ANNALS OF ECONOMICS AND FINANCE **26-2**, 643–665 (2025)

Disaggregation Reverses the Risk-Free Rate Puzzle

Matthew S. Wilson*

In the macro data, it appears that people are buying too many bonds. They help agents smooth consumption, but the rate of return is low and aggregate consumption is already very smooth. However, consumption at the household level is far more volatile. This makes bonds more appealing. Instead of solving the puzzle, the main result is reversed: households are not buying enough bonds. Under some specifications, the unconditional Euler equation is not rejected, but in those cases the forecast errors are inconsistent with rational expectations.

Key Words: Risk-free rate puzzle; Bonds; Interest rates; Consumer Expenditure Survey.

JEL Classification Numbers: E4, G1.

1. INTRODUCTION

There is a wide variety of macro models, but they tend to share the same foundations. The representative agent has a CES utility function. The discount factor, β, is 0.99 if the data is quarterly and 0.96 if it is annual. The coefficient of relative risk aversion is between one and four. This framework has several implications for asset pricing. Unfortunately, the model fails many of these tests. For reasonable values of the discount factor and CES parameter, the model generates a risk-free rate that is too high. Alternatively, the model could take the risk-free rate from the data as given and attempt to explain people's bond purchases. In that case, they are buying too many bonds. This is the risk-free rate puzzle.

This simple asset-pricing test appears to refute the standard macro framework. Perhaps this is proof that we need to fundamentally change the assumptions of macroeconomic theory. But before we discard the all the research built on these foundations, a closer investigation is warranted. Is

* Economics Department, University of Richmond, 102 UR Dr, Richmond, VA 23173. Email: mwilson7@richmond.edu. George Evans, Bruce McGough, anonymous reviewers, and seminar participants at the University of Oregon Macro Group and Abilene Christian University provided helpful feedback and discussion. Any remaining errors are my own.

1529-7373/2025

there a small change to the framework that can solve the puzzle? For instance, some papers test the model with micro data instead of macro data. As I demonstrate later, this would be appropriate if markets are incomplete. There is measurement error in the household data, but I show how to correct for it.

I fail to solve the puzzle. However, my results indicate a new direction for future research. Many papers have attempted to explain why households are buying so many bonds. However, in the household data, the puzzle is flipped. Now it appears that households are not buying enough bonds. Solving this new puzzle will require a different approach. I also demonstrate the role of seasonality. In some cases, lowering the discount factor appears to solve the puzzle. However, then there is autocorrelation in the agents' forecast errors. This violates the rational expectations hypothesis; people should not be consistently overestimating or underestimating. Most of the previous literature does not test for autocorrelation.

Many papers address the puzzle by modifying the utility function (e.g. Weil 1989, Boldrin et al. 2001, Epstein and Zin 1990). Huggett (1993) and Telmer (1993) explore the effects of incomplete markets. Bansal and Coleman (1996) study transaction costs. Gabaix (2012) uses rare disasters to explain many puzzles. However, all of these papers are based on macro data. The puzzle in the micro data is different.

Kocherlakota and Pistaferri (2009) use household data. They show that the puzzle can be solved only if the discount factor is unrealistically low — less than 0.5. Wilson (2020) notes that their tests are biased. Furthermore, Kollmann (2009) and Basu et al. (2011) demonstrate that their conclusions are not robust to changes in the sample design. Constantinides and Ghosh (2017) add an extensive set of assumptions to the standard macro framework and estimate 14 parameters with GMM. There is no consensus on how to deal with outliers in the household survey data. Therefore, I also explore if my results are sensitive to changes in the outlier criteria. The point estimates change, but the overall conclusion is unaffected.

Barr et al. (2012) extend the framework of Kocherlakota and Pisaferri (2009) to analyze inflation-indexed bonds in the UK. Jacobs (1999) uses micro data to study the risk-free rate puzzle. His source is the Panel Survey of Income Dynamics (PSID), which only tracks food consumption, not total consumption. Wilson (2020) also works with household data, but his focus is on the equity premium puzzle. He demonstrates the effect of measurement error and suggests a solution. With some modifications, I follow his approach.

Within the asset pricing literature, there is a subfield that focuses on seasonality. Kamstra et al. (2015, 2017) document seasonal patterns in asset holdings and treasury bond yields. The non-seasonally adjusted data is much more volatile. Consequently, bonds become more attractive due to their stability. This can potentially solve the equity premium and risk-free rate puzzles. Ferson and Harvey (1992) generalized Miron's (1986) model and studied the equity premium. While seasonality plays a substantial role, these changes do not solve the puzzle. Heaton (1995) reaches the same conclusion. I find that there are significant seasonal effects, but the puzzle remains even after accounting for them.

I build on Kocherlakota and Pistaferri (2009) and Jacobs (1999), who analyze the risk-free rate puzzle with household data. With an approach similar to Wilson (2020), I correct for the measurement error bias that Kocherlakota and Pistaferri (2009) neglected. My consumption data is broader than Jacobs (1999), who only considers food consumption. Additionally, I examine the impact of seasonality in the micro data. The asset pricing and seasonality literature focuses on macro data. Household consumption is more volatile, which increases the appeal of bonds. Unfortunately, this reverses the puzzle instead of solving it. The model's risk-free rate is too low.

2. THE RISK-FREE RATE PUZZLE IN THE MACRO DATA

The model is rejected under all conditions: quarterly, annual, seasonally adjusted, and not seasonally adjusted.

Each household i has the same CRRA utility function. At time t, it receives income y_{it}. The gross real rate of return is R_t; the household earns this on the bonds it held in the previous period ($b_{i,t-1}$). The household decides on how to allocate this money between consumption (c_{it}) and buying bonds (b_{it}). The term z_{it} may include seasonal variables that affect preferences.

$$\max E_{t_0} \sum_{t=t_0}^{\infty} \beta^t \left(\frac{c_{it}^{1-\sigma} - 1}{1-\sigma} \right) e^{z_{it}} \text{ subject to } c_{it} + b_{it} = y_{it} + R_t b_{i,t-1}. \quad (1)$$

The household treats y_{it} and R_t as exogenous. Then we get the Euler equation.

$$e^{z_{it}} c_{it}^{-\sigma} = \beta e^{z_{i,t+1}} E_t R_{t+1} c_{i,t+1}^{-\sigma} \quad (2)$$

After applying the Law of Iterated Expectations, the Euler equation can be expressed as follows.

$$E\left(\left(\frac{e^{z_{i,t+1}}}{e^{z_{it}}} \right) \left(\frac{c_{i,t+1}}{c_{it}} \right)^{-\sigma} R_{t+1} \right) - \beta^{-1} = 0 \quad (3)$$

Aggregate consumption per capita in time t is C_t. The ratio of household to aggregate consumption is h_{it}. This can be plugged into equation (3):

$$h_{it} = \frac{c_{it}}{C_t} \tag{4}$$

$$E\left(\left(\frac{e^{z_{i,t+1}}}{e^{z_{it}}}\right)\left(\frac{h_{i,t+1}C_{t+1}}{h_{it}C_t}\right)^{-\sigma} R_{t+1}\right) - \beta^{-1} = 0. \tag{5}$$

If markets are complete, then households can buy insurance against idiosyncratic income shocks. As a result, household consumption will depend only on aggregate shocks and $h_{i,t+1}/h_{it}$ will be equal to one. This is the justification for using macro data. The proof is in Wilson (2020). If instead markets are incomplete, then there is no guarantee that $h_{i,t+1}/h_{it}$ will be constant, so in Section 3, we will turn to the household-level data.

If there are complete markets, then there is a representative agent and equation (5) can be rewritten as follows.

$$E\left(\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma} R_{t+1}\right) - \beta^{-1} \equiv E(M) = 0. \tag{6}$$

Following the macro orthodoxy, I set the quarterly discount factor β to 0.99. Later I will consider other values for β. The values for R_t come from the real rate of return on 90-day Treasury bills. This is also standard in the macro literature (Mehra and Prescott 1985, Kocherlakota 1996).

Following the literature, I only consider consumption of nondurables and services. Durable consumption is harder to measure. For example, the purchase of a new car is recorded as a one-time jump in expenditures. However, the consumption of the car persists for several years. The time periods do not match. On the other hand, a banana bought in the first quarter will be consumed in the first quarter; nondurable expenditures match nondurable consumption.

The sample period is 1989-2013, excluding the first quarter of 1996. This time period was selected in order to match the micro data in Section 3. As explained later, the first quarter of 1996 was dropped due to a change in the household survey. The results in the macro data are not sensitive to whether that quarter is included or not.

Following Kocherlakota (1996), I try several different calibrations and test whether the Euler equation is rejected. I.e., instead of estimating the CES parameter, I plug in different values for it. Wilson (2020) explains the advantages of this approach over GMM. First, Toda and Walsh (2015, 2017) demonstrate that measurement error biases GMM estimates in the household data. This section is about the macro data, but for comparability, I use the same methods for both datasets. Second, the model might still be rejected even for the best-fitting value of σ, an issue that Jacobs (1999) mentions. Wilson (2020) provides a third reason. It is important to show if parameter values currently in use are justified. For instance, a popular utility function is log preferences ($\sigma = 1$). Perhaps a different value of σ fits the data better, but $\sigma = 1$ may still be defensible.

For CRRA utility, it is well known that the elasticity of intertemporal substitution (EIS) is the inverse of the relative risk aversion (i.e., $1/\sigma$). I present the results in terms of CRRA parameter since it seemed natural to use whole number increments. Vissing-Jørgensen (2002) summarized the empirical research on the EIS and does not see any evidence for it being greater than one. It is possible to show larger values of σ. However, this would cause problems in the household data that comes later. The next section illustrates that when measurement error is raised to a high exponent, there is too much noise to reliably correct for the bias.

TABLE 1.
The Risk-Free Rate Puzzle, based on quarterly aggregate data. The model predicts that the mean of M is zero. $N = 99$ quarters

Coeff. of Relative	**SA**		**NSA, NSP**		**NSA, SP**	
Risk Aversion (σ)	**Mean M**	**P-value**	**Mean M**	**P-value**	**Mean M**	**P-value**
1	−0.0084 (0.0011)	0.000	−0.0067 (0.0099)	0.500	−0.0094 (0.0025)	0.000
2	−0.0083 (0.0015)	0.000	0.0052 (0.0212)	0.806	−0.0101 (0.0044)	0.024
3	−0.0082 (0.0021)	0.000	0.0284 (0.0337)	0.400	−0.0104 (0.0064)	0.109
4	−0.0081 (0.0027)	0.004	0.0642 (0.0479)	0.183	−0.0103 (0.0085)	0.228

In Table 1, I test equation (6). First, I show the results for the seasonally adjusted (SA) consumption data. z_t is set to zero; we do not have to control for seasonality if the data is already seasonally adjusted. For all reasonable values of the CRRA parameter, the mean of M is significantly less than zero in the SA data.

$$\beta E_t(R_{t+1} MU_{t+1}) - MU_t < 0. \tag{7}$$

In other words, the marginal utility of consuming today exceeds the expected marginal utility of buying bonds and consuming later. It looks like people should be buying fewer bonds.

The middle columns show the results with quarterly non-seasonally adjusted (NSA) data and no seasonal preferences (NSP), which means that $z_t = 0$. Seasonal adjustment smooths the consumption data. This diminishes the appeal of bonds. The rate of return is very low; agents mainly buy bonds because they are a safe way to smooth consumption. However, that is unnecessary if consumption is already very smooth. The NSA data is volatile; this should motivate people to buy bonds. The point estimates are positive, indicating that people should be buying more of them. However, the standard errors are large enough that the Euler equation is not rejected.

In the last columns of Table 1, I included seasonal preferences z_t. I.e., I assumed that z_t was a linear function of four quarterly dummy variables[1]. This is shown in Equation (8):

$$z_t = D_{Q1,t} + \theta_2 D_{Q2,t} + \theta_3 D_{Q3,t} + \theta_4 D_{Q4,t}. \tag{8}$$

The Euler equation with seasonal preferences z_t is shown in Equation (9). I used GMM to estimate the seasonal θ terms. For the rest of the paper, "SP" refers to "seasonal preferences" (i.e. estimating z_t), while "NSP" means "no seasonal preferences" (i.e. $z_t = 0$).

$$\begin{aligned} & E\left(\left(\frac{e^{D_{Q1,t+1}+\theta_2 D_{Q2,t+1}+\theta_3 D_{Q3,t+1}+\theta_4 D_{Q4,t+1}}}{e^{D_{Q1,t}+\theta_2 D_{Q2,t}+\theta_3 D_{Q3,t}+\theta_4 D_{Q4,t}}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{\sigma} R_(t+1)\right) - \beta^{-1} \\ \equiv\ & E(M) = 0 \end{aligned} \tag{9}$$

While consumption fluctuates in the NSA data, marginal utility should be more stable. Some of the swings are due to seasonal preferences, e.g., Christmas shopping and summer vacation.

Nevertheless, the standard errors are still higher than in the SA data. Once seasonal preferences are incorporated, NSA consumption stabilizes; bonds become less attractive. The point estimates of M are negative, which means that people should buy fewer bonds. However, the variance is high, so the Euler equation is not always rejected. This makes it easier for the model to pass the test, but as I will show later, the puzzle remains unsolved.

To gain intuition, it is useful to decompose the Euler equation, as shown in Equation (10). In Equation (11), I rearranged it to find the model's

[1] There is no coefficient on $D_{Q1,t}$, since then the coefficients cannot be identified. $\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)$ would be $e^{\theta_{t+1}-\theta_t}$. If $\hat{\theta}$ minimizes the GMM criterion, then so does $\hat{\theta} = \hat{\theta} + k$. Thus, there is no unique solution. The CES parameter σ is calibrated. This is because Toda and Walsh (2015) showed that GMM estimates of the CES parameter are biased in the household data.

risk-free rate[2]. In the data, the quarterly risk-free rate is 1.0017. This implies an annual rate of 0.7%.

$$Cov\left(\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}, R_{t+1}\right)+E\left[\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\right]E[R_{t+1}]-\beta^{-1}=0 \tag{10}$$

$$E[R_{t+1}] = \frac{\beta^{-1} - Cov\left(\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}, R_{t+1}\right)}{E\left[\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\right]} \tag{11}$$

Table 2 shows that the covariance term is not driving the difference. Instead, it is the volatility of the NSA data. Mathematically, this inflates the $E\left[\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\right]$ term. People dislike these fluctuations, so an investment that smooths consumption is more desirable here. Thus, in the NSA-NSP data, the model's risk-free rate is lower. However, it is still too high.

TABLE 2.

Summary statistics for SA and NSA data. $N = 99$ quarters

$\sigma = 1$	$Cov\left(\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-1}, R_{t+1}\right)$	$V\left[\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{-1}\right]$	R_{t+1}
SA	0.0000	0.00004	1.0101
NSA, NSP	−0.0004	0.01041	1.0085
NSA, SP	0.0000	0.00047	1.0112

Under some specifications, the model was not rejected in Table 1 when NSA data was used, so I ran additional tests. If agents have rational expectations, then there should be no autocorrelation in the forecast errors (Wilson 2020). Autocorrelation implies that $E[M_t|M_{t-1}, M_{t-2}, \ldots] \neq 0$, which is inconsistent with the following equation:

$$E_t\left(\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{C_{t+1}}{C_t}\right)^{\sigma} R_{t+1}\right) - \beta^{-1} \equiv E_t(M_{t+1}) = 0. \tag{12}$$

I ran a Breusch-Godfrey test for autocorrelation. It regresses the residuals on lags of themselves. If there is no autocorrelation, then the R^2 should be low. Table 3 presents the results.

[2]This involves dividing a sample covariance by a sample mean; Wilson (2020) proves that this is biased. I retain this section in order to compare with the literature, where computing the model's risk-free rate is common.

TABLE 3.

Breusch-Godfrey test. The model predicts that there is no autocorrelation. $N = 99$ quarters

	NSA, NSP		NSA, SP	
Coeff. of Relative Risk Aversion (σ)	χ^2	P-value	χ^2	P-value
1	35.668	0.000	4.413	0.036
2	33.202	0.000	6.727	0.010
3	30.290	0.000	7.360	0.007
4	27.432	0.000	7.516	0.006

In the NSA-NSP data, there is strong evidence of autocorrelation. The autocorrelation weakens when seasonal preferences are included, but it remains significant. This demonstrates that the autocorrelation issue is more than just seasonality.

To sidestep the seasonality issue, we could switch to annual data. However, this leads to "time aggregation bias" (Grossman et al. 1987). A researcher looking at the annual data only sees total consumption and not the ups and downs within the year. It looks like households are smoothing more than they really are. This deepens the risk-free rate puzzle. If consumption is already very smooth, then buying lots of bonds is unjustified and the Euler equation is rejected, as shown in Table 4.

TABLE 4.

The Risk-Free Rate Puzzle, based on annual aggregate data. The model predicts that the mean of M is zero. $N = 25$ years

Coeff. of Relative Risk Aversion (σ)	Mean M	P-value
1	-0.0405 (0.0055)	0.000
2	-0.0453 (0.0090)	0.000
3	-0.0496 (0.0130)	0.001
4	-0.0535 (0.0170)	0.005

Thus, there is a puzzle in the macro data. However, in order to aggregate, I assumed that markets were complete. The next step is to turn to the micro data, since that can be used even if markets are incomplete.

3. THE RISK-FREE RATE PUZZLE IN THE MICRO DATA

3.1. Data and Measurement Error

Every quarter, the BLS interviews about 5,000 households for its Consumer Expenditure Survey (CEX). Each household stays in the sample for five quarters, even if it stops responding to the survey. Once the five quarters elapse, it is replaced by a new household. The sample is staggered so that about 20% of the households exit the survey each quarter. There was a major change in the first quarter of 1996 (BLS 1997). Following the advice in the BLS documentation, I dropped this period from the sample. I used data from 1989-2014, matching the time period in the previous section. The respondents are not fully representative of the population; I correct for this by using the BLS's sample weights.

Consumption data is collected in all of the interviews except the first one. Since the interview usually happens in the middle of a quarter, it takes two interviews in order to piece together a quarter's consumption. For example, suppose I want to find $\frac{c_{i,t+1}}{c_{it}}$. I need interviews 2 and 3 to obtain c_{it} and interviews 3 and 4 for $c_{i,t+1}$. Similarly, interviews 3, 4, and 5 are required for $\frac{c_{i,t+2}}{c_{i,t+1}}$. Thus, I can only get two observations of a household's quarterly consumption growth. For many households, there is only one observation because they did not complete all the interviews. Because there are so few observations per household, Wilson (2020) argued that it is better to cluster over time rather than cluster over households. The motivation is that aggregate shocks such as recessions affect all households at the same time. For the rest of the paper, I will cluster over time.

In other papers using CEX data, there has been much disagreement about outliers. Each paper uses its own criterion for identifying outliers. Perhaps this is why their conclusions are so different: Cogley (2002) claims that the equity premium puzzle is not solved in the CEX data, while Balduzzi and Yao (2007) and Brav et al. (2002) disagree. Earlier, I mentioned that the results in Basu et al. (2011) and Kollman (2009) did not match Kocherlakota and Pistaferri (2009). Due to these controversies, I tried several different criteria for excluding outliers. For my baseline results, I dropped households that spent less than $1000 per person on nondurables. Even if they are very frugal, they must be underreporting their spending on three months of housing and food. Other households reported spending fluctuations that were implausibly volatile. I dropped observations if nondurable consumption growth was less than 0.5 or greater than two. For comparison, I also reran the results using the outlier criteria in Brav et al. (2002) and Kocherlakota and Pistaferri (2009).

I exclude observations that have obvious errors. Specifically, I dropped households that reported negative consumption. Additionally, I drop incomplete income reporters and "topcoded" consumption. In order to protect confidentiality, the BLS censors (or "topcodes") some observations that are above a given threshold (BLS 2014). For example, credit card debt higher than \$51,323 is recorded as \$51,323 (BLS 2014). In most cases, topcoding only affects income data rather than consumption; I only drop the observations if the reported consumption is altered.

According to Branch (1994) and BLS (1994), the CEX captures about 95% of the households' total consumption. It also includes data on housing consumption. For homeowners, housing consumption is the rent they would receive if they were to lease their home. In spite of its strengths, there is measurement error in the CEX. Household i's true consumption is c_{it}. On the CEX survey, it reports consumption $\hat{c}_{it}$. The measurement error term is ε_{it}.

$$c_{it} = \hat{c}_{it}\varepsilon_{it}. \tag{13}$$

If the measurement error is i.i.d., then the true Euler equation is (15). However, a researcher who does not observe the measurement error might test (16).

$$\eta = \left(\frac{\varepsilon_{i,t+1}}{\varepsilon_{it}}\right)^{-\sigma}, \tag{14}$$

$$E\left(\left(\frac{e^{z_{i,t+1}}}{e^{z_{it}}}\right)\left(\frac{\hat{c}_{i,t+1}}{\hat{c}_{it}}\right)^{-\sigma} R_{t+1}\right) E(\eta) - \beta^{-1} \equiv E(m) = 0, \tag{15}$$

$$E\left(\left(\frac{e^{z_{i,t+1}}}{e^{z_{it}}}\right)\left(\frac{\hat{c}_{i,t+1}}{\hat{c}_{it}}\right)^{-\sigma} R_{t+1}\right) - \beta^{-1} \equiv E(\hat{m}). \tag{16}$$

Even though $E(\varepsilon_{it}) = E(\varepsilon_{i,t+1})$, we cannot assume that the errors cancel out in equation (15). Brav et al. (2002) prove that $E(\eta) \geq 1$, so tests of (16) are biased. However, I find a way to correct for this. Rewriting the Euler equation and taking expected values:

$$\begin{aligned} m &= \eta(\hat{m} + \beta^{-1}) - \beta^{-1}. && (17)\\ E(m) &= E(\eta)E(\hat{m}) + \beta^{-1}(E(\eta) - 1). && (18) \end{aligned}$$

Though $E(\eta) \geq 1$, it is unclear if the true mean $E(m)$ is greater than $E(\hat{m})$. To gain intuition, revisit Equation (10). The measurement error makes the data more volatile. Consequently, bonds become more appealing due to their stability. This raises $E(\hat{m})$. On the other hand, the noise will drive the covariance term toward zero. In the macro data, the covariance term was already close to zero, but its true value in the micro data is unknown. Thus, the overall effect is ambiguous.

To perform hypothesis tests, I also have to correct the standard errors. By substitution:

$$V(m) = E[(\eta(\hat{m} + \beta^{-1}) - \beta^{-1} - E[\eta(\hat{m} + \beta^{-1}) - \beta^{-1})])^2]. \tag{19}$$

Equation (19) can be rewritten as follows, since the measurement error is independent.

$$V(m) = (V(\eta) + E(\eta)^2)(V(\hat{m}) + E(\hat{m} + \beta^{-1})^2) - E(\eta\hat{m} + \eta\beta^{-1})^2. \tag{20}$$

Now I exploit η's independence again in order to break up the last term and simplify.

$$V(m) = (V(\eta) + E(\eta)^2)V(\hat{m}) + V(\eta)((E(\hat{m}) + \beta^{-1})^2). \tag{21}$$

Because $E(\eta) \geq 1$, the true variance is larger than the variance of $\hat{m}$. This may be surprising, since usually measurement error inflates the variance. However, the risk-free rate puzzle is an exception since the mismeasurement affects the terms unevenly: $E\left(\left(\frac{c_{i,t+1}}{c_{it}}\right)^{-\sigma} R_{t+1}\right)$ is impacted but β^{-1} is not. This biases the mean and the variance. The special case of no measurement error corresponds to $V(\eta) = 0$ and $E(\eta) = 1$. If that is plugged in, then $V(m) = V(\hat{m})$; i.e., the corrected and uncorrected variances would be the same.

To make these adjustments, we have to know the mean and the variance of the measurement error. Based on Branch (1994) and BLS (1994), the mean of ε_{it} is 0.95^{-1}. Finding the variance is harder. Geisen et al. (2011) reinterviewed a subsample of CEX. They asked respondents to provide receipts for the purchases that they reported on the survey. However, the authors do not report the variance of the measurement error. After assuming a distribution, Wilson (2020) is able to deduce the variance in their paper. He studies two cases. In the baseline scenario, the measurement error is normally distributed and the standard deviation is 0.053. However, Toda and Walsh (2015) argue that there is a double power law in the data and propose a double Pareto distribution. In that case, Wilson (2020) finds that the standard deviation is 0.057. In addition, I also use the generalized Pareto distribution (Equation 22). Following the same procedure as Wilson (2020), I ran simulations and set $\alpha = 7.95$, $k = 2.8$, and $u = 0.6071$. With these parameters, I match the magnitude of overreporting, the magnitude of underreporting, and the average measurement error in Geisen et al. (2011).

$$f(x) = \begin{cases} \frac{\alpha\left(\frac{k+x-u}{k}\right)^{-1-\alpha}}{k}, & x \geq u \\ 0, & x < u \end{cases} \tag{22}$$

Since η is a ratio of measurement errors raised to an exponent, I could not calculate its moments analytically for any of the distributions. Instead, I used 10,000 simulations. Each simulation had 10,000 observations. The results are in Table 5.

TABLE 5.

Summary statistics for η. $N = 10,000$ simulations

	Normal Distribution		**Double Pareto**		**Generalized Pareto**	
Coeff. of Relative Risk Aversion (σ)	Mean	Variance	Mean	Variance	Mean	Variance
1	1.0031	0.0063	1.0037	0.0076	1.0064	0.0162
2	1.0126	0.0261	1.0150	0.0461	1.0291	0.0698
3	1.0287	0.0615	1.0344	0.1613	1.0692	0.1778
4	1.0514	0.1170	1.0642	113.0237	1.1288	0.3760

The double Pareto distribution has fatter tails. This generates more extreme outliers than the normal distribution, even though they have similar variances. In η, these outliers are raised to an exponent, magnifying their effects. This is why η has a higher mean and variance in the double Pareto scenario. The effect is especially strong when $\sigma = 4$. In that case, the corrected standard errors will be extremely large, so the tests will have almost no power. Consequently, I will restrict my attention to the range $\sigma \in [1, 3]$.

Though the corrections appear to be small, they are not negligible. Following Wilson (2020), I ran 10,000 simulations of 10,000 observations each; the true mean of m was set to zero. A hypothesis test should reject the null $E(m) = 0$ in 5% of the simulations. However, if I test $E(\hat{m}) = 0$, the null is rejected too frequently. When $\sigma = 1$, the null is rejected 8.18% of the time, and that is significantly greater than 5%. For $\sigma \in \{2, 3\}$, the measurement error has a larger effect, and the null is rejected even more frequently. Thus, if I were to rely upon tests of $E(\hat{m}) = 0$, it would be unclear what is driving my results: are households really buying too few bonds or are my tests biased? That is why the corrections are necessary.

3.2. Baseline Results

In the aggregate data, it appeared that people were buying too many bonds. I tested Equation (23) with household-level CEX data; the discount factor, β, was set to 0.99. The results are in Table 6. In all specifications, the model was rejected with a p-value less than 0.001.

$$E\left(\left(\frac{e^{z_{t+1}}}{e^{z_t}}\right)\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma} R_{t+1}\right) - \beta^{-1} \equiv E(m_{t+1}) = 0 \tag{23}$$

TABLE 6.
The Risk-Free Rate Puzzle, based on quarterly CEX data for nondurable household consumption. The model predicts that the mean of m is zero. $N = 219,879$

	Normal		**Double Pareto**		**Generalized Pareto**	
	NSP	SP	NSP	SP	NSP	SP
Coeff. of Relative Risk Aversion (σ)	Mean m	Mean m	Mean m	Mean m	Mean m	Mean m
1	0.0200 (0.0021)	0.0197 (0.0015)	0.0206 (0.0021)	0.0203 (0.0015)	0.0234 (0.0021)	0.0231 (0.0015)
2	0.1126 (0.0047)	0.1118 (0.0030)	0.1152 (0.0048)	0.1144 (0.0030)	0.1309 (0.0049)	0.1301 (0.0031)
3	0.2879 (0.0087)	0.2861 (0.0056)	0.2951 (0.0091)	0.2933 (0.0057)	0.3390 (0.0095)	0.3371 (0.0058)

The corrected means and standard errors are slightly larger if we assume that the distribution is Pareto instead of Gaussian. The gap widens when the exponent is increased. The double Pareto distribution generates more outliers; the biasing effect of the outliers is stronger when they are raised to a power. Thus, a bigger correction is needed for the double Pareto case, especially when the exponent is high.

Seasonal preferences contribute to the results. Due to Christmas, there is a surge in fourth quarter spending followed by a plunge in the first quarter; this adds volatility. Bonds help people smooth their consumption. This means that bonds become more appealing when there are high fluctuations. However, once seasonal preferences are incorporated, marginal utility is smoother, so bonds become less attractive. Thus, the mean of m falls, though it remains positive and significant. I.e., expected marginal utility in the future is greater than today's marginal utility. Thus, people could raise their utility by sacrificing present consumption to buy more bonds. In the data, people aren't buying too many bonds — they are buying too few. This flips the risk-free rate puzzle.

$$\beta E_t(MU_{t+1}R_{t+1}) > MU_t \tag{24}$$

In Table 7, I decompose the Euler equation as in Equation (10).

The covariance between the rate of return and marginal utility ratio is negligible. The volatility in the household data raises the expected marginal utility ratio. The growth rate is raised to an exponent, so the fluctuations do not cancel out. Instead, they push the mean upwards. This increases the mean of m, which causes the puzzle to reverse. Note that Table 7 does not imply that consumption is shrinking ($E[1/X] \neq 1/E[X]$). Consumption is rising and its growth rate is very similar to the PCE's. Strikingly, the

TABLE 7.

Summary statistics for household nondurable consumption. $N = 219,879$

	$Cov\left(\left(\frac{e^{z_{i,t+1}}}{e^{z_{it}}}\right)\left(\frac{c_{i,t+1}}{c_{it}}\right)^{-1}, R_{t+1}\right)$	$E\left[\left(\frac{e^{z_{i,t+1}}}{e^{z_{it}}}\right)\left(\frac{c_{i,t+1}}{c_{it}}\right)^{-1}\right]$	Risk-Free Rate
NSP	−0.0001	1.0257	0.9849
SP	0.0000	1.0255	0.9850

model's risk-free rate is below one. I.e., households will buy bonds even if they do not keep up with inflation. When volatility is high, households should be willing to sacrifice consumption in order to achieve stability.

However, the point estimates in Table 6 are sometimes very close to zero. This suggests that changing the discount factor (β) might solve the puzzle. With a lower discount factor, people would be more impatient and buy fewer bonds, potentially solving the puzzle. Table 8 shows several different calibrations. I do not consider $\sigma = 3$ because it would require a discount factor that is implausibly low. Back in Table 6, the sample mean was relatively far from zero when $\sigma = 3$, so β would have to be much lower than 0.99.

TABLE 8.

The Risk-Free Rate Puzzle, based on quarterly CEX data for nondurable consumption. The model predicts that the mean of m is zero. $N = 219,879$.

	Normal		**Double Pareto**		**Generalized Pareto**	
	NSP	SP	NSP	SP	NSP	SP
Parameters	Mean m	Mean m	Mean m	Mean m	Mean m	Mean m
$\beta = 0.96$,	-0.0116^{***}	-0.0114^{***}	-0.0110^{***}	-0.0108^{***}	-0.0082^{***}	-0.0080^{***}
$\sigma = 1$	(0.0021)	(0.0015)	(0.0021)	(0.0015)	(0.0021)	(0.0015)
$\beta = 0.97$,	−0.0008	−0.0006	−0.0002	−0.0001	0.0025	0.0027
$\sigma = 1$	(0.0021)	(0.0015)	(0.0021)	(0.0015)	(0.0021)	(0.0015)
$\beta = 0.98$,	0.0097^{***}	0.0099^{***}	0.0103^{***}	0.0105^{***}	0.0130^{***}	0.0132^{***}
$\sigma = 1$	(0.0021)	(0.0015)	(0.0021)	(0.0015)	(0.0021)	(0.0015)
$\beta = 0.88$,	-0.0137^{**}	-0.0135^{***}	-0.0110^{*}	-0.0109^{***}	0.0046	0.0047
$\sigma = 2$	(0.0047)	(0.0032)	(0.0048)	(0.0032)	(0.0049)	(0.0032)
$\beta = 0.89$,	−0.0009	−0.0008	0.0017	0.0019	0.0174^{***}	0.0175^{***}
$\sigma = 2$	(0.0047)	(0.0032)	(0.0048)	(0.0032)	(0.0049)	(0.0032)
$\beta = 0.90$,	0.0116^{*}	0.0117^{***}	0.0142^{**}	0.0144^{***}	0.0299^{***}	0.0300^{***}
$\sigma = 2$	(0.0047)	(0.0032)	(0.0048)	(0.0032)	(0.0049)	(0.0032)

* Significant at the 5% level
** Significant at the 1% level
*** Significant at the 0.1% level

When $\sigma = 1$, a small modification to β suffices. A value of $\beta = 0.97$ implies an annual discount factor of 0.89, which seems reasonable. If $\sigma = 2$, much larger changes are required. Now the quarterly discount factor β is 0.89, which leads to an annual discount factor of just 0.62. This is implausibly low. As an illustration, consider a worker planning to retire in 10 years. The discount factor over this period would be $\beta^{4(10)}$, which is less than 0.01. Even when retirement is only 10 years away, the worker will barely care, since they discount the future so heavily.

TABLE 9.

Breusch-Godfrey test. The model predicts that there is no autocorrelation.

Specification	**NSP**		**SP**	
$\beta = 0.97,\ \sigma = 1$	χ^2	P-value	χ^2	P-value
Household, 1 lag	9906	0.000	9945	0.000
Aggregate, 1 lag	0.402	0.526	0.006	0.936
Aggregate, 4 lags	40.732	0.000	19.883	0.001
$\beta = 0.89,\ \sigma = 2$				
Household, 1 lag	7264	0.000	7278	0.000
Aggregate, 1 lag	0.543	0.461	0.183	0.669
Aggregate, 4 lags	43.118	0.000	19.154	0.001

Nevertheless, adjusting β does not solve the puzzle. If agents have rational expectations, then there should be no autocorrelation. Table 9 shows that the null is strongly rejected for the households' forecast errors[3]. Seasonality does not explain the autocorrelation; the problem remains even when seasonal preferences are accounted for. Instead, it could be due to heterogeneity. Different households may have different discount factors and risk aversion. Imposing homogeneity may have caused the forecast errors to be consistently positive for some households and negative for others. This shows up as autocorrelation, but perhaps each household is rational and following its own preferences. Since I only have one or two observations per household, I cannot estimate each one's preferences. Instead, I checked the aggregate forecast error, $m_t^{ag}g$.

$$m_t^{agg} = \sum_i m_{it} \tag{25}$$

[3]Wilson (2020) reached the opposite conclusion in his paper on the equity premium puzzle. However, the results can be reconciled; autocorrelation in $E_t\left(\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma} R_{t+1}^{bonds}\right) - \beta^{-1}$ and $E_t\left(\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma} R_{t+1}^{stocks}\right) - \beta^{-1}$ may cancel out when the Euler equations are subtracted from each other; Wilson (2020) tested $E_t\left(\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma} (R_{t+1}^{stocks} - R_{t+1}^{bonds})\right)$.

In aggregate, household heterogeneity should cancel out and the autocorrelation should vanish. This would justify the use of representative agents. If one lag is used, then there is no evidence of autocorrelation in the aggregate forecast error. However, with four lags (i.e. one year), the null is rejected in all specifications. This is inconsistent with rational expectations.

3.3. Robustness

If I change the criteria for excluding outliers, my main results still hold. Brav et al. (2002) write, “First, we delete from the sample households with consumption reported in fewer than three consecutive quarters. Second, we delete the consumption growth $\frac{c_{it}}{c_{i,t-1}}$ if $\frac{c_{it}}{c_{i,t-1}} < 1/2$ *and* $\frac{c_{i,t+1}}{c_{it}} > 2$ [emphasis in the original]. Third, we delete the consumption growth $\frac{c_{it}}{c_{i,t-1}}$ if it is greater than five.” They also drop rural households. In the beginning, the CEX only surveyed urban households; Brav et al. (2002) drop rural households in the rest of the sample in order to be consistent. Constantinides and Ghosh (2017) use the same criteria. Table 10 shows what happens when I use those rules in my sample. In all cases, the p-values were less than 0.001.

TABLE 10.

The Risk-Free Rate Puzzle, Brav et al. (2002) sample. $N = 219,099$

	Normal		**Double Pareto**		**Generalized Pareto**	
	NSP	SP	NSP	SP	NSP	SP
Coeff. of Relative Risk Aversion (σ)	Mean m	Mean m	Mean m	Mean m	Mean m	Mean m
1	0.0189	0.0186	0.0194	0.0192	0.0222	0.0219
	(0.0022)	(0.0016)	(0.0022)	(0.0016)	(0.0022)	(0.0016)
2	0.1351	0.1343	0.1378	0.1370	0.1537	0.1530
	(0.0062)	(0.0067)	(0.0063)	(0.0067)	(0.0064)	(0.0071)
3	0.4771	0.5121	0.4854	0.5213	0.5357	0.5774
	(0.0513)	(0.0740)	(0.0538)	(0.0747)	(0.0557)	(0.0791)

The mean is still positive and significant under all specifications. A slight change in β could fix this, but then there would still be strong evidence of autocorrelation, as shown in Table 11.

Kocherlakota and Pistaferri (2009) used a different set of criteria. They drop households with zero food consumption. Households that skip an interview and then reenter the sample are excluded. Students living in dorms are also dropped along with the five biggest and five smallest observations. There is more noise in this sample, but that is not necessarily bad. Tighter outlier criteria reduce measurement error but at the risk of excluding valid observations. Table 12 shows that the puzzle deepens in this sample. The

TABLE 11.

Breusch-Godfrey test, Brav et al. (2002) sample.

Specification	**NSP**		**SP**	
$\sigma = 1$	χ^2	P-value	χ^2	P-value
Household, 1 lag	5729	0.000	5746	0.000
Aggregate, 1 lag	0.615	0.433	0.017	0.897
Aggregate, 4 lags	38.129	0.000	15.112	0.005
$\sigma = 2$				
Household, 1 lag	86.92	0.000	84.188	0.000
Aggregate, 1 lag	0.075	0.785	0.032	0.857
Aggregate, 4 lags	29.497	0.000	38.047	0.000

p-values were always below 0.001. People were already buying too few bonds if my outlier criteria or Brav et al.'s (2002) are used. If the sample is more volatile, bonds become even more appealing to smooth consumption.

TABLE 12.

The Risk-Free Rate Puzzle, Kocherlakota and Pistaferri (2009) sample. $N = 218,789$

	Normal		**Double Pareto**		**Generalized Pareto**	
	NSP	SP	NSP	SP	NSP	SP
Coeff. of Relative Risk Aversion (σ)	Mean m	Mean m	Mean m	Mean m	Mean m	Mean m
1	0.0328 (0.0025)	0.0322 (0.0019)	0.0334 (0.0025)	0.0328 (0.0019)	0.0362 (0.0025)	0.0356 (0.0019)
2	0.1992 (0.0081)	0.1993 (0.0088)	0.2021 (0.0082)	0.2022 (0.0089)	0.2189 (0.0084)	0.2192 (0.0092)
3	0.8361 (0.0807)	0.8509 (0.0830)	0.8464 (0.0847)	0.8613 (0.0835)	0.9088 (0.0877)	0.9246 (0.0864)

If σ is one or two, the point estimates are close to zero, though they remain significantly positive. As usual, tweaking β can solve this issue but not the autocorrelation problem. Table 13 summarizes.

The autocorrelation is weaker in this sample because there is more noise. Raising the noise to a higher power weakens the autocorrelation even further. Nevertheless, the null is still rejected. Perhaps this could be fixed by continuing to loosen the outlier criteria, introducing more noise. That would push the sample mean further from zero, but then the discount factor could be lowered appropriately. However, a successful result would not be robust. Additionally, the measurement error probably biases us against finding autocorrelation. Diluting the autocorrelation by increasing the measurement error is not a satisfactory solution to the puzzle.

TABLE 13.
Breusch-Godfrey test, Kocherlakota and Pistaferri (2009) sample. The model predicts that there is no autocorrelation.

Specification	**NSP**		**SP**	
$\sigma = 1$	χ^2	P-value	χ^2	P-value
Household, 1 lag	8613	0.000	8616	0.000
Aggregate, 1 lag	0.809	0.369	0.000	0.985
Aggregate, 4 lags	35.393	0.000	18.217	0.001
$\sigma = 2$				
Household, 1 lag	84.67	0.000	82.397	0.000
Aggregate, 1 lag	0.192	0.661	0.292	0.589
Aggregate, 4 lags	13.026	0.011	19.933	0.001

In the equity premium literature, many papers study limited asset market participation (e.g. Mankiw and Zeldes 1991, Cogley 2002, Kocherlakota and Pistaferri 2009, Balduzzi and Yao 2007). Lots of households do not own stocks or bonds, so the Euler equation might not be binding for them. In theory, negative stock holdings can be interpreted as shorting, but most people do not know how to do that. Thus, it makes sense to exclude households that do not own stocks. However, the risk-free rate puzzle is about bonds. Negative bond holdings are considered loans. Perhaps some households face borrowing constraints; this prevents them from satisfying their Euler equation. I tried restricting the sample to bondholders. However, the CEX does not track this very well. It asks if households own securities, but it does not make them specify whether these assets are all stocks, all bonds, or a mixture. Table 14 is based on households that report a positive amount of securities, but not all of them are necessarily bondholders. The sample shrinks dramatically; Cogley (2002) argues that the CEX underreports the number of stock- and bondholders. However, this should not bias the results.

The results are very similar to the baseline shown in Table 6. The mean of m is always positive and significant, so the Euler equation is rejected. Reducing β can fix that problem. In Table 15, I show that there is strong autocorrelation at the household level. However, the results change when we aggregate and account for seasonal preferences. There is no evidence of autocorrelation if $\sigma = 2$, even when four lags are used. This shows that the autocorrelation observed previously was driven by households that did not hold securities.

However, the puzzle remains unsolved. In Table 16, we see that the quarterly discount factor has to be reduced to 0.89 in order to not reject the Euler equation. This implausibly low and has unrealistic implications, which were discussed in Section 3.2.

TABLE 14.
The Risk-Free Rate Puzzle, stock- and bondholders only. The model predicts that the mean of m is zero. $N = 34,683$

	Normal		**Double Pareto**		**Generalized Pareto**	
	NSP	SP	NSP	SP	NSP	SP
Coeff. of Relative Risk Aversion (σ)	Mean m	Mean m	Mean m	Mean m	Mean m	Mean m
1	0.0200 (0.0035)	0.0188 (0.0020)	0.0206 (0.0036)	0.0194 (0.0020)	0.0234 (0.0036)	0.0222 (0.0020)
2	0.1162 (0.0079)	0.1125 (0.0044)	0.1189 (0.0081)	0.1152 (0.0044)	0.1346 (0.0083)	0.1308 (0.0044)
3	0.3011 (0.0144)	0.2931 (0.0083)	0.3085 (0.0152)	0.3004 (0.0083)	0.3528 (0.0158)	0.3445 (0.0086)

TABLE 15.
Breusch-Godfrey test, stock- and bondholders only. The model predicts that there is no autocorrelation. $N = 34,683$

Specification	**NSP**		**SP**	
$\sigma = 1$	χ^2	P-value	χ^2	P-value
Household, 1 lag	1943	0.000	1953	0.000
Aggregate, 1 lag	1.881	0.170	1.924	0.165
Aggregate, 4 lags	47.129	0.000	9.943	0.041
$\sigma = 2$				
Household, 1 lag	1428	0.000	1424	0.000
Aggregate, 1 lag	2.480	0.115	2.086	0.149
Aggregate, 4 lags	45.615	0.000	8.232	0.083

TABLE 16.
The Risk-Free Rate Puzzle, stock- and bondholders only with seasonal preferences. The model predicts that the mean of m is zero. $N = 34,683$

	Normal		**Double Pareto**		**Generalized Pareto**	
Parameters	Mean m	P-value	Mean m	P-value	Mean m	P-value
$\beta = 0.88$, $\sigma = 2$	-0.0137 (0.0043)	0.002	-0.0111 (0.0043)	0.011	0.0045 (0.0044)	0.301
$\beta = 0.89$, $\sigma = 2$	-0.0010 (0.0043)	0.821	0.0017 (0.0043)	0.701	0.0173 (0.0044)	0.000
$\beta = 0.90$, $\sigma = 2$	0.0115 (0.0043)	0.009	0.0141 (0.0043)	0.001	0.0298 (0.0044)	0.000

4. CONCLUSION

The standard macro framework makes two claims. First, the marginal utility of consuming today should equal the expected marginal utility of consuming next period. This is how people maximize their utility. Second, there should be no autocorrelation in the forecast errors. This is implied by the rational expectations hypothesis. The data does not support these twin claims. This conclusion is not sensitive to whether macro or micro data is used, which outliers are excluded, whether the sample is restricted to bondholders, or if seasonal preferences are incorporated.

Thus, there is an asset-pricing puzzle. However, the puzzle is different in the micro data. Household consumption is more volatile, so bonds are desirable in order to smooth consumption. Bonds are so attractive in this dataset that people should be buying more of them. In the macro data, consumption was already smooth, so it appeared that people were buying too many bonds. Many papers tried to resolve this issue, but a new approach is needed since now the puzzle is reversed.

The equity premium puzzle is closely related. In Brav et al. (2002), the coefficient of relative risk aversion had to be at least two in order to solve the puzzle. Balduzzi and Yao (2007) achieved the same result in a few of their specifications, but in most of their paper, risk aversion had to be higher. In Wilson (2020), relative risk aversion had to be at least 2.7. However, these results are difficult to reconcile with the risk-free rate puzzle. I found that when risk aversion was higher, the discount factor had to be lowered. When relative risk aversion is equal to one, a modest reduction in β suffices. However, for the higher risk aversion in these equity premium papers, a much more dramatic change would be required.

We should be willing to relax the assumption of $\beta = 0.99$. It does not find support here or in laboratory experiments. In his famous paper "Theory Ahead of Business Cycle Measurement", Prescott (1986) wrote, "What sort of science would economics be if micro studies used one [leisure] share parameter and aggregate studies another?" Currently, macro studies use one discount factor and experimental studies use another. This inconsistency needs to be resolved by lowering β. However, in order to address the risk-free rate puzzle, I found that β had to be implausibly small. For discussion of behavioral economics and asset-pricing puzzles, see Semenov (2009).

We could drop the rational expectations hypothesis. In many specifications, adjusting the discount factor appeared to work, but then the assumption of no autocorrelation was rejected. There is a large body of research on adaptive learning and other alternatives to rational expectations. However, I am unsure if macroeconomists would consider that to be a solution. The no-autocorrelation condition is widely seen as a strength —

not a weakness — of rational expectations. It seems unrealistic that people would make consistent overestimates or underestimates. Macroeconomists are increasingly willing to consider alternatives to rational expectations, but not because they believe there is autocorrelation in the forecast errors.

Seasonality may be part of the solution. Miron (1986) presents evidence that seasonal effects are larger for some goods than for others. Perhaps a new utility function could give different weights to different categories of goods in each season. Here is an illustration.

$$U = \sum_{t=0}^{\infty} \sum_{s=1}^{4} \beta^{4t+s} \left(\omega_s \frac{\left(\sum_{k=1}^{K} \gamma_{sk} c_{itsk}\right)^{1-\sigma}}{1-\sigma} \right) \tag{26}$$

There are K categories of goods. The weight for good k in season s is γ_{sk}. The overall seasonal weight is ω_s. I do not expect this generalized seasonal CRRA function to become mainstream. If it solves the puzzle, the usual CRRA utility could be justified since it is a special case; the extra features that solve the puzzle are probably unnecessary for other macro projects. Non-CRRA utility has been tried (e.g. Epstein-Zin preferences), but the motive was to explain the macro puzzle, which is reversed in the micro data.

The micro data presents a new version of an old puzzle. This paper does not have a solution, but the suggestions above may lead us to an answer.

REFERENCES

Balduzzi, Pierluigi, and Tong Yao, 2007. Testing Heterogeneous-Agent Models: An Alternative Aggregation Approach. *Journal of Monetary Economics* **54(2)**, 369-412.

Bansal, Ravi, and Wilbur John Coleman, 1996. A Monetary Explanation of the Equity Premium, Term Premium, and Risk-Free Rate Puzzles. *Journal of Political Economy* **104(6)**, 1135-1171.

Barr, David, Parantap Basu, and Kenji Wada, 2012. Uninsurable Risk and the Determination of Real Interest Rates: An Investigation Using UK Indexed Bonds. Working paper.

Basu, Parantap, Andrei Semenov, and Kenji Wada, 2011. Uninsurable Risk and Financial Market Puzzles. *Journal of International Money and Finance* **30(6)**, 1055-1089.

BLS (Bureau of Labor Statistics), 1997. Consumer Expenditure Interview Survey, 1996: Interview Survey and Detailed Expenditure Files. Washington, DC: United States Department of Labor [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1997. https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/20/studies/2794?archive=ICPSR&sortBy=7&paging.startRow=26

BLS (Bureau of Labor Statistics), 2013. Consumer Expenditure Interview Survey Public Use Microdata: 2012 Users' Documentation. Washington, DC: United States Department of Labor.

BLS (Bureau of Labor Statistics), 2014. Consumer Expenditure Interview Survey Public Use Microdata: 2013 Users' Documentation. Washington, DC: United States Department of Labor.

BLS (Bureau of Labor Statistics), 2014. Consumer Expenditure Survey: Frequently Asked Questions. Washington, DC: United States Department of Labor. http://www.bls.gov/cex/csxfaqs.htm#q23.

Boldrin, Michele, Lawrence J. Christiano, and Jonas DM Fisher, 2001. Habit Persistence, Asset Returns, and the Business Cycle. *American Economic Review* **91(1)**, 149-166.

Branch, E. Raphael, 1994. The Consumer Expenditure Survey: A Comparative Analysis. Monthly Labor Review: December 1994.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy, 2002. Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence. *Journal of Political Economy* **110(4)**, 793-824.

Cogley, Timothy, 2002. Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey. *Journal of Monetary Economics* **49(2)**, 309-334.

Constantinides, George M., 1990. Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy* **98(3)**, 519-543.

Constantinides, George M., and Anisha Ghosh, 2017. Asset Pricing with Countercyclical Household Consumption Risk. *The Journal of Finance* **72(1)**, 415-460.

Epstein, Larry G., and Stanley E. Zin, 1990. 'First-order' Risk Aversion and the Equity Premium Puzzle. *Journal of Monetary Economics* **26(3)**, 387-407.

Epstein, Larry G., and Stanley E. Zin, 1991. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. *Journal of Political Economy*, 263-286.

Ferson, Wayne E., and Campbell R. Harvey, 1992. Seasonality and Consumption?Based Asset Pricing. *The Journal of Finance* **47(2)**, 511-552.

Gabaix, Xavier, 2012. Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance. *The Quarterly Journal of Economics* **127(2)**, 645-700.

Geisen, Emily, Ashley Richards, Charles Strohm, and Joan Wang, 2011. US Consumer Expenditure Records Study. US Census Bureau.

Grossman, Sanford J., Angelo Melino, and Robert J. Shiller, 1987. Estimating the Continuous-Time Consumption-Based Asset-pricing Model. *Journal of Business & Economic Statistics* **5(3)**, 315-327.

Huggett, Mark, 1993. The Risk-Free rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics and Control* **17(5-6)**, 953-969.

Jacobs, Kris, 1999. Incomplete Markets and Security Prices: Do Asset-Pricing Puzzles Result from Aggregation Problems? *The Journal of Finance* **54(1)**, 123-163.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi, 2015. Seasonal Variation in Treasury Returns. *Critical Finance Review* **4(1)**, 45-115.

Kamstra, Mark J., Lisa A. Kramer, Maurice D. Levi, and Russ Wermers, 2017. Seasonal Asset Allocation: Evidence from Mutual Fund Flows. *Journal of Financial and Quantitative Analysis* **52(1)**, 71-109.

Kocherlakota, Narayana, and Luigi Pistaferri, 2009. Asset Pricing Implications of Pareto Optimality with Private Information. *Journal of Political Economy* **117(3)**, 555-590.

Kollmann, Robert, 2009. Household Heterogeneity and the Real Exchange Rate: Still a Puzzle. CEPR Discussion Paper No. DP7301. Available at SSRN: https://ssrn.com/abstract=1433896

Mankiw, N. Gregory, and Stephen P. Zeldes, 1991. The Consumption of Stockholders and Nonstockholders. *Journal of Financial Economics* **29(1)**, 97-112.

Miron, Jeffrey A., 1986. Seasonal Fluctuations and the Life Cycle-Permanent Income Model of Consumption. *Journal of Political Economy* **94(6)**, 1258-1279.

Semenov, Andrei, 2009. Departures from Rational Expectations and Asset Pricing Anomalies. *The Journal of Behavioral Finance* **10(4)**, 234-241.

Telmer, Chris I., 1993. Asset-Pricing Puzzles and Incomplete Markets. *The Journal of Finance* **48(5)**, 1803-1832.

Toda, Alexis Akira, and Kieran Walsh, 2015. The Double Power Law in Consumption and Implications for Testing Euler Equations. *Journal of Political Economy* **123(5)**, 1177-1200.

Toda, Alexis Akira, and Kieran James Walsh, 2017, Fat Tails and Spurious Estimation of Consumption-Based Asset Pricing Models. Journal of Applied Econometrics 32, no. 6: 1156-1177.

Vissing-Jørgensen, Annette, 2002. Limited Asset Market Participation and the Elasticity of Intertemporal Substitution. *Journal of Political Economy* **110(4)**, 825-853.

Weil, Philippe, 1989. The Equity Premium Puzzle and the Risk-free Rate Puzzle. *Journal of Monetary Economics* **24**, 401-421.

Wilson, Matthew S., 2020. Disaggregation and the Equity Premium Puzzle. *Journal of Empirical Finance* **Vol.58**, 1-18.