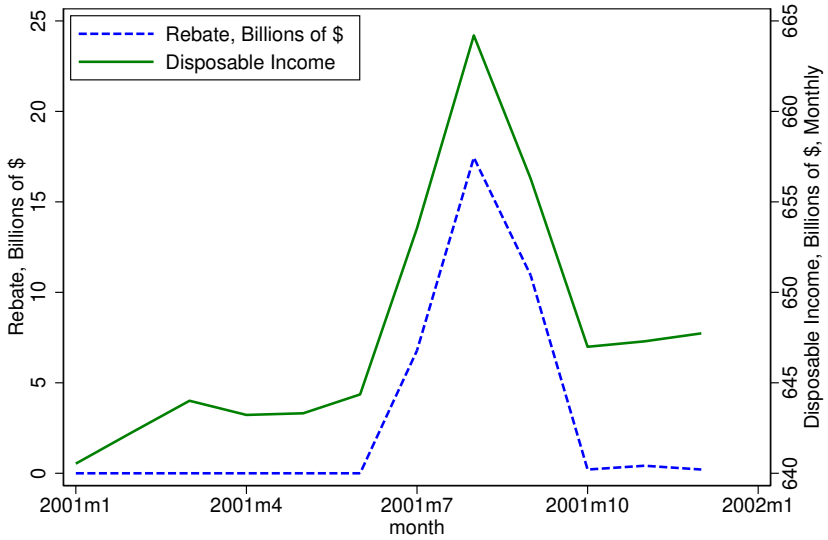
3.1 Background

The 2001 tax rebates are unique among U.S. tax rebates of the 21st century in that they were accompanied by a semi-permanent tax cut that would raise the take-home pay of households for the next ten years.[1] In June 2001, President Bush signed the Economic Growth and Tax Relief Reconciliation Act, which reduced the rate on the bottom tax bracket from 15 to 10 percent. All taxpayers that filed for the 2000 tax year were given a "rebate" equal to the net difference between the 2000 and 2001 tax rates. Most rebate checks were either $300 or $600, depending on filing status. There was no phase out for higher income taxpayers, resulting in two-thirds of US households receiving a check. The rebates totaled $38 billion, equal to 6 percent of monthly disposable income, and the checks were distributed primarily from late July through late September of 2001. The timing of the distribution across households was effectively random since it depended on the last two-digits of the recipient's social security number.

Figure 2 shows the aggregate amount of tax rebates paid (blue dashed line) alongside nominal disposable income (green solid line). Household disposable income rose essentially one-for-one with the distributed rebates, with the effect being clearly evident in the aggregate data.

1. The tax change ended up becoming permanent following legislation in 2010

Figure 2. 2001 Tax Rebates: Aggregate Payments and Nominal Income

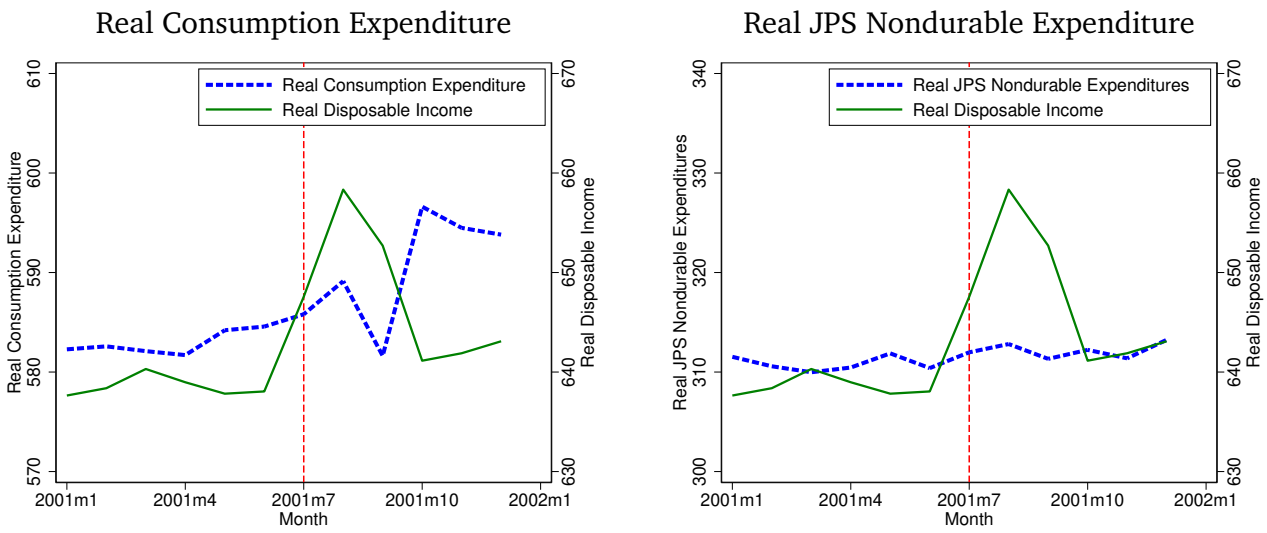


Various studies analyzed how the increase in household disposable income from the 2001 tax rebates translated into household spending. Shapiro and Slemrod (2003a) appended a question to the 2001 and 2002 Michigan Survey of Consumers asking households if they received the rebate and whether they planned on mostly spending, mostly saving, or mostly paying off debt with the rebate. Among rebate recipients, 22 percent said they would mostly spend the rebate. Subsequent work by Johnson et al. (2006) utilized the random timing of rebate receipt to conclude that households spent 40 percent of the rebate checks on nondurable consumption in the three months after receiving the rebate and then another 30 percent in the subsequent three month period. Johnson et al. (2006) found no statistically significant effect of rebate receipt on durable consumption. They nevertheless concluded that the 2001 tax rebates led to a significant increase in aggregate consumption demand.

Figure 3 shows the path of total real personal consumption expenditures (Panel A) and the nondurable component analyzed by Johnson et al. (2006) (Panel B) along with real disposable income. While the spike in disposable income is clearly visible there is no corresponding spike in total consumption or this particular nondurable subcategory.[2]

2. There is a spike in total PCE in October 2001, which reflects a spike in motor vehicle expenditures. The Federal Reserve Greenbook and Shapiro and Slemrod (2003b) attribute this spike to the itnroduction of 0% financing.

Figure 3. 2001 Tax Rebates: Real Disposable Income and Consumer Expenditures



Notes: Consumption Expenditure is total PCE as defined by the BLS. JPS Nondurable Expenditure is the subset of consumer expenditures that are included by Johnson et al. (2006) in their measure of nondurable expenditure.

Shapiro and Slemrod (2003b) also studied the personal saving rate data around the time of the rebates and compared the spike to historical spikes in the personal saving rate. They concluded that the spike in the saving rate was consistent with all the rebate being saved in July and August. The September high saving rate is also consistent with that story but 9/11 complicates the narrative.

Our analysis formalizes the tension between the large MPCs estimated by Johnson et al. (2006) and the non-response of the corresponding aggregate expenditures, and provides a resolution of this tension.

4 Counterfactuals

What would have been the path of consumption in 2001 had there been no tax rebates? We use the micro estimates for the nondurable MPC from Johnson et al. (2006) to simulate two different counterfactual paths for nondurable consumption expenditures in the absence of the 2001 rebate payments. The micro counterfactual simply subtracts the induced spending implied by the micro MPC estimates from aggregate nondurable consumer expenditures using BEA data, similar in spirit to Sahm et al. (2012). We then create a macro, or general equilibrium (GE), counterfactual by feed-

ing the rebate path into a two-good two-agent New Keynesian (TG-TANK) model calibrated to the estimated nondurable micro MPC from Johnson et al. (2006), simulating the impulse response of aggregate nondurable consumption, and subtracting it from the data. In contrast to the micro counterfactual, this macro counterfactual captures general equilibrium effects implied by the TG-TANK model.

4.1 Micro Counterfactual

We begin by constructing the micro counterfactual. Johnson et al.'s (2006) micro MPC estimates are based on a subset of household spending that is a mix of nondurable goods, some services, and some durable goods. This category, which we abbreviate *JPS ND*, accounted for only 53 percent of total personal consumption expenditures (PCE) in 2001. To be consistent across the micro and macro levels, we create our counterfactuals using aggregate BEA data on consumer expenditure based on the same subset of goods and services as JPS ND.

We construct the implied amount of JPS ND aggregate spending directly induced by the tax rebates for Johnson et al.'s (2006) two main sets of estimates. The first considers only the contemporaneous effect during the current quarter. That MPC estimate is $\beta = 0.375$ (Johnson et al., 2006, Table 2, Panel C) and we apply it to the three-month moving average of aggregate rebate receipts:

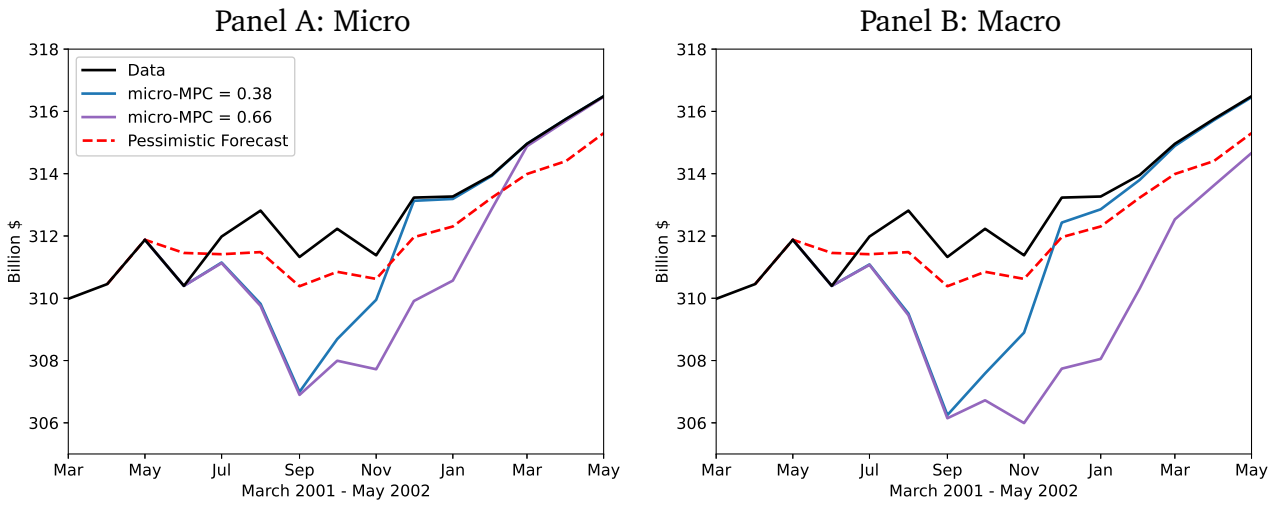
$$\textit{3-Month Induced Spending}_t = 0.375\left(\textit{rebate}_t + \textit{rebate}_{t-1} + \textit{rebate}_{t-2}\right)/3.$$

This equation incorporates the assumption that households will evenly spread out their induced rebate spending over three months.

The second set of estimates are from Johnson et al.'s (2006) Table 4, Panel A where they allow for lagged spending responses and find MPCs of 0.386 for the first quarter and an additional 0.273 for the subsequent quarter. In this case, we specify induced spending as:

$$\begin{aligned}\textit{6-Month Induced Spending}_t = {} & 0.386\left(\textit{rebate}_t + \textit{rebate}_{t-1} + \textit{rebate}_{t-2}\right)/3 \\ & +0.273\left(\textit{rebate}_{t-3} + \textit{rebate}_{t-4} + \textit{rebate}_{t-5}\right)/3.\end{aligned}$$

Figure 4. Micro and Macro Counterfactual Consumption Expenditures



Source: BEA and authors' calculations.
Panel A: Counterfactual expenditure paths are created by subtracting the nondurable consumption spending induced by the rebate from the path of actual nondurable consumption. Nondurable consumption expenditures are defined following Johnson et al. (2006). The 3-month estimates are taken from their Table 2, the 6-month estimates from their Table 4.
Panel B: Based on Two-Good, Two-Agent NK model simulations and actual data on rebates and nondurable consumption expenditures as measured by Johnson et al. (2006).

We then construct the micro counterfactual by subtracting the path of induced spending from the path of the actual aggregate JPS ND category spending.

Panel A of Figure 4 shows the path of aggregate BEA spending on the JPS ND categories as the solid black line and two micro counterfactuals in blue and purple. The first counterfactual is based on the contemporaneous estimates from Johnson et al.'s (2006) Table 2 and the second is based on the Table 4 estimates that allow lagged effects as well. The actual JPS ND spending category increased slightly during the fall of 2001, but the counterfactuals imply that spending in the JPS ND category would have fallen by around $5 billion from May (the month before the rebate passed) to its trough in September if there had been no rebate. The recovery is much slower when the Johnson et al. (2006) 6-month estimates are used. Both counterfactuals exhibit the same V-shaped we found for the 2008 rebate counterfactual in Orchard et al. (2023). The graph also shows a "pessimistic forecast" as a dashed red line. We will discuss the details of this forecast later in Section 5 when we assess plausibility.

4.2 Macro Counterfactual

The micro counterfactual does not allow for any general equilibrium forces, so next we construct a macro counterfactual by calibrating a standard TANK model to the US economy in 2001, simulating the economy's response to the tax rebates, and then subtracting that economy's simulated response from actual spending. Our model incorporates sticky wages and prices, variable utilization of capital, and investment adjustment costs. The model has two types of households, forward-looking optimizers and hand-to-mouth, that follow the consumption rules defined below. Our model is similar to Ramey (2021), which extends Galí et al.'s (2007) fiscal NK model. As most of the model is rather standard, we describe only the household side of the model here and leave the full details for the appendix.

4.2.1 Optimizing Households

A fraction $1-\gamma$ of households are forward-looking optimizers. They maximize utility

$$\sum_{t=0}^{\infty}\beta^t\left[\frac{\left[s_1^{\frac{1}{\iota}}(C_{1t}^o)^{\frac{\iota-1}{\iota}}+(1-s_1)^{\frac{1}{\iota}}(C_{2t}^o)^{\frac{\iota-1}{\iota}}\right]^{\frac{\iota}{\iota-1}(1-\frac{1}{\sigma})}}{1-\frac{1}{\sigma}}-\nu\frac{(H_t^o)^{1+\phi}}{1+\phi}\right]$$

subject to their budget constraint

$$A_t^o=\frac{R_{t-1}}{\Pi_t}A_{t-1}^o-C_{1t}^o-C_{2t}^o+W_tH_t^o-T_t^o+\text{Profits}_t$$

where C_{1t}^o is nondurable consumption as defined by Johnson et al. (2006), C_{2t}^o is all other consumption, and H_t^o is hours worked, A_t^o are holdings of the nominal bond, R_t is the gross nominal interest rate, Π_t is the gross inflation rate measured in the price of the consumption good, W_t is the real wage, T_t^o are net taxes (i.e. taxes less transfers), and *Profits* is real profit income. The parameter s_1 is the share of JPS nondurable expenditure in total consumption, ι is the elasticity of substitution across consumption goods, σ is the intertemporal elasticity of substitution for the consumption basket, and ϕ is the Frisch elasticity.

Optimizing households pick an optimal sequence for nondurable consumption, other consumption, and assets, $\{C_{1t}^o, C_{2t}^o, A_t^o\}_{t=0}^{\infty}$. Labor supply is not chosen by the household, but instead by a union.

The first order conditions for consumption and assets are:

$$\begin{aligned}
\lambda_t &= (C_t^o)^{-\frac{1}{\sigma}} \\
C_{1t}^o &= s_1 C_t^o \\
C_{2t}^o &= (1-s_1) C_t^o \\
\lambda_t &= \beta \frac{R_t}{\Pi_{t+1}} \lambda_{t+1}
\end{aligned}$$

where λ is the Lagrange multiplier on the household budget constraint and C_t^o is the sum of the two types of consumption.

4.2.2 Hand-to-Mouth Households

In order for lump-sum transfers to have general equilibrium effects, we require non-Ricardian households. We adopt Galí et al.'s (2007) assumption that a certain fraction γ consume hand-to-mouth. Relative to their set-up, our hand-to-mouth households may consume their income over several periods rather than all at once.

We assume that in steady state, hand-to-mouth households have the same after-tax income as optimizing households,

$$WH^m - T^m = WH^o - T^o + \text{Profits}$$

where variables superscripted by m denote the hand-to-mouth household.

We then directly specify dynamic marginal propensities to match lagged effects implied by the micro MPC estimates. Thus, expenditures on JPS nondurables are

$$C_{1t}^m - C_1^m = \sum_{l=0}^{L} mpc_l [W_{t-l} H_{t-l}^m - T_{t-l}^m - (WH^m - T^m)] \prod_{k=1}^{l} \frac{R_{t-k}}{\Pi_{t-k+1}}$$

where mpc_l is the marginal propensity to spend on nondurable consumption today out of income l periods ago. Income that was saved l periods ago for nondurable consumption today accrues real interest $\prod_{k=1}^{l} \frac{R_{t-k}}{\Pi_{t-1+1}}$.

We follow Johnson et al. (2006) and treat the residual expenditure category as not responding to the rebate. Thus, hand-to-mouth households simply consume a constant

amount of this good,

$$C_{2t}^m = C_2^m$$

The marginal utility of consumption for the hand-to-mouth household is

$$\lambda_t^m = (C_t^m)^{-\frac{1}{\sigma}}$$

where C_t^m is the sum of the two types of consumption.

4.2.3 The Remainder of the Model

The remainder of our model is rather standard so we briefly summarize its features here and provide a detailed description in the appendix. Product market firms are monopolistically competitive and face Calvo-style adjustment costs on output prices. They also face adjustment costs on capital, but have the option to vary the utilization of existing capital at the cost of higher depreciation. Labor is differentiated and unions set wages as a markup on reservation wages. Wages are also subject to Calvo-style adjustment costs. Given the wage set by the union, households stand ready to supply as many hours to the labor market as required by firms. The government sets the interest rate according to a Taylor rule, issues transfers and faces an intertemporal budget constraint. We assume that any temporary transfers are financed by deficits in the short run.

4.3 Results

In order to simulate the effects of the 2001 tax rebates, we calibrate the share of rule-of-thumb households to the micro MPC estimates of nondurable spending from Johnson et al. (2006) (0.375). With the model beginning in steady state, we shock the model with a path of government transfers that matches the actual path of aggregate tax rebate payments in 2001. We then subtract the simulated model deviations from the actual path of nondurables expenditure (defined according to JPS) to create the counterfactual path of spending had there been no rebate.

Panel B of Figure 4 shows the path of actual nondurable spending from the BEA using the JPS nondurable definition and the macro counterfactual path with GE effects. The macro counterfactuals show a larger drop and a more pronounced V-shape in JPS

nondurable expenditure than the micro counterfactual. That GE forces amplify the micro effects is typical of HANK-style models (Auclert et al., 2018).

Following Orchard et al. (2023), we define the *GE-MPC* to be the general equilibrium response of consumption to the rebate. The GE-MPC, which is the micro MPC plus any general equilibrium amplification effects, is approximately equal to the GDP multiplier if the economy is closed and the stimulus is temporary. Table 1 shows the GE-MPC for different values of the micro MPC. The first row, shows the results with a micro MPC of 0.38. The GE-MPC is 0.5. The second row, shows that a micro MPC of 0.66 over 6-months yields a GE-MPC of 1.19. Thus, the GE-MPC is higher than the micro MPC, and the GE-MPC rises non-linearly with the level of the micro MPC. The amplification and non-linearity is due to the standard Keynesian multiplier effect.

Table 1. General Equilibrium Marginal Propensity to Consume

MPCs	
micro	GE
0.38	0.50
0.66	1.19

5 Is the Macro Counterfactual Plausible?

We now assess whether the macro counterfactual is plausible. This step of the methodology is particularly challenging for the 2001 rebates because the terrorist attacks of September 11th occurred during the third month of the distribution of the rebates. The macro counterfactual shown in Figure 4 implies that JPS's definition of nondurable spending would have declined by 1.4 percent in the three months ending in September 2001. That is a very large drop by historical standards, at least since the monthly data began in 1959, and is exceeded only by COVID-19 and is almost the same magnitude as the decline after the failure of Lehman Brothers in 2008. The question is whether 9/11 could have caused that fall. To assess that possibility, we review the commentary and forecasts surrounding 9/11.

Panel A of Figure **??** shows that, prior to the passage of the tax change and rebates in June 2001, professional forecasters in the Survey of Professional Forecasters (SPF) were

expecting small positive consumption growth in 2001Q3.[3] The May 9th Greenbook noted the decline in consumer sentiment and it predicted anemic growth in consumer spending in the second quarter with a pickup in the third quarter (Federal Reserve Board of Governors (2001a). Since these forecasts are for total consumer expenditures, we also combine the separate services and nondurable goods growth forecasts from the FED Greenbook to create a more appropriate comparison with the JPS nondurables category. Panel B of Figure 5 shows that this forecast lies above actual JPS nondurables expenditure in 2001.[4]

There were further signs of slowing in late summer. A September 7th article in *The New York Times* (Section C, page 8) reported disappointing sales for retailers in August. While overall retail sales had risen, there seemed to be a shift in demand to discounters, resulting in a decline in same-store sales at numerous types of outlets such as department stores. Several commenters noted that there was no evidence of a boost from the rebates, with one commenter calling the rebates a "nonevent."

The University of Michigan index of consumer sentiment had fallen steadily from November 2000 to April 2001 by a cumulative 18%. The index bounced up by 4% in May and stayed at that level through August, but then fell 11% from August to September. According to the Federal Reserve's September 27, 2001 Greenbook (Federal Reserve Board of Governors (2001b), the preliminary consumer sentiment data through September 10th indicated that sentiment had already fallen 8% from August. Thus, two-thirds of the September fall in consumer sentiment happened before the terrorist attacks. The Greenbook also noted that The Conference Board measure of consumer sentiment fell more than 16 points from August to September, but that 88 percent of the responses were received before September 11th. Thus, most of the fall seemed to predate the attacks.

The late September Greenbook estimated that the disruptions caused by 9/11 would likely subtract a quarter of a percentage point (not annualized) from the growth in personal consumption from the second to third quarter. Since September is the last month of the third quarter, this estimate implies that total personal consumption fell $2.3 billion in September as a result of 9/11. However, the subsequent Greenbook noted that

3. Since the responses for the SPF are due by the middle of the quarter, all results would have been collected prior to the tax law passage in June 2001.

4. The May 9, 2001 Greenbook did not know of the passage of the Tax Rebates, but the FED staff assumed that rebates would be included in the Tax bill to be sent to households in the fourth quarter of 2001.

the surge in motor vehicle sales in October in response to the dealer incentives suggested that consumer sentiment might have bounced back more quickly than anticipated.

Nevertheless, it is clear that even apart from the terrorist attacks, economic events were reducing consumer confidence. As a supplement to the professional forecasts, we create our own forecasts based on a time series model augmented with negative information that came out after the tax legislation was enacted. Our objective is to create a more pessimistic forecast than the ones based only on information available through May.

Our forecasts are based on a parsimonious time series model with the following variables: log real consumption using JPS's definition, log real disposable income, log total consumption deflator, and consumer sentiment from the Michigan Survey.[5] The current value of consumer sentiment enters the equations for consumption, income and price. We include six lags of all variables and estimate the model on monthly data from 1978 (when the monthly consumer sentiment index becomes available) through 2019.

We construct two versions of the forecasts. The *regular* forecast estimates the time series model just described and creates forecasts using information available only through May 2001. The *pessimistic* forecast augments the model with the current value and six lags of a dummy variable for recession and current value and one lag of a dummy variable for 9/11 and treats them as exogenous in the construction of forecasts. Thus, we are implicitly assuming the forecaster knows the economy is in recession from April through November 2001 and that 9/11 occurs.

Panel B of Figure 5 shows the resulting forecasts. The regular forecast without the recession and 9/11 substantially overshoots the actual data. The pessimistic forecast that includes the two sets of dummy variables tracks just below the actual data. Neither forecast shows a dramatic V-shaped decline in consumer expenditures.

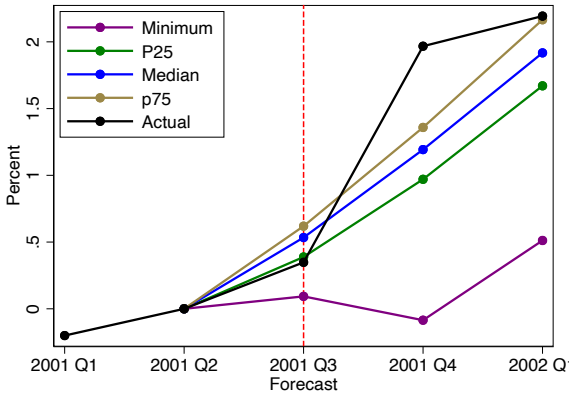
Figure 4 in the last section showed this pessimistic forecast along our micro and macro counterfactuals. The counterfactuals lie substantially below this forecast. This casts doubt that these counterfactuals, and thus the large aggregate effects of the rebate, are plausible.

Could it be that our forecasts understate the effects of 9/11? The top panel of Table 2 shows that the cumulative drop in expenditure through December 2001 implied by our

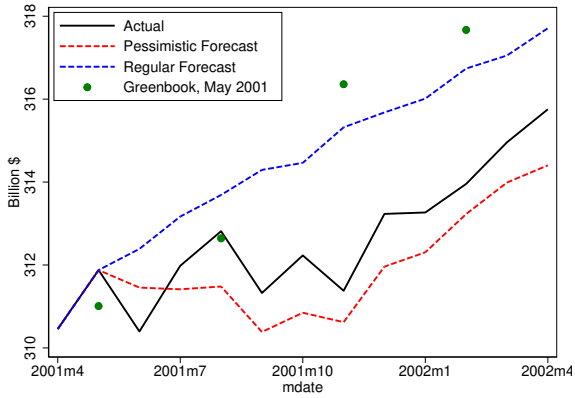
5. We initially included other variables such as gasoline prices and the Gilchrist and Zakrajšek (2012) bond spread but they did not affect the forecasts since movements in those variables weren't important during the 2001 episode.

Figure 5. Real Nondurable Consumption Forecasts

Panel A: Real Consumption Growth



Panel B: Real JPS Nondurable Consumption



Panel A: SPF moments are period-by-period.
Panel B: The pessimistic model includes a recession dummy, a 9/11 dummy, and lags as exogenous variables in all equations. All other variables based on information through May 2001. Nondurable consumption expenditures are defined following Johnson et al. (2006).

pessimistic forecasts is comparable to the the late September Greenbook's assessment of the effects of 9/11 on consumption. We also compare our estimates to the implied effect of the drop in consumer confidence in September 2001 using the estimates in Barsky and Sims (2012). These are very similar to our estimates, both in terms of the implied decline in September as well as the cumulative decline through December 2001.

The bottom panel of Table 2 show that our macro counterfactuals imply an effect of the 2001 rebate on JPS nondurable consumption through December 2001 that was an order of magnitude larger. This suggests it is unlikely that the effects of the rebate were offset by 9/11. Even for September, the counterfactual implies that the rebate had at least twice the effect of 9/11. This contrasts strongly with contemporary views that the rebate was a "nonevent."

Since the counterfactuals do not appear plausible, this leaves two possibilities: either the micro MPC estimates are too large or the macro model needs to substantially dampen the aggregate effects of the rebate. We investigate the first possibility next and find evidence that, when appropriately measured, the micro MPC estimates are substantially smaller than those reported by Johnson et al. (2006).

Table 2. Estimated 9/11 Impact on JPS Nondurables vs Our Counterfactuals

Forecast of 9/11 Effect	September 2001 Impact	Through December 2001
Our Pessimistic Forecast	\$0.7bn	\$2.4bn
September Greenbook	\$2.3bn	\$2.3bn
Barsky-Sims (2012)	\$0.9bn	\$2.8bn
Macro Counterfactual:		
micro MPC = 0.375	\$5.1bn	\$17.2bn
micro MPC = 0.66	\$5.2bn	\$25.8bn

6 Micro MPC Estimates

Here we reexamine the frequently cited estimates from Johnson et al. (2006) that imply large MPCs on nondurable expenditures of 0.4 to 0.66. Similar to the re-analyses of the 2008 rebates by Orchard et al. (2023), Borusyak and Jaravel (2017), and Borusyak et al. (2022), we find that the MPC estimates for the 2001 tax rebates are upward biased. However, unlike Orchard et al. (2023) we find that the main reason for the bias is not econometric issues, but rather how nondurable expenditures are defined. We replicate the Johnson et al. (2006) results using their definition of nondurable consumption and estimate similar MPCs of 0.32 to 0.64. However, if we instead define nondurables using the BEA's definitions, we estimate that the nondurable MPC is one-third to one-half the size and statistically insignificant. Moreover, the MPC on total consumption is estimated to be close to zero but estimated very imprecisely.

6.1 Baseline Johnson et. al. (2006) Results

Johnson et al. (2006) use both the variation in treatment time (i.e., the month in which the household received the rebate) and the treatment size (i.e. the dollar value of the rebate check) to estimate the causal impact of receiving a rebate on household spending using a standard difference-in-differences (DID) event-study methodology. For brevity, we show only our replication results for the 2SLS version of Johnson et al. (2006). Our replication follows Johnson et al. (2006) exactly, except that we exclude households that receive more than one rebate check since these households are not valid comparisons to treated households (Orchard et al., 2023).

For the contemporaneous specification, Johnson et al. (2006) estimate the following regression,

$$C_{i,t} - C_{i,t-1} = \sum_s \beta_{0s} month_s + \beta_1' \mathbf{X}_{i,t-1} + \beta_2 \widehat{ESP_{i,t}} + u_{i,t} \tag{1}$$

where t indexes the interview (performed once every three months), and i indexes individual households. The regression includes fixed effects for each month ($month_s$), household controls for age and change in household size $X_{i,t}$, and the main variable of interest, $\widehat{ESP_{i,t}}$, is the predicted size of the rebate or Economic Stimulus Payment (ESP) which is instrumented by an indicator for rebate receipt.

Table 3 column 1 shows the result of our replication. We estimate a 0.32 MPC for Johnson et al.'s (2006) definition of nondurables spending. As compared to Johnson et al. (2006) our replication has a slightly lower estimated MPCs and sample sizes, partially due to our exclusion of households that receive more than one rebate.

6.2 Nonrobustness to Different Categorizations of Spending

We now demonstrate that the results of Johnson et al. (2006) depend on a particular definition of nondurables spending that omits many services but includes some goods classified by the BEA as durables. We then show that the estimated MPCs are not robust to alternate definitions of nondurables spending.

Johnson et al. (2006) follow Lusardi (1996) by defining nondurable expenditure as the sum of several aggregate categories in the consumer expenditure survey.[6] We construct an alternate definition of nondurable goods and services based on the BLS staff's concordance between the Consumer Expenditure Survey categories and the BEA's aggregate personal consumption spending categories (Bureau of Labor Statistics, 2019). Hereafter, we refer to the Johnson et al. (2006) definition of nondurables as the JPS definition and the one based on the BLS concordance as the BEA definition.

The JPS and BEA definitions of nondurables differ substantially. For example, services such as housing services, vehicle repair and leases, and foreign travel are included in the BEA definition, but not in the JPS definition. Jewelry and watches are classfied

6. Specifically nondurable spending is defined as spending on several Consumer Expenditure Survey aggregate categories including food, alcoholic beverages, tobacco, utilities, personal care, household operations, public transportation, gas and motor oil, apparel, health care, and miscellaneous expenses.

Table 3. Household Spending Response to Rebate by Subcategory: 2SLS

Panel A: Full Sample

	JPS Definitions	BEA Definitions			
	Nondurables	ND Goods	Services	Durables	Total
	(1)	(2)	(3)	(4)	(5)
Rebate Amount	0.32**	0.06	0.03	−0.16	−0.07
	(0.15)	(0.06)	(0.22)	(0.50)	(0.58)
Observations	12,018	12,018	12,018	12,018	12,018

Panel B: Rebate Only Sample

	JPS Definitions	BEA Definitions			
	Nondurables	ND Goods	Services	Durables	Total
	(1)	(2)	(3)	(4)	(5)
Rebate Amount	0.33	−0.06	−0.21	0.37	0.10
	(0.32)	(0.10)	(0.40)	(0.81)	(0.97)
Observations	5,875	5,875	5,875	5,875	5,875

Notes: Standard errors, in parentheses, are clustered at the household level. Significance is indicated by: $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$. All regressions include interview (time) fixed effects, as well as household level controls for age, change in number of adults, and change in number of children. The 2SLS results use $I(Rebate > 0)$ and the other regressors as instruments for the rebate amount.

as nondurables by JPS, but as durables by the BEA.[7] Additionally, some loan payments are included in the JPS definition of nondurables, but are mostly excluded from BEA consumer spending.

We reestimate Equation (1) using the BEA definitions of nondurable goods and services. Table 3, Panel A Columns 3 and 4, show the results of this estimation. The point estimates of the MPC for BEA nondurables spending are small and insignificant: 0.06 for nondurable goods and 0.03 for services. Summing together, the nondurable MPC using the BEA definitions is only 0.09.[8] The durable goods MPC is negative but imprecisely estimated, and so is the total MPC based on BEA definitions.

7. In the Consumer Expenditure Survey, apparel includes jewelry and watches and health care includes some medical devices, but these are classified as durable goods by the BEA.

8. When we instead sum the categories and then estimate the MPC, the value is also 0.09 (with a standard error of 0.24).

In Panel B, we restrict the estimation to the sample of households receiving rebates, so the only variation used is the randomized timing. In this subsample our MPC estimates for nondurable goods and services are negative but also substantially noisier than in Panel A. The MPC for total consumption is positive but close to zero. In what follows we focus on the full sample because it yields relatively more precise estimates.

To understand better why estimated nondurable MPCs using the BEA and JPS definitions are so different, we estimate the MPCs separately for each of the 135 categories in the BLS' Consumer Expenditure Survey-BEA concordance. We then create a mapping between these categories and nondurable spending in the JPS definition. Appendix Table B.1 provides a full accounting of estimated MPCs for each sub-category and Table 4 provides a summary.

On average, expenditure categories included in the JPS definition of nondurables have a higher MPC than the excluded categories. Table 4 Column 1 shows that among all BEA spending categories, the average MPC is positive for categories included in JPS nondurables and negative for other categories. Column 2 shows that summed together, the nondurables JPS MPCs mapped into BEA categories is 0.26, [9] while the excluded categories' MPCs sum to -0.37, resulting in a total PCE MPC that is actually negative.

The remaining columns of Table 4 show that even within the subcategories of nondurable goods, services, and durables, the categories included in JPS nondurables tend to have higher MPCs than the excluded categories. Among nondurable goods, the excluded nondurable good categories do have a similar MPC to the included categories; however, nondurable goods account for only one-seventh of the JPS nondurable MPC. Within services, the JPS-included categories account for 0.17 of the 0.26 JPS nondurable MPC. In contrast, if all BEA service categories are included, the services MPC falls to 0.03. Finally, as we discussed previously, some BEA durables are included in the JPS definition of nondurables. These included durables, which include items such as jewelry and medical devices, have a higher contribution to the JPS nondurable MPC than the included BEA nondurable goods.

As Appendix Table B.1 shows, the micro estimates for the BEA subcategories are imprecise, vary substantially, and are sometimes negative. Of the 135 BEA subcategories that we consider, only 15 have estimated micro MPCs that are both positive and

9. The sum of the JPS nondurable MPCs mapped into BEA categories does not match the total MPC shown in Column 2 of Table 3 since some JPS categories are not included in total BEA consumption and because the BLS' correspondence attributes only a fraction of Consumer Expenditure Survey categories such as home improvement to consumption (the rest is attributed to investment).

Table 4. Sum of Disaggregated MPCs

BEA Category	All	ND Goods	Services	Durables
In JPS Definition	0.261	0.038	0.165	0.058
Not in JPS Definition	−.328	0.022	−.135	−.215
All	−.067	0.06	0.031	−.157
Observations	135	26	71	38

Notes: Average MPC is the average coefficient from 2SLS regressions of BEA subcategories on rebate amount using the indicator for rebate receipt as an instrument. Sum MPC is the sum of all rebate amount coefficients in the specified category.

statistically significant; 13 of these 15 categories are included in the JPS definition of nondurables.

Johnson et al. (2006) also estimate a specification that allows the rebate to have effects that last for a full six months. Table 5 reports MPCs for specifications of (1) that include a lagged rebate variable. Using JPS definitions for the full sample (Panel A, Column 1) we estimate a cumulative 6-month MPC of 0.64, similar to Johnson et al. (2006).[10] Using BEA definitions of nondurable goods and services reduces the estimated 6-month MPC by half and eliminates the statistically significance. Adding the response of durable expenditures yields an even lower 6-month MPC for total expenditures of 0.11. The MPC estimates for the rebate-only sample in Panel B fall from 0.8 using JPS definitions (Column 1) to substantially negative -3.8 for total PCE.

In short, we find that the MPCs reported by Johnson et al. (2006) are not representative of MPCs on personal consumption expenditure. Using either standard BEA categories or total consumer expenditures we estimate micro MPCs that are small, at or below 0.1, and not statistically significant. These results suggest that the overall stimulative effect of the 2001 rebate was likely very modest. To the extent that the MPCs for JPS categories are large this likely reflects noise at the micro level.

7 Lessons from the 2001 Rebate Natural Experiment

We reexamined the micro estimates based on the natural experiment of the 2001 rebates because the implied macro counterfactual path of consumer spending did not seem entirely plausible. Our results from the last section offer a full resolution of the

10. Because the consumption equation is estimated in first-differences, the 6-month cumulative MPC is equal to twice the coefficient on the current rebate plus the coefficient on the lagged rebate.

Table 5. Household Spending Response to Rebate by Subcategory: 2SLS with Lag

Panel A: Full Sample

	JPS Definitions	BEA Definitions			
	Nondurables	ND Goods	Services	Durables	Total
	(1)	(2)	(3)	(4)	(5)
Rebate Amount	0.32**	0.07	0.04	−0.15	−0.05
	(0.14)	(0.06)	(0.22)	(0.49)	(0.57)
Lag Rebate Amount	0.00	0.07	0.05	0.08	0.20
	(0.12)	(0.06)	(0.19)	(0.40)	(0.45)
6-month MPC	0.64	0.20	0.13	-0.22	0.11
6-month MPC S.E.	(0.29)	(0.12)	(0.46)	(0.98)	(1.15)
Observations	12,018	12,018	12,018	12,018	12,018

Panel B: Rebate Only Sample

	JPS Definitions	BEA Definitions			
	Nondurables	ND Goods	Services	Durables	Total
	(1)	(2)	(3)	(4)	(5)
Rebate Amount	0.37	−0.02	−0.47	−0.47	−0.97
	(0.36)	(0.11)	(0.46)	(0.81)	(0.98)
Lag Rebate Amount	0.07	0.08	−0.46	−1.48*	−1.87*
	(0.28)	(0.11)	(0.42)	(0.88)	(1.03)
6-month MPC	0.80	0.04	-1.41	-2.42	-3.80
6-month MPC S.E.	(0.88)	(0.29)	(1.18)	(2.09)	(2.47)
Observations	5,875	5,875	5,875	5,875	5,875

Notes: Standard errors, in parentheses, are clustered at the household level. Significance is indicated by: $^{*}\,p < 0.1$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$. All regressions include interview (time) fixed effects, as well as household level controls for age, change in number of adults, and change in number of children. The 2SLS results use $I(Rebate > 0)$ and the other regressors as instruments for the rebate amount.

implausible counterfactual. Specifically, we explain the implausible decline of the counterfactual path to be the result of micro estimates that are based on very noisy data. Johnson et al.'s (2006) 6-month cumulative MPC is estimated to be 0.66 based on their categories. However, we have shown that the results are not robust to alternative categorizations, such as those by the BEA. In fact, our estimated 6-month cumulative MPC based on total consumer spending is approximately zero. If we use this estimate to

construct the macro counterfactual, the path is approximately equivalent to the path of the actual data, suggesting no discernible stimulative effect of the rebates. Thus, we our revisions of the micro estimates have resolved our implausible counterfactual.

It is important to highlight another confounding factor in the 2001 rebate natural experiment that makes it unsuitable for drawing inferences about the effects of temporary stimulus. The 2001 tax rebates were considered to be a down payment on the semi-permanent tax cut that would go into effect in January 2002 (Shapiro and Slemrod, 2003b). As such, the rebate cannot be interpreted as a one-time change in disposable income. How does this affect the interpretation of our estimates?

In the full sample regressions, we exploit the fact that some households receive the rebate and others do not. But that same variation is also correlated with exposure to the permanent tax cut. Thus, the full sample MPC estimates likely reflect some anticipation effect of the permanent tax cut. The magnitude of this effect depends on whether households expect the tax cut to be permanent (i.e., funded by a future reduction in government spending) or offset by higher future taxes. The rebate-only sample exploits only the variation in the timing of rebate receipt and thus not on overall eligibility. In this sense, these estimates more cleanly isolate non-Ricardian behavior, such as liquidity constraints.

Since we estimate only small MPCs that are similar across the full and rebate-only samples, our results are consistent with households being largely Ricardian. This conclusion is also consistent with the behavior of aggregate consumer expenditure.

8 Other Applications of the Methodology

In this section, we briefly summarize two previous applications of our methodology to provide further illustration and comparison. The examples demonstrate that each case study usually requires different methods and yields different answers for the reconciliation of the micro estimates with the aggregate data.

8.1 The 2008 Rebate MPCs

In Orchard et al. (2023), we constructed macro counterfactuals using the leading estimates of MPCs from the 2008 rebate and found counterfactual declines in consumer spending that were even more extreme than the ones for the 2001 rebate. However,

the details behind the 2001 and 2008 case studies differed substantially. The 2001 tax rebate accompanied a permanent tax change and did not have an income cap, while the 2008 tax rebate was a temporary transfer that was targeted to low- and medium-income earners. The estimated MPC for motor vehicles was negative for the 2001 rebates, while spending on motor vehicles was a large part of the estimated MPCs for the 2008 rebates. Finally, micro MPC estimates from the 2001 rebate were robust to alternate DID econometric methods, but fell apart when other consumption aggregations were considered, while the 2008 micro MPC estimates fell if improved DID methods were employed.

Specifically, we showed that the micro MPC estimates for the 2008 rebate based on a standard DID methodology were upward biased for three reasons: (1) omission of a lagged rebate, (2) heterogenous treatment effects leading to forbidden comparisons, and (3) a negative correlation between rebate receipt and lagged spending. After correcting for these biases, the micro MPC for total PCE fell from a range of 0.5 to 0.9 down to 0.3, with almost all of the expenditure allocated to motor vehicles. These results were consistent with earlier findings of Borusyak and Jaravel (2017) and Borusyak et al. (2022), who found substantially reduced MPCs when they applied their new DID method to another set of 2008 rebate estimates. Because the 2001 rebate was distributed over a slightly shorter time period and estimates vary much less across months of disbursement, the 2001 estimates are not sensitive to the econometric issues we highlighted for the 2008 rebate.

We also found that because most of the induced spending was on motor vehicles, the stimulative effects of the tax rebates were dampened rather than amplified in general equilibrium. During the roll-out of the 2008 tax rebates, the relative price of new motor vehicle prices spiked relative to trend. We captured this effect in our two-good TANK model by including a durable good (standing in for motor vehicles) following McKay and Wieland (2021, 2022). With standard values for the relative supply elasticity of durable goods, the model was able to replicate the increase in relative motor vehicle purchases seen in the data and predicted that the consumption response following the tax rebates is substantially dampened in equilibrium. The crowding out was large because durable demand is inherently more elastic than nondurable demand. In contrast, with the 2001 tax rebates we found no statistically significant increase in durable goods spending, so similar macro dampening did not occur in 2001.

8.2 ARRA Multipliers

A second application of our method was to cross-state multipliers estimated during the Great Recession. Numerous papers exploited the variation of federal government outlays across states resulting from the American Recovery and Reinvestment Act (ARRA) passed in February 2009.[11] Various categories of the spending were based on pre-existing formulas, so researchers were able to construct Bartik-style (Bartik (1991) instruments by state to avoid endogoneity problems.

Chodorow-Reich (2019) synthesized the estimates by standardizing the econometric framework and found that all the main sets of instruments produced estimates of approximately two job-years created per $100,000 spent. Using an approximation for converting the job creation numbers to an output multiplier, he found a state-level output multiplier of approximately two. Building on earlier theoretical work by Farhi and Werning (2016) and Nakamura and Steinsson (2014) on how to translate the cross-state employment and output multipliers to national multipliers, Chodorow-Reich (2019) concluded that the estimated state-level multipliers were likely a *lower bound* for the national multipliers. Thus, the estimates implied that the ARRA had quantitatively large aggregate effects on employment and output.

Ramey (2019a) conducted a plausibility test of these estimates. In particular, she used Chodorow-Reich's (2019) estimated impulse responses for employment to calculate a counterfactual unemployment rate had there had been no ARRA. According to the counterfactual, the national unemployment rate would have risen to 15.5% by the end of 2009 and would have remained elevated for at least a year more. This counterfactual suggests a rise in the unemployment rate on par with the rise during the first two years of the Great Depression, which is surprising in light of the very different response of the Federal Reserve during the two historical episodes.

In subsequent work Ramey (2019b) revisited the leading cross-state ARRA estimates and questioned the use of those estimates for assessing national effects. In particular, she suggested that since the estimates weighted each state's per capita outcomes equally and did not account for the crowd-in effect on state and local spending, the estimates might not be nationally representative. When she reestimated the cross-state model using population weights and including state and local spending, she found that the estimate of the job-years created per $100K spent fell from 2 to 0.9.

11. Examples include Chodorow-Reich et al. (2012), Wilson (2012), and Conley and Dupor (2013). See Chodorow-Reich (2019) for a review and synthesis of this literature.

The counterfactual implied by the revised estimates might be considered to be more plausible. Using the estimate of 0.9 rather than 2 produces a counterfactual suggesting that the unemployment rate would have risen to 12.4% rather than 15.5% had there been no ARRA spending. Further work at reconciliation might re-examine the theory that predicts that the cross-state estimates are a lower bound on the national multipliers.

9 Conclusion

In this paper we have argued that researchers should take the "credibility revolution" one step further and assess the *plausibility* of micro estimates and macro models. We suggest that researchers calibrate macroeconomic models using micro estimates to create macro counterfactuals of historical episodes. An assessment of the plausibility of those counterfactuals can help determine which existing estimates are more reasonable and can also guide researchers to the micro estimates and macro models that deserve more scrutiny.

We have illustrated our method by assessing the plausibility of the nondurable MPC estimates of the 2001 tax rebates from Johnson et al. (2006). We made two main contributions: (1) we showed that a standard macroeconomic model calibrated to Johnson et al. (2006) leads to a macroeconomic counterfactual implying that consumption would have fallen far more than what would have been expected given macroeconomic conditions at that time; and (2) we reconciled the micro estimates with the aggregate effects after revisiting the micro estimates. We found that the estimates from Johnson et al. (2006) were not robust to the definition of consumption categories and moreover did not apply to total consumption. We concluded that the micro data were simply too noisy to give a clear answer. We also echoed Shapiro and Slemrod's (2003b) argument that the quasi-permanent nature of the tax cuts make estimates from 2001 rebate natural experiment unsuitable for determining the effects of temporary transfers. Finally, we also reviewed two previous applications, one to the 2008 rebate and the other to the cross-state ARRA multipliers. For the 2008 rebate, we highlighted how revisions to both the micro estimates and the macro model were necessary for the reconciliation of the micro estimates with the aggregate data.

We conclude with a few comments on the assessment of plausibility. While all steps in our methodology require some degree of subjective judgment, such as the choice of particular micro estimates from a paper and the multitude of choices involved in

constructing and calibrating a macro model, the step that is the most subjective is the assessment of plausibility. While plausibility, like credibility, is in the eyes of the beholder, a researcher can nonetheless put that assessment on a firmer foundation by comparing the counterfactual to other historical episodes and to forecasts at the time. With those comparisons laid out, researchers and policymakers trying to decide which of a range of micro estimates to use will have a new tool for narrowing the range.

References

Angrist, Joshua D and Jörn-Steffen Pischke, 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *Journal of economic perspectives* 24(2): 3–30.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub, 2018. "The intertemporal keynesian cross." Technical report, National Bureau of Economic Research.

Barsky, Robert B and Eric R Sims, 2012. "Information, animal spirits, and the meaning of innovations in consumer confidence." *American Economic Review* 102(4): 1343–1377.

Bartik, Timothy J, 1991. "Who benefits from state and local economic development policies?" .

Borusyak, Kirill and Xavier Jaravel, 2017. "Revisiting event study designs." *Available at SSRN 2826228* .

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, 2022. "Revisiting event study designs: Robust and efficient estimation." *arXiv preprint arXiv:2108.12419* .

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

Bureau of Labor Statistics, Staff, 2019. "Comparing expenditures from the Consumer Expenditure Surveys with the Personal Consumer Expenditures: Results of the CE/PCE Concordance." Technical report, Bureau of Labor Statistics.

Chodorow-Reich, Gabriel, 2019. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" *American Economic Journal: Economic Policy* 11(2): 1–34.

Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston, 2012. "Does state fiscal relief during recessions increase employment? Evidence from the American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy* 4(3): 118–145.

Congressional Budget Office, Staff, 2009. "Did the 2008 Tax Rebates Stimulate Short-Term Growth?" Technical report, Economic and Budget Issue Brief.

Conley, Timothy G and Bill Dupor, 2013. "The American Recovery and Reinvestment Act: solely a government jobs program?" *Journal of monetary Economics* 60(5): 535–549.

Council of Economic Advisers, Staff, 2010. "The Economic Impact of the American Recovery and Reinvestment Act." Technical report, Executive Office of the President.

Farhi, Emmanuel and Iván Werning, 2016. "Fiscal Multipliers: Liquidity Traps and Currency Unions." In *Handbook of Macroeconomics. Vol. 2*, edited by John B. Taylor and Harald Uhlig., pp. 2417–2492. Amsterdam: Elsevier.

Federal Reserve Board of Governors, Staff, 2001a. "Current Economic and Financial Conditions: Summary and Outlook." Technical report, May 9, 2001 Greenbook.

Federal Reserve Board of Governors, Staff, 2001b. "Current Economic and Financial Conditions: Summary and Outlook." Technical report, September 27, 2001 Greenbook.

Feldstein, Martin, 2008. "The Tax Rebate Was a Flop; Obama's Stimulus Plan Won't Work Either." *The Wall Street Journal,* August 8.

Galí, Jordi, David López-Salido, and Javier Vallés, 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association* 5(1): 227–270.

Gilchrist, Simon and Egon Zakrajšek, 2012. "Credit spreads and business cycle fluctuations." *American economic review* 102(4): 1692–1720.

Huntley, Jonathan and Valentina Michelangeli, 2014. "Can tax rebates stimulate consumption spending in a life-cycle model?" *American Economic Journal: Macroeconomics* 6(1): 162–189.

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.

Lusardi, Annamaria, 1996. "Permanent income, current income, and consumption: Evidence from two panel data sets." *Journal of Business & Economic Statistics* 14(1): 81–90.

McKay, Alisdair and Johannes F Wieland, 2021. "Lumpy durable consumption demand and the limited ammunition of monetary policy." *Econometrica* 89(6): 2717–2749.

McKay, Alisdair and Johannes F Wieland, 2022. "Forward guidance and durable goods demand." *American Economic Review: Insights* 4(1): 106–22.

Nakamura, Emi and Jón Steinsson, 2014. "Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions." *American Economic Review* 104(3): 753–792.

Orchard, Jacob, Valerie A Ramey, and Johannes F Wieland, 2023. "Micro MPCs and macro counterfactuals: the case of the 2008 rebates." Technical report, Working Paper.

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): 2530–53.

Ramey, Valerie A, 2011. "Can government purchases stimulate the economy?" *Journal of Economic Literature* 49(3): 673–685.

Ramey, Valerie A, 2019a. "Fiscal Policy: Tax and Spending Multipliers." *Evolution or Revolution?: Rethinking Macroeconomic Policy after the Great Recession* p. 121.

Ramey, Valerie A, 2019b. "Ten Years after the Financial Crisis: What Have We Learned from the Renaissance in Fiscal Research?" *Journal of Economic Perspectives* 33(2): 89–114.

Ramey, Valerie A, 2021. "The Macroeconomic Consequences of Infrastructure Investment." In *Economic Analysis and Infrastructure Investment*, edited by Edward L. Glaeser and James M. Poterba., p. Chapter 4. Chicago: University of Chicago Press.

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): 216–50.

Shapiro, Matthew D and Joel Slemrod, 2003a. "Consumer response to tax rebates." *American Economic Review* 93(1): 381–396.

Shapiro, Matthew D and Joel Slemrod, 2003b. "Did the 2001 tax rebate stimulate spending? Evidence from taxpayer surveys." *Tax policy and the economy* 17: 83–109.

Taylor, John B, 2009. "The lack of an empirical rationale for a revival of discretionary fiscal policy." *American Economic Review* 99(2): 550–55.

Wilson, Daniel J, 2012. "Fiscal spending jobs multipliers: Evidence from the 2009 American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy* 4(3): 251–282.

Appendix

A Model

A.1 Optimizing Households

A fraction $1-\gamma$ of households maximizes utility subject to their budget constraints. The utility function of each household is

$$\sum_{t=0}^{\infty}\beta^t\left[\frac{\left[s_1^{\frac{1}{\iota}}(C_{1t}^o)^{\frac{\iota-1}{\iota}}+(1-s_1)^{\frac{1}{\iota}}(C_{2t}^o)^{\frac{\iota-1}{\iota}}\right]^{\frac{\iota}{\iota-1}(1-\frac{1}{\sigma})}}{1-\frac{1}{\sigma}}-\nu\frac{(H_t^o)^{1+\phi}}{1+\phi}\right]$$

where C_{1t}^o is nondurable consumption as defined by Johnson et al. (2006), C_{2t}^o is all other consumption, and H_t^o is hours worked. The parameter s_1 is the share of JPS nondurable expenditure in total consumption, ι is the elasticity of substitution across consumption goods, and ϕ is the Frisch elasticity.

The optimizing household budget constraint is

$$A_t^o=\frac{R_{t-1}}{\Pi_t}A_{t-1}^o-C_{1t}^o-C_{2t}^o+W_tH_t^o-T_t^o+\text{Profits}_t^k+\text{Profits}_t^s$$

where R_t is the gross nominal interest rate, Π_t is the gross inflation rate measured in th price of the consumption good, A_t^o are holdings of the nominal bond, W_t is the real wage, T_t^o are net taxes (i.e. taxes less transfers), Profits^k are profits of the capital good producing firms, and Profits^s are profits of the sticky-price firms

Optimizing households pick an optimal sequence for nondurable consumption, other consumption, and assets, $\{C_{1t}^o,C_{2t}^o,A_t^o\}_{t=0}^{\infty}$. Labor supply is not chosen by the household, but instead by a union. The first order conditions for consumption and assets are:

$$\begin{aligned}\lambda_t&=(C_t^o)^{-\frac{1}{\sigma}}\\C_{1t}^o&=s_1C_t^o\\C_{2t}^o&=(1-s_1)C_t^o)\\\lambda_t&=\beta\frac{R_t}{\Pi_{t+1}}\lambda_{t+1}\end{aligned}$$

where λ is the Lagrange multiplier on the household budget constraint.

A.2 Hand-to-Mouth Households

In order for lump-sum transfers to have general equilibrium effects, we require non-Ricardian households. We adopt Galí et al.'s (2007) assumption that a certain fraction γ consume hand-to-mouth. Relative to their set-up, our hand-to-mouth households may consume their income over several periods rather than all at once.

We assume that in steady state, hand-to-mouth households have the same after-tax income as optimizing households,

$$WH^m - T^m = WH^o - T^o + \text{Profits}^k + \text{Profits}^s$$

where variables superscripted by m denote the hand-to-mouth household.

We then directly specify dynamic marginal propensities to match lagged effects implied by the micro MPC estimates. Thus, expenditures on JPS nondurables are

$$C_{1t}^m - C_1^m = \sum_{l=0}^{L} mpc_l [W_{t-l} H_{t-l}^m - T_{t-l}^m - (WH^m - T^m)] \prod_{k=1}^{l} \frac{R_{t-k}}{\Pi_{t-k+1}}$$

where mpc_l is the marginal propensity to spend on consumption goods today out of income l periods ago. Income that was saved l periods ago for consumption today accrues real interest $\prod_{k=1}^{l} \frac{R_{t-k}}{\Pi_{t-1+1}}$.

We follow Johnson et al. (2006) and treat the residual expenditure category as not responding to the rebate. Thus, hand-to-mouth households simply consume a constant amount of this good,

$$C_{2t}^m = C_2^m$$

The marginal utility of consumption for the hand-to-mouth household is

$$\lambda_t^m = (C_t^m)^{-\frac{1}{\sigma}}$$

where C_t^m is the sum of the two types of consumption.

A.3 Wages

A continuum of unions indexed by j provide differentiated labor services to the final good firm that are subsitutable with elasticity ϵ^w. Each period there is a iid probability θ^w that the union cannot adjust the contract wage. In this case, wages will adjust by a fraction χ^w of last periods inflation.

The union imposes the same work hours on optimizing and hand-to-mouth households:

$$H_t^m = H_t^o = H_t$$

The demand for hours from union j at time $t+s$ conditional on having last reset wages at time t is

$$H^d_{t+s}(j) = H^d_{t+s}\left(\frac{W_t(j)(\frac{P_{t+s-1}}{P_{t-1}})^{\chi^w}(\frac{P_t}{P_{t+s}})}{W_{t+s}}\right)^{-\epsilon^w} = H^d_{t+s}W^{\epsilon^w}_{t+s}\left(\frac{P_{t+s}}{P_t}\right)^{\epsilon^w}\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{-\epsilon^w\chi^w} W_t(j)^{-\epsilon^w}$$

where P_t is the price level at time t.

If the union can adjust its wage at time t it picks the optimal wage to maximize the expected discounted utility of the representative household while this wage prevails:

$$\max_{w^*_t}\sum_{s=0}^{\infty}(\beta\theta^w)^s H^d_{t+s}W^{\epsilon^w}_{t+s}\left(\frac{P_{t+s}}{P_t}\right)^{\epsilon^w}\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{-\epsilon^w\chi^w}\left[\tilde{\lambda}_{t+s}\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{\chi^w}\left(\frac{P_{t+s}}{P_t}\right)^{-1}(W^*_t)^{1-\epsilon^w} - \nu H^{\phi}_{t+s}(W^*_t)^{-\epsilon^w}\right]$$

where $\tilde{\lambda} = (1-\gamma)\lambda_t + \gamma\lambda^m_t$

The first order condition for the union is:

$$\begin{aligned}
&(\epsilon^w-1)\sum_{s=0}^{\infty}(\beta\theta^w)^s H^d_{t+s}W^{\epsilon^w}_{t+s}\left(\frac{P_{t+s}}{P_t}\right)^{\epsilon^w-1}\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{-\chi^w(\epsilon^w-1)}\tilde{\lambda}_{t+s}(W^*_t)^{1-\epsilon^w}\\
&= \epsilon^w\nu\sum_{s=0}^{\infty}(\beta\theta^w)^s H^d_{t+s}H^{\phi}_{t+s}W^{\epsilon^w}_{t+s}\left(\frac{P_{t+s}}{P_t}\right)^{\epsilon^w}\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{-\epsilon^w\chi^w}(W^*_t)^{-\epsilon^w}
\end{aligned}$$

We write it recursively using

$$\begin{aligned}
F_{1t} &= \nu H^d_t H^{\phi}_t W^{\epsilon^w}_t (W^*_t)^{-\epsilon^w} + \beta\theta^w\Pi^{\epsilon^w}_{t+1}\Pi_t^{-\chi^w\epsilon^w}\left(\frac{W^*_t}{W^*_{t+1}}\right)^{-\epsilon^w}F_{1,t+1}\\
F_{2t} &= H^d_t W^{\epsilon^w}_t\tilde{\lambda}_t(W^*_t)^{1-\epsilon^w} + \beta\theta^w\Pi^{\epsilon^w-1}_{t+1}\Pi_t^{-\chi^w(\epsilon^w-1)}\left(\frac{W^*_t}{W^*_{t+1}}\right)^{1-\epsilon^w}F_{2,t+1}\\
\epsilon^w F_{1t} &= (\epsilon^w-1)F_{2t}
\end{aligned}$$

Wage dispersion across unions lead to inefficiency in the labor types used by firms. This creates a wedge between hours worked H_t and effective hours worked H^d_t, which we denote by s^w_t,

$$H_t = s^w_t H^d_t,$$

and which evolves according to,

$$s^w_t = (1-\theta^w)\left(\frac{W^*_t}{W_t}\right)^{-\epsilon^w} + \theta\left(\frac{W_{t-1}}{W_t}\right)^{-\epsilon^w}\Pi^{\epsilon^w}_t s^w_{t-1}$$

A.4 Production of capital goods

The representative capital goods firm chooses investment I_t, the capital stock K_t, and the utilization rate u_t to maximize profits,

$$\begin{aligned}
&\max_{\{K_{t+s},I_{t+s},u_{t+s}\}} \sum_{s=0}^{\infty} \beta^s \lambda_{t+s} \text{Profits}^k_{t+s} \\
&\text{s.t. } \text{Profits}^k_t = R^k_t u_t K_{t-1} - I_t \\
&\qquad K_t = (1-\delta(u_t))K_{t-1} + I_t\left[1 - S\left(\frac{I_t}{I_{t-1}}\right)\right]
\end{aligned}$$

where R^k_{t+s} is the rental rate of capital paid by the final goods firm, $S\left(\frac{I_t}{I_{t-1}}\right)$ is an investment adjustment cost, and $\delta(u)$ is the depreciation rate of capital which is increasing in utilization.

Let ζ_t be the Lagrange multiplier on the capital accumulation equation and define Tobin's q as the relative value of capital to nondurable consumption,

$$q_t = \frac{\zeta_t}{\lambda^o_t}.$$

Then the first order conditions for the representative capital producing firms are,

$$\begin{aligned}
1 &= q_t\left[1 - S\left(\frac{I_t}{I_{t-1}}\right) - \left(\frac{I_t}{I_{t-1}}\right)S'\left(\frac{I_t}{I_{t-1}}\right)\right] + \beta\frac{\lambda_{t+1}}{\lambda_t}q_{t+1}\left(\frac{I_{t+1}}{I_t}\right)^2 S'\left(\frac{I_{t+1}}{I_t}\right) \\
q_t &= \beta\frac{\lambda_{t+1}}{\lambda_t}R^k_{t+1}u_{t+1} + \beta(1-\delta(u_{t+1}))\frac{\lambda_{t+1}}{\lambda_t}q_{t+1} \\
R^k_t &= \delta'(u_t)q_t
\end{aligned}$$

A.5 Production of final goods

Final output Y_t is produced using a Cobb-Douglas production function with capital share α,

$$s_t Y_t = Z_t(u_t K_{t-1})^\alpha (H^d_t)^{1-\alpha}$$

where Z_t is aggregate TFP. The wedge s_t captures a distortion from price dispersion, which is described below.

The cost minimization for the representative final goods firm is

$$\begin{aligned}
&\min R^k_t u_t K_{t-1} + W_t H^d_t \\
&\text{s.t. } Z_t(u_t K_{t-1})^\alpha (H^d_t)^{1-\alpha} = s_t Y_t
\end{aligned}$$

which yields the following first order conditions for capital and labor,

$$R_t^k = \xi_t \alpha \frac{s_t Y_t}{u_t K_{t-1}}$$
$$W_t = \xi_t (1-\alpha) \frac{s_t Y_t}{H_t^d}$$

where ξ_t is the Lagrange multiplier on the production function. Dividing the two first order conditions yields the optimal capital-labor ratio,

$$\frac{u_t K_{t-1}}{H_t^d} = \frac{\alpha}{1-\alpha} \frac{W_t}{R_t^k},$$

which in turn yields the marginal cost of output is,

$$MC_t = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}(R_t^k)^{\alpha} W_t^{1-\alpha} \frac{1}{Z_t}$$

With perfect competition among final goods firms, the real final goods price is equal to marginal cost,

$$p_t^f = MC_t,$$

and final good firms make zero profits.

A.6 Prices

A continuum of retailers purchases final goods at price p_t^f and differentiates these goods with elasticity of substitution ϵ. Retailers can only reset their price with probability θ. The profit maximization problem for setting the reset price is

$$\max_{p_t^*} \sum_{s=0}^{\infty} \beta^s \left(\frac{\lambda_{t+s}}{\lambda_t} \right) \theta^s Y_{t+s} \left[(p_t^*)^{1-\epsilon} \left(\frac{P_{t+s}}{P_t} \right)^{\epsilon-1} - (p_t^*)^{-\epsilon} \left(\frac{P_{t+s}}{P_t} \right)^{\epsilon} p_{t+s}^f \right]$$

The first order condition for the optimal reset price is

$$\epsilon \sum_{s=0}^{\infty} \beta^s \left(\frac{\lambda_{t+s}}{\lambda_t} \right) \theta^s Y_{t+s} \left(\frac{P_{t+s}}{P_t} \right)^{\epsilon} (p_t^*)^{-\epsilon-1} p_{t+s}^f = (\epsilon - 1) \sum_{s=0}^{\infty} \beta^s \left(\frac{\lambda_{t+s}}{\lambda_t} \right) \theta^s Y_{t+s} \left(\frac{P_{t+s}}{P_t} \right)^{\epsilon-1} (p_t^*)^{-\epsilon}$$

which we write recursively as

$$
\begin{aligned}
X_{1t} &= Y_t p_t^f (p_t^*)^{-\epsilon-1} + \beta\theta \left(\frac{\lambda_{t+1}}{\lambda_t}\right)\left(\frac{P_{t+1}}{P_t}\right)^{\epsilon}\left(\frac{p_t^*}{p_{t+1}^*}\right)^{-\epsilon-1} X_{1,t+1} \\
X_{2t} &= Y_t (p_t^*)^{-\epsilon} + \beta\theta \left(\frac{\lambda_{t+1}}{\lambda_t}\right)\left(\frac{P_{t+1}}{P_t}\right)^{\epsilon-1}\left(\frac{p_t^*}{p_{t+1}^*}\right)^{-\epsilon} X_{2,t+1} \\
\epsilon X_{1t} &= (\epsilon-1) X_{2t}
\end{aligned}
$$

The optimal reset price determines aggregate inflation

$$1 = (1-\theta)(p_t^*)^{1-\epsilon} + \theta \Pi_t^{-(1-\epsilon)}$$

as well as the relative price distortion

$$
\begin{aligned}
s_t &= \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} di \\
&= (1-\theta)(p_t^*)^{-\epsilon} + \theta \int_0^1 \left(\frac{P_{t-1}(i)}{P_t}\right)^{-\epsilon} di \\
&= (1-\theta)(p_t^*)^{-\epsilon} + \theta \Pi_t^{\epsilon} s_{t-1}
\end{aligned}
$$

Due to monopoly power, the sticky-price firms make non-zero profits in equilibrium equal to

$$\text{Profits}_t^s = Y_t(1-p_t^f)$$

A.7 Government

The central bank sets the gross nominal interest rate according to the following interest rate rule,

$$R_t = (1-\rho_r)R_{t-1} + \rho_r \left[R + \phi_\pi(\Pi_t - \bar{\Pi}) + \phi_y\left(\frac{Y_t}{\bar{Y}} - 1\right)\right]$$

where ρ_r determines the degree of interest rate smoothing, ϕ_π the response to deviations of inflation from target, and ϕ_y the response to deviations of output from target.

The government issues one-period nominal bonds at gross interest R_t to cover debt repayment and any fiscal deficit.

$$B_t = \frac{R_{t-1}}{\Pi_t} B_{t-1} - T_t$$

To balance the budget over time, taxes are an increasing function of the debt level,

$$T_t = T + \phi_b(B_{t-k} - \bar{B}) - \epsilon_t.$$

We allow for a lag of k periods in the response of taxes to debt. The shock ϵ_t represents a one-time deficit financed transfer from the government to households.

A.8 Market Clearing

The goods market clears if total expenditure equals output.

$$Y_t = C_t + I_t + X_t$$

The bond market clears of bonds supplied by the government equal bonds held by households,

$$B_t = A_t$$

A.9 Functional Forms

We assume the following functional forms:

$$\delta(u_t) = \delta_0 + \delta_1(u_t - 1) + \delta_2(u_t - 1)^2$$
$$S\left(\frac{I_t}{I_{t-1}}\right) = \frac{\kappa}{2}\left(\frac{I_t}{I_{t-1}} - 1\right)^2$$

B Empirical Appendix

B.1 MPCs Estimated by BEA Consumer Expenditure Subgroups

Table B.1. MPC by BEA Subcategory

BEA Category		MPC	S.E.	JPS (2006) Non-durable
Total BEA Expenditures		**-0.11**	**0.59**	
BEA Durables		**-0.20**	**0.51**	
Furnishings		0.11	0.22	
	Small electric appliances	0.00	0.00	No
	Appliances	0.01	0.02	No
	Dishes/flatware	0.00	0.01	No

	Floor coverings	0.05	0.04	No
	Furniture	0.01	0.10	No
	Lighting and Decorative items	0.07	0.07	No
	Nonelectric cookware	0.00	0.00	No
	Tools/hardware	-0.01	0.01	No
	Window covers	-0.02	0.01	No
Motor vehicles and parts		-0.26	0.41	
	New autos	-0.25	0.25	No
	New light trucks	-0.15	0.27	No
	Tires	0.00	0.01	No
	Used autos	0.12	0.20	No
	Used light trucks	0.02	0.21	No
	Vehicle Parts	0.00	0.01	No
Other durable goods		0.06	0.03	
	Educ. books	0.01	0.01	No
	Therapeutic medical equip.	0.00	0.01	Yes
	Eyeglasses and contacts	0.02	0.01	Yes
	Jewelry	0.03	0.02	Yes
	Luggage	0.00	0.00	No
	Telephone equip.	0.00	0.00	No
	Watches	0.01	0.00	Yes
Recreational goods and vehicles		-0.11	0.19	
	Audio equipment	0.01	0.01	No
	Bicycles and accessories	0.00	0.00	No
	Other rec. vehicles	-0.16	0.18	No
	Other video equip.	0.00	0.00	No
	Personal computers and equipment	0.03	0.02	No
	Photo. equipment	-0.01	0.00	No
	Pleasure boats	0.00	0.04	No
	Rec. books	0.00	0.00	Yes
	Info process equip.	0.00	0.00	No
	Motorcycles	-0.03	0.04	No
	Music instruments	0.00	0.01	No
	Purchased music	0.00	0.00	No
	Software and accessories	0.00	0.00	No
	Sporting equip.	0.03	0.04	No
	TV	0.02	0.01	No
	Video discs/other	0.00	0.00	No
BEA Nondurable Goods		**0.06**	**0.06**	

Clothing		0.04	0.03	
	Children's clothing	0.01	0.00	Yes
	Clothing materials	0.00	0.00	Yes
	Footwear	0.01	0.01	Yes
	Men's clothing	0.01	0.02	Yes
	Women's clothing	0.01	0.02	Yes
Food and beverages at home		0.01	0.04	
	Alcohol (home)	0.00	0.01	Yes
	Food at home	0.01	0.04	Yes
Gasoline and other energy goods		-0.02	0.02	
	Other fuels	0.00	0.00	Yes
	Fuel oil	0.01	0.01	Yes
	Gasoline	-0.03	0.02	Yes
	Lubricants	0.00	0.00	Yes
Other nondurable goods		0.02	0.02	
	Cleaning products	0.00	0.00	No
	Other medical products	0.00	0.00	No
	Pets and pet products	0.00	0.01	No
	Prescription drugs	0.02	0.01	No
	Sewing items	0.00	0.00	No
	Horticulture goods	0.00	0.00	No
	Nonelectric personal care	0.00	0.00	Yes
	Photo supplies	0.00	0.00	No
	Electric personal care	0.00	0.00	Yes
	Games	0.00	0.01	No
	Linens	0.01	0.01	No
	Misc. printed material	0.00	0.00	No
	Newspapers	0.01	0.00	Yes
	Tobacco	-0.01	0.01	Yes
BEA Services		**0.03**	**0.22**	
Food services and accomodations		0.08	0.06	
	Hotels	0.04	0.03	No
	Alcohol (restaurants)	-0.01	0.01	Yes
	Food (restaurants)	0.00	0.03	Yes
	Housing at school	0.01	0.02	No
	School meals	0.03	0.01	Yes
Financial services and insurance		0.00	0.05	
	Financial service charges	0.00	0.00	Yes

	Net household insurance	0.03	0.01	No
	Life insurance	-0.02	0.04	No
	Medical care	0.01	0.02	Yes
	Net vehicle insurance	-0.02	0.03	No
Health care		0.07	0.03	
	All other medical services	0.00	0.00	Yes
	Hospitals	0.04	0.02	Yes
	Med. Lab.	0.00	0.01	Yes
	Nursing homes	0.01	0.01	Yes
	Physician services	0.00	0.01	Yes
	Dental	0.03	0.02	Yes
	Home healthcare	0.00	0.01	Yes
	"Outpatient care and allied services"	0.00	0.01	Yes
Housing and utilities		-0.08	0.13	
	Electricity	0.01	0.01	Yes
	Natural gas	-0.02	0.01	Yes
	Garbage	0.00	0.00	Yes
	Group housing	0.01	0.01	No
	Owner imputed rent	-0.11	0.13	No
	Rent	0.02	0.03	No
	Water/sewage	0.01	0.00	Yes
Other services		0.05	0.12	
	Child care	0.00	0.01	Yes
	Clothing repair, rental, and alterations	0.00	0.00	Yes
	Commercial and vocational schools	0.01	0.01	No
	Day care	0.00	0.01	Yes
	Higher education	0.00	0.04	No
	Miscellaneous personal care	0.00	0.00	Yes
	Other household services	0.00	0.01	Yes
	Social and civic organizations	0.01	0.02	No
	Social assistance	0.01	0.03	Yes
	Footwear repair	0.00	0.00	Yes
	K-12 Schools	-0.03	0.04	No
	Cell service	0.01	0.01	Yes
	Domestic services	0.01	0.01	Yes
	Foreign travel	0.02	0.02	No
	Funeral service	0.00	0.05	Yes
	Furniture repair	0.00	0.00	No

	Internet	0.00	0.00	Yes
	Laundry	0.00	0.01	Yes
	Legal service	0.02	0.05	Yes
	Moving	0.00	0.01	Yes
	Repair of appliances	0.00	0.00	No
	Salons	0.01	0.00	Yes
	Tax Services	-0.01	0.01	Yes
	Telecommunication service	0.00	0.01	Yes
Recreation services		-0.08	0.08	
	Cable or satellite television and radio	0.00	0.00	No
	Gambling	0.00	0.01	Yes
	Membership clubs	-0.10	0.08	No
	Photo processing	0.00	0.00	No
	Photo studios	0.00	0.00	No
	Spectator sports	0.01	0.01	No
	Amusement parks and other	0.00	0.01	No
	Cinema and other entertainment	-0.01	0.01	No
	Maintenance rec. equip.	0.00	0.00	No
	Repair TV and audio equip.	0.00	0.00	No
	Vet service	0.01	0.01	No
	Video rental	0.00	0.00	No
Transportation services		0.00	0.05	
	Intercity buses	0.00	0.00	Yes
	Motor vehicle maintenance	-0.02	0.03	No
	Parking and tolls	0.00	0.00	No
	Water transportation	0.00	0.00	Yes
	Motor vehicle rental	-0.01	0.01	No
	Air	0.02	0.02	Yes
	Intracity mass transit	0.00	0.00	Yes
	Motor vehicle leasing	0.00	0.03	No
	Taxi	0.01	0.00	Yes
	Train	0.00	0.01	Yes

Note: From 2SLS regressions of consumption category on rebate amount. Subcategory MPCs may not sum to category MPCs due to rounding. BEA category names have been shortened.