Publication Bias in Measuring Intertemporal Substitution*

Tomas Havranek

Czech National Bank and Charles University, Prague

tomas.havranek@ies-prague.org

November 15, 2013

Abstract

I examine 2,735 estimates of the elasticity of intertemporal substitution in consumption (EIS) reported in 169 published studies. The literature shows strong publication bias: researchers report negative and insignificant estimates less often than they should, which pulls the mean estimate up by about 0.5. When I correct the mean for the bias, for macro estimates I get zero, even though the reported t-statistics are on average two. The corrected mean of micro estimates for asset holders is around 0.3–0.4. Calibrations of the EIS greater than 0.8 are inconsistent with the bulk of the empirical evidence.

Keywords: Elasticity of intertemporal substitution, consumption, publication bias, meta-analysis

JEL Codes: E21, C83

1 Introduction

The elasticity of intertemporal substitution in consumption, a key input into macroeconomic models, has been estimated by hundreds of researchers. I argue that the findings of the literature are biased upwards because of the tendency to preferentially select positive and significant estimates for publication. The publication bias is so strong that the literature could consistently produce statistically significant estimates even if the underlying EIS was zero.

*An online appendix with data, code, and additional results is available at meta-analysis.cz/eis.

I exploit an assumption that researchers make when estimating the EIS: the regression coefficient has a t-distribution. This assumption implies that the reported estimates should not be correlated with their standard errors. But when I regress the estimates on their standard errors I get a coefficient of about two—even if I control for 30 variables reflecting the context in which researchers obtain the estimates. The finding indicates that the reported t-statistic tends to equal two no matter how large the underlying elasticity is. The constant in this regression is zero, which suggests that the mean underlying elasticity beyond publication bias is negligible. Therefore the mean EIS reported in the literature, 0.5, is entirely due to publication bias. Researchers may put as much as 75% of all negative estimates into their file drawers.

A quarter of the studies use micro data to estimate the elasticity. The micro studies report a positive elasticity even after correction for publication bias: on average about 0.2. The corrected mean estimate reaches 0.3–0.4 for asset holders, which I consider the literature's best shot for the calibration of the EIS. Vissing-Jorgensen (2002) argues that including non-asset holders creates a downward bias in the estimated elasticity, because the corresponding Euler equation is not necessarily valid for households not participating in asset markets. My results suggest that the empirical literature on the EIS does not support calibrations greater than 0.8, the largest upper bound I get for the asset holders' elasticity.

Publication bias is often unintentional and at the level of individual studies can even be beneficial. Suppose, for example, that a researcher estimates a negative elasticity of intertemporal substitution. A negative EIS implies convex utility, so the estimate is probably a statistical artifact. One should get negative estimates from time to time when the underlying EIS is positive but small, yet it would be a mistake to build conclusions on them. The problem is that no upper limit exists which would mirror the lower limit of zero given by the theory. If many researchers discard negative estimates but most report large positive ones, our inference from the literature gets biased.

Researchers in medicine, among others, have long been concerned with publication bias. Some of the methods I use were developed by medical scientists (such as the funnel plot; Egger *et al.*, 1997), and the best medical journals now require registration of clinical trials before publication of results (Stanley, 2005). Similarly the American Economic Association has agreed to establish a registry for randomized control trials "to counter publication bias" (Siegfried, 2012, p. 648), with the eventual intention to make registration necessary for submission to the Association's journals. It appears infeasible to impose this requirement in other fields of empirical economics, although there is little reason to believe they are free of the bias. As an empirical, non-experimental researcher I have to choose among many methods, and if the result is not to my liking, I can try a different specification. With noisy data I might be tempted to search for a specification that produces a large estimate, which offsets large standard errors and delivers statistical significance. Publication bias in economics has been mentioned by, for example, DeLong & Lang (1992), Card & Krueger (1995), Stanley (2001), and Ashenfelter & Greenstone (2004).[1]

Because the empirical literature yields little consensus concerning the appropriate value of the EIS, calibrations routinely differ by an order of magnitude; as Table 1 documents, the typical range of calibrations lies between 0.2 and 2. I find it unsatisfactory that the empirical research of the last three decades, the work of hundreds of economists, has not been systematically synthesized to provide clearer guidance for the calibration of this crucial parameter. One of the first surveys on the micro evidence from consumption Euler equations, Browning & Lusardi (1996), puts it in the following way:

> It is frustrating in the extreme that we have very little idea of what gives rise to the different findings. (...) We still await a study which traces all of the sources of differences in conclusions to sample period; sample selection; functional form; variable definition; demographic controls; econometric technique; stochastic specification; instrument definition; etc. (p. 1833)

In this paper I try to assign a pattern to the differences in the reported estimates of the EIS and compute the mean corrected for potential publication and misspecification biases. The publication bias seems to dwarf the effects of alleged misspecifications.

[1]Most of the studies mentioning publication bias are meta-analyses, as is this paper. Stanley & Jarrell (1989) introduce the framework of meta-analysis in economics. Applications include Smith & Huang (1995) on the willingness to pay for air quality, Ashenfelter *et al.* (1999) on the return to education, Görg & Strobl (2001) on the impact of foreign investment on domestic productivity, Disdier & Head (2008) on the effect of distance on trade, Card *et al.* (2010) on the impact of active labor market policy, Chetty *et al.* (2011) on labor supply elasticities, Havranek & Irsova (2011) on vertical spillovers from foreign investment, and Rusnak *et al.* (2013) on the transmission of monetary policy to prices. When registries of empirical research are missing, meta-analysis represents the only way to correct for publication bias.

Table 1: Authors calibrate the elasticity of intertemporal substitution differently

Study	EIS	Comments on the calibration
House & Shapiro (2006)	0.2	p. 1837: "Most empirical evidence indicates that the elasticity of intertemporal substitution is substantially less than one (see Hall, 1988). Our calibration is roughly the average estimate in Hall (1988), Campbell & Mankiw (1989), and Barsky *et al.* (1997)."
Piazzesi *et al.* (2007)	0.2	p. 550: "We follow Hall (1988), who estimates σ [EIS] to be around 0.2. Studies based on micro data find values for σ that are somewhat higher, but not by much. For example, Runkle (1991) reports an estimate of 0.45 using micro data on food consumption. Attanasio & Browning (1995) report estimates using CEX data between $[0.48, 0.67]$."
Chari *et al.* (2002)	0.2	p. 546: "The literature has a wide range of estimates for the curvature parameter σ [the inverse of the EIS]. We set σ to 5 and show later that this value is critical for generating the right volatility in the real exchange rate."
Trabandt & Uhlig (2011)	0.5	p. 311: "For the intertemporal elasticity of substitution, a general consensus is followed for it to be close to 0.5."
Jeanne & Ranciere (2011)	0.5	p. 920: "The benchmark risk aversion [the inverse of the EIS] and its range of variation are standard in the growth and real business cycle literature."
Jin (2012)	0.5	p. 2130: "The intertemporal elasticity of substitution is set to the standard value."
Rudebusch & Swanson (2012)	0.5	p. 121: "We set the curvature of household utility with respect to consumption, φ, to 2, implying an intertemporal elasticity of substitution in consumption of 0.5, which is consistent with estimates in the micro literature (e.g., Vissing-Jorgensen, 2002)."
Smets & Wouters (2007)	0.67	p. 593: "These [values for the EIS and other parameters] are all quite standard calibrations."
Bansal & Yaron (2004)	1.5	p. 1492: "The magnitude for the EIS that we focus on is 1.5. Hansen & Singleton (1982) and Attanasio & Weber (1989) estimate the EIS to be well in excess of 1.5. More recently, Vissing-Jorgensen (2002) and Guvenen (2006) also argue that the EIS is well over 1."
Ai (2010)	2	p. 1357: "Empirical evidence on the magnitude of the EIS parameter is mixed. While Hansen & Singleton (1982), Attanasio & Weber (1989), and Vissing-Jorgensen (2002) estimate the EIS parameter to be larger than one, other studies, for example, Hall (1988), Campbell (1999), and Browning *et al.* (1999), argue that the EIS parameter is well below one. (...) Bansal *et al.* (2007) estimate the EIS parameter to be 2.43 with a standard deviation of 1.3."
Barro (2009)	2	p. 252: "Because of the shortcomings of macroeconomic estimates of the EIS, it is worthwhile to consider microeconomic evidence. The Gruber (2006) analysis is particularly attractive because it uses cross-individual differences in after-tax real interest rates that derive from arguably exogenous differences in tax rates on capital income."
Colacito & Croce (2011)	2	p. 159: "The intertemporal elasticity of substitution is equal to two, a number consistent with the literature on long-run risks. (...) Hall (1988) and many follow-up studies estimate this number to be below unity. Guvenen (2006) reproduces capital and consumption fluctuations as long as most of the wealth is held by a small fraction of the population with a high elasticity of intertemporal substitution. Attanasio & Weber (1989) document an intertemporal elasticity of substitution greater than one in the United Kingdom."

Notes: The table lists baseline calibrations of the elasticity of intertemporal substitution in selected studies. Many other authors assume EIS = 1 and use logarithmic utility.

2 Data

I search in Google Scholar for studies that estimate the EIS using consumption Euler equations. Most studies use a log-linear version of the Euler equation and regress consumption growth on the real rate of return following Hall (1988). My search query is available in the online appendix along with all data and the list of studies examined after the search. A total of 169 published papers report an estimate of the EIS and its standard error or a statistic from which the standard error can be computed. I collect all estimates from the papers and also codify 30 variables reflecting the context in which researchers obtain the elasticities (Table A1 in the Appendix). I add the last study to the data set on January 1, 2013, and terminate the search. The oldest study was published in 1981, and the ten most recent ones in 2012. The 169 studies combined provide 2,735 estimates, which makes this paper, to my knowledge, the largest meta-analysis conducted in economics. Doucouliagos & Stanley (2013) survey 87 economic meta-analyses and report that the largest one includes 1,460 estimates from 124 studies.

About 200 unpublished papers provide estimates of the EIS as well, but I focus on published studies only. I have three reasons for this restriction. First, publication status is a simple indicator of quality. Second, it would take many months to collect all information from the unpublished studies. Third, I am interested in publication bias, which may also appear in working papers (for example, when researchers use their prior beliefs concerning the correct value of the elasticity as a model selection check), but is usually associated with published studies. Rusnak *et al.* (2013), on the other hand, find little difference in the extent of publication bias between published and unpublished studies that examine the transmission of monetary policy to prices. I additionally collect estimates of the coefficient of relative risk aversion if the coefficient also determines the EIS—in the typical isoelastic utility function, the same parameter determines both risk aversion and the inverse of the EIS. In this case I approximate the standard error of the EIS by the delta method.

The mean reported estimate from all studies is 0.5. For the computation I exclude estimates that are larger than 10 in absolute value because they would influence the unweighted average heavily. (Later in the analysis I use precision as the weight, and these large estimates are usually imprecise, so I leave them in the data set.) The mean estimate reported in the literature thus corresponds to the consensus value referred to by Trabandt & Uhlig (2011), Jeanne & Ranciere

(2011), Jin (2012), and Rudebusch & Swanson (2012). But the arithmetic mean is driven by studies reporting many estimates. As a next step I select the median estimate from each study: the mean of medians is even larger than the mean of all estimates and reaches 0.7. For micro studies (42 out of the 169 studies in the data set) I get a mean EIS of 0.8. The data set also includes 33 studies published in the top five general interest journals; these studies report the EIS to be 0.9 on average.

If I stopped here I would argue that the empirical evidence of the last three decades, when more weight is given to micro studies and the best journals, is consistent with calibration of the EIS close to one. Logarithmic utility would seem to be a good approximation of the isoelastic utility function. This conclusion could be a mistake, though, since not all estimates have the same probability of being published. If researchers intentionally or unintentionally suppress negative or insignificant estimates, the mean reported elasticity gets biased upwards.

3 Publication Bias

Researchers estimating the EIS assume that the regression parameter is approximately normally distributed, and the ratios of the regression estimates of the EIS to their standard errors are assumed to have a t-distribution. This implies that the numerator should be statistically independent from the denominator and the reported estimates are uncorrelated with their standard errors (Card & Krueger, 1995):

$$EIS_{ij} = EIS_0 + \beta \cdot SE(EIS_{ij}) + u_{ij}, \tag{1}$$

where EIS_{ij} and $SE(EIS_{ij})$ are the i-th estimates of the elasticity of intertemporal substitution and their standard errors reported in the j-th studies; u_{ij} is a normal disturbance term. The coefficient β should be zero in the absence of publication bias. The bias has two potential sources. First, researchers may discard negative estimates, which are inconsistent with the theory since they imply a convex utility function. In this case I would obtain a positive estimate of β because of the heteroskedasticity of (1): With low standard errors the estimates lie close to the mean underlying elasticity. As the standard errors increase, the estimates get more dispersed, and some get large. If researchers discard the negative estimates but keep the large positive ones, a positive correlation between EIS_{ij} and $SE(EIS_{ij})$ arises. Second, researchers (or editors or referees) may prefer statistically significant estimates. In that case researchers need large estimates of the EIS to offset the standard errors, and again I obtain a positive estimate of β. If the underlying elasticity is zero but all researchers desire a positive estimate significant at the 5% level, they need t-statistics of about two, so the estimated β will be close to two.

The constant in regression (1), EIS_0, denotes the underlying effect corrected for publication bias: the mean EIS conditional on standard errors approaching zero. I have noted that (1) is heteroskedastic, and the degree of heteroskedasticity is determined by the estimates' standard errors. To achieve efficiency I use weighted least squares with the inverse of the standard error, the estimates' precision, as the weight (Stanley, 2008). In all regressions that include multiple estimates from one study I cluster standard errors at the study level. I prefer to estimate the equation with study fixed effects to remove the influence of the studies' characteristics.

The first column of Table 2 reports the baseline result. The estimated β is approximately two and the constant equals zero, suggesting strong publication bias and zero underlying elasticity on average. Because 80% of the reported estimates are positive, as much as 1,641 (60% of 2,735) negative estimates may be missing in the literature because of publication selection. This result suggests that researchers report only a quarter of all negative estimates. Moreover, half of the positive estimates have a t-statistic above two, which would indicate that researchers discard 90% of estimates if we accepted that the underlying elasticity was zero for all studies.

Table 2: The reported estimates are correlated with their standard errors

	FE	BE	Median	IV	Micro	Top	Country
SE	2.115***	3.020***	2.719***	1.659*	1.496**	1.466*	2.117***
	(0.205)	(0.573)	(0.397)	(0.850)	(0.717)	(0.825)	(0.216)
Constant	0.0145	0.0303***	0.0322***	0.0340	0.174***	0.171*	0.0144
	(0.00881)	(0.00656)	(0.00893)	(0.0363)	(0.0554)	(0.0887)	(0.00928)
Observations	2,735	2,735	2,735	2,735	512	566	2,735
Studies	169	169	169	169	42	33	169

Notes: The table presents the results of regression $EIS_{ij} = EIS_0 + \beta \cdot SE(EIS_{ij}) + u_{ij}$. EIS_{ij} and $SE(EIS_{ij})$ are the i-th estimates of the elasticity of intertemporal substitution and their standard errors reported in the j-th studies. Estimated by weighted least squares with the inverse of the reported estimate's standard error taken as the weight. Standard errors of regression parameters are clustered at the study level and shown in parentheses. FE = study fixed effects. BE = between effects. Median = only median estimates of the EIS reported in the studies are included. IV = the number of observations is used as an instrument for the standard error. Micro = only micro estimates of the EIS are included. Top = only estimates of the EIS from the top 5 journals are included. Country = country and study fixed effects. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

In the second and third column of the table I use the average and median values reported in the studies. The evidence for publication bias gets stronger when between-study instead of within-study variation is used. The estimated magnitude of the bias increases to 2.7–3, and both regressions identify a significant but small underlying EIS of about 0.03. In the fourth column of the table I use the number of observations as an instrument for the reported standard error. Because the methods employed in the studies probably influence the reported elaticities and can also influence the standard errors, the explanatory variable in (1) may be endogenous. Larger studies report more precise estimates, and I believe it is reasonable to assume that the number of observations is little correlated with the choice of methodology (but I control for method choices in the next section). The instrumental variable estimation with study fixed effects yields a smaller estimate for publication bias, around 1.7. The estimate is marginally insignificant at the 5% level (p-value = 0.051), but one can achieve significance by excluding outlying values of the EIS larger than 10 in absolute value.

Publication bias seems to be smaller in micro studies and studies published in the top five journals, but not by much—by approximately 25% compared to all studies. Moreover, micro studies and studies published in top journals show a positive EIS, about 0.2, even after correction for publication bias. In the last column of the table I add country fixed effects to the baseline specification, because the estimates cover 120 different countries (although more than half of all the estimates use data on the US).[2] The result is similar to the case when I employ only study fixed effects: the magnitude of the publication bias is large and the mean EIS corrected for the bias is negligible.

It is difficult to say at this point which of the two potential sources of publication bias drives the results in Table 2. A graphical inspection of the data suggests that both sources play a role. Figure 1 shows the so-called funnel plot, which is often used in medical meta-analyses to detect publication bias (Egger *et al.*, 1997). The horizontal axis measures the magnitude of the estimate of the EIS, while the vertical axis measures the estimate's precision, the inverse of the standard error. The most precise estimates should be concentrated close to the underlying effect at the top of the figure, while the imprecise estimates at the bottom should be more dispersed. The normal distribution of the estimates assumed by researchers ensures that in the absence of publication bias the figure is symmetrical, forming an inverted funnel.

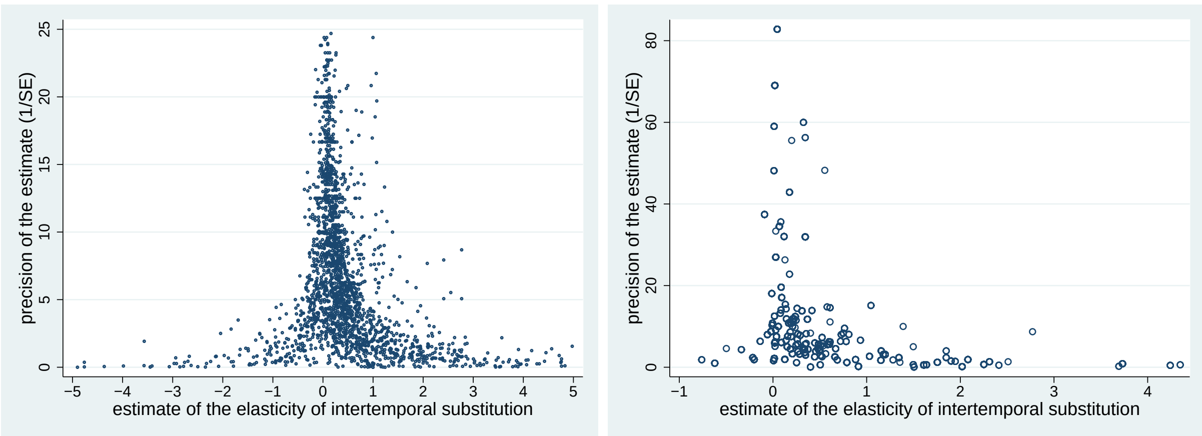
[2] In Havranek *et al.* (2013) we examine the cross-country heterogeneity in the estimates of the EIS.

Panel (a) of Figure 1 shows the funnel plot with all estimates of the EIS. The most precise estimates are positive but small. Researchers publish negative estimates less often than positive estimates with the same precision, which makes the arithmetic average of the reported estimates much larger than the precision-weighted average. It is easier to see the pattern of publication bias in panel (b) of Figure 1, where I show only the median estimates reported in the studies. In fact, equation (1) estimated in Table 2 can be interpreted as a test of the funnel's asymmetry. The weighted least squares version of equation (1) follows from rotating the axes of the funnel plot and dividing the values on the new vertical axis by the standard error to remove heteroskedasticity.

Figure 1: Negative estimates are underreported

(a) All estimates

(b) Median estimates reported in studies

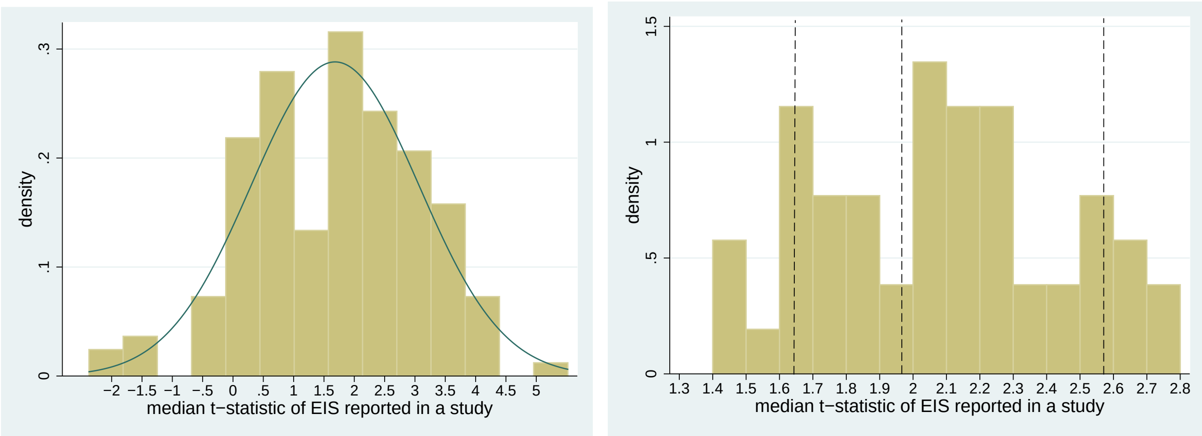


Notes: In the absence of publication bias the funnel should be symmetrical around the most precise elasticity estimates. This expectation follows from the assumptions that researchers make when estimating the elasticity. I exclude estimates with extreme magnitude or precision from the figure but include all in the regressions.

Figure 2 shows the consequences of the second source of publication bias: selection of estimates for statistical significance. The figure depicts the distribution of the median t-statistics of the estimates reported in the studies. I use median values here because some studies report many estimates with similar t-statistics, which would distort the histogram. From panel (a) we can see that estimates marginally insignificant at the 5% level are reported less often than they should be. Panel (b) offers a closer look at the distribution of the t-statistics around two. When the t-statistics reach the critical level corresponding to statistical significance at the 10%, 5%, and 1% level, estimates seem to get published more frequently.

Figure 2: Marginally insignificant estimates are underreported

(a) Histogram of median t-statistics (b) Median t-statistics around $t = 2$



Notes: In the absence of publication bias the distribution of the t-statistics should be approximately normal. The dashed lines in panel (b) denote critical values most often used for determining significance. I exclude estimates with large t-statistics from the figure but include all in the regressions.

Hedges (1992) introduces a model of the second source of publication bias. He assumes that the probability of publication of an estimate is determined by its statistical significance and only changes when the t-statistic reaches a psychological barrier. I prefer the funnel asymmetry test, because it is more flexible and captures both sources of publication bias. Nevertheless, I estimate Hedges' model and report the results in the online appendix. The results of Hedges' model also suggest strong publication bias in the literature.

4 Heterogeneity

So far I have assumed that all differences in the estimates are due to sampling error and publication bias. But in reality the estimates come from studies that use various data sets and methods, which may themselves lead to systematically different results. To explain the differences in the results I collect 30 variables reflecting the utility function used in the study, characteristics of the data, design of the analysis, definition of variables, estimation characteristics, and publication characteristics. I add these variables, described in Table A1 in the Appendix, to regression (1), which yields

$$EIS_{ij} = EIS_0 + \beta \cdot SE(EIS_{ij}) + \boldsymbol{\gamma}\mathbf{X}_{ij} + v_{ij}, \tag{2}$$

where $\mathbf{X}$ is a vector of the 30 characteristics of the estimates. The constant in the regression still denotes the underlying EIS corrected for publication bias, but now the constant must be interpreted together with $\mathbf{X}$. In this specification the constant represents the underlying EIS conditional on $\mathbf{X} = \mathbf{0}$.

My intention is to find out whether the estimated coefficient for publication bias survives the addition of variables reflecting heterogeneity. Moreover, I would like to estimate the corrected EIS reported in micro studies, and especially for asset holders, while controlling for method characteristics. Vissing-Jorgensen (2002) argues that the EIS of asset holders represents the underlying elasticity better than does the mean over all households. Guvenen (2006) shows that the dilemma between the large EIS required by most macro models and small empirical estimates can arise when the elasticity differs across groups of people: the rich (or asset holders) have a higher EIS than the rest of the population. The EIS of the asset holders determines fluctuations in investment and output, which makes the estimate more suitable for calibration—at least if the model focuses on inference concerning aggregates linked to wealth.

Table 3 presents the results of regression (2); the estimated coefficients for the control variables are reported in Table A2 in the Appendix. Some variables have the same value for all estimates reported in a study, so I do not use fixed effects as I need both between and within-study variation. I use sampling weights equal to the inverse of the number of estimates reported in a study to take into account that some studies report more estimates than others. The first column of the table includes only dummy variables for micro data and asset holders additionally to the standard error. The estimation yields a large coefficient for publication bias (2.5) and a negligible EIS beyond the bias for macro studies (0.02). The corrected elasticity for micro studies is 0.22, and for micro estimates related to asset holders the elasticity reaches 0.36 ($= 0.0237 + 0.200 + 0.136$) with a narrow 95% confidence interval $[0.33, 0.39]$.

In the next columns of the table I add groups of control variables. The estimated magnitude of the publication bias decreases from 2.5, but oscillates around two. The difference between micro and macro studies increases with more control variables. Because of the many variables in the regression it is difficult to discern the corrected underlying elasticity; it depends on the values of the control variables. I take the last column of the table and choose a preferred value for each variable to get an estimate conditional on my definition of "best practice." I put quotes

Table 3: Micro estimates are larger than macro estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SE	2.465***	1.926***	1.864***	2.109***	1.975***	1.961***	1.809***
	(0.394)	(0.251)	(0.243)	(0.268)	(0.261)	(0.262)	(0.248)
Micro data	0.200***	0.209***	0.269***	0.350***	0.476***	0.502***	0.430***
	(0.0250)	(0.0308)	(0.0495)	(0.0986)	(0.0854)	(0.0865)	(0.106)
Asset holders	0.136***	0.174***	0.195***	0.189***	0.228***	0.236***	0.316***
	(0.0303)	(0.0365)	(0.0626)	(0.0565)	(0.0482)	(0.0460)	(0.0586)
Constant	0.0237**	0.00512	-27.52	-37.61*	-26.33	-32.58	-43.89
	(0.0109)	(0.00322)	(21.42)	(21.97)	(18.44)	(22.45)	(43.05)
Utility definition		Included	Included	Included	Included	Included	Included
Data characteristics			Included	Included	Included	Included	Included
Design char.				Included	Included	Included	Included
Variable def.					Included	Included	Included
Estimation char.						Included	Included
Publication char.							Included
Observations	2,735	2,735	2,735	2,735	2,735	2,735	2,735
Studies	169	169	169	169	169	169	169

Notes: The table presents the results of regression $EIS_{ij} = EIS_0 + \beta \cdot SE(EIS_{ij}) + \boldsymbol{\gamma}\mathbf{X}_{ij} + u_{ij}$. EIS_{ij} and $SE(EIS_{ij})$ are the i-th estimates of the elasticity of intertemporal substitution and their standard errors reported in the j-th studies. $\mathbf{X}$ is a vector of the estimates' characteristics described in Table A1. Regression coefficients for control variables are reported in Table A2. Estimated with sampling weights equal to the inverse of the number of estimates reported in the j-th studies. Standard errors are clustered at the study level and shown in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

around best practice here because no study can address all potential problems in the literature simultaneously and some of my variables do not capture methodology—for example, the number of citations. We can also imagine the result as an aggregated EIS with more weight given to the estimates' characteristics that I consider in some way better than others.

Concerning the potential problems in estimation, Campbell & Mankiw (1989) illustrate why first lags of variables should not be used as instruments because of the time aggregation of consumption. Attanasio & Weber (1995) note that the use of food as a proxy for nondurable consumption can produce biased estimates if food is not separable from other consumption goods. Lawrance (1991) and Gruber (2006) argue that micro studies should include time dummies for the identification to come from cross-sectional variation and not from time series variation correlated with consumption. Ogaki & Reinhart (1998) show that assuming separability between durables and nondurables can produce a downward bias. Mulligan (2002) argues that the rate of return should be measured as the expected return on a representative unit of capital.

Attanasio & Weber (1993) note that estimating Euler equations on macro data can lead to a bias because of, for example, the omission of demographic factors. Attanasio & Low (2004) show that log-linearized Euler equations only give consistent estimates when the available time span of the data is long. Carroll (2001) is skeptical about the use of log-linearized consumption Euler equations in general, because higher-order terms may be endogenous to omitted variables. Campbell (1999) suggests that estimating the EIS with consumption growth as the dependent variable, instead of the inverse estimation with the rate of return as the dependent variable, circumvents the problem of weak instruments. Other problems have been suggested in the literature,[3] but the ones I mention here have been addressed by many empirical studies, which allows me to examine their influence on the results.

My definition of best practice is the following. I prefer if the first lag of variables is not included among instruments—which is to say I plug in value "0" for the dummy variable *First lag instrument* in column 7 of Table A2. I prefer if nondurable consumption, not food or total consumption, is used as the dependent variable; if micro studies include time dummies; if the model allows for nonseparability between durables and nondurables; if the rate of return is measured as the return on capital; if the researcher uses micro data; if the researcher estimates the exact Euler equation; if the EIS is estimated directly in a regression with consumption as the dependent variable; if the study differentiates between the EIS and the coefficient of relative risk aversion; if the regression is estimated by the general method of moments; and if the study is published in a top journal. I also plug in the maximum number of cross-sectional units used, the maximum number of years of the data period, the maximum average year of the data, the maximum number of citations of the study, and the maximum impact factor of the outlet. I set all other variables to their sample means.

The resulting estimate of the EIS for asset holders is a linear combination of regression parameters conditional on my definition of best practice. I get a point estimate of 0.33, which is close to the estimate unconditional on methodology (0.36). The estimate conditional on best practice, however, has a much wider 95% confidence interval: $[-0.2, 0.8]$. I would get a different estimate if I chose a different definition of best practice, but in general the alleged misspecifications do not seem to have a systematic effect on the estimated EIS. I believe it is safe to say that calibrations of the EIS greater than 0.8 are inconsistent with the published literature estimating the elasticity.

[3] For example, Bansal & Yaron (2004) argue that ignoring time-varying consumption volatility leads to a downward bias in the macro estimates of the EIS, but Beeler & Campbell (2009) question the extent of the bias.

The order in which I add the groups of control variables into Table 3 is arbitrary. I also derive the best-practice estimate from the last column with all control variables, even though most of them are insignificant: the last column is probably not the best possible model for explaining the heterogeneity in the estimates of the EIS. So, as a robustness check, I employ Bayesian model averaging to address model uncertainty. Bayesian model averaging runs many regressions that include the possible subsets of all explanatory variables and constructs a weighted average over these regressions. The results, available in the online appendix, are consistent with those reported here.

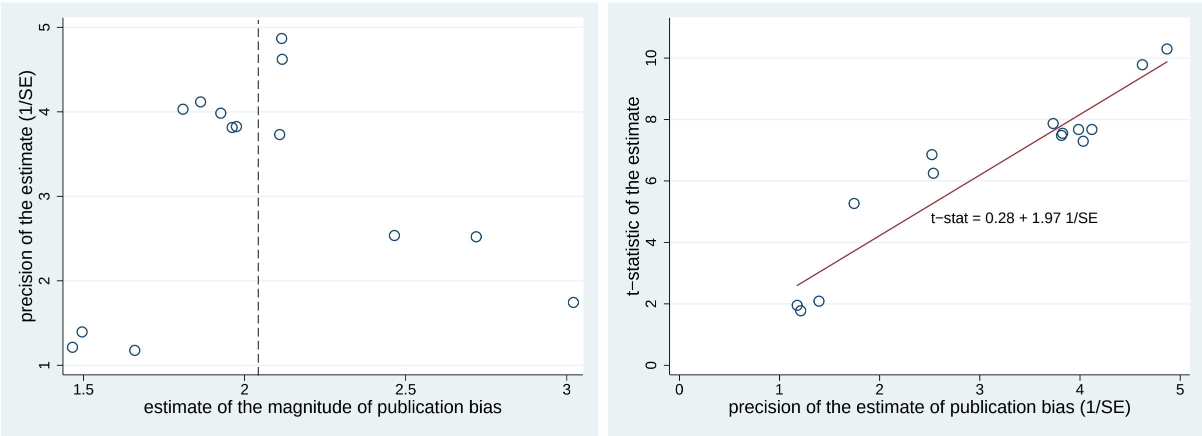
5 Conclusion

My preferred estimate of the elasticity of intertemporal substitution is 1/3. This number corresponds to the mean elasticity reported for asset holders after correction for publication bias and alleged misspecifications. The estimate represents my best guess based on the published literature. So if the studies share a misspecification that influences their results in one direction, my estimate is biased as well. I control for 30 variables reflecting the different contexts in which researchers estimate the EIS, but some other estimation aspects are idiosyncratic, and I cannot examine them meaningfully. Nevertheless, the reported estimates do not increase with the year of publication, which suggests that newer methods do not bring a substantial improvement in identification—or that the elasticity actually is small.

All empirical work is prone to publication selection because of what researchers think is a plausible value of the estimated coefficient. For example, in this paper I would be puzzled if my estimate for publication bias was negative. It would make little sense to me why negative elasticities should have a higher probability of publication. Similarly I would be disappointed with an insignificant estimate. Therefore I might have engaged in some publication selection myself, which can be tested since I report 14 different estimates of the bias. Panel (a) of Figure 3 shows the resulting funnel plot; the funnel seems to be quite symmetrical and the most precise estimates lie close to two. The test of funnel asymmetry in panel (b) corroborates the impression that the estimate of publication bias around two is unbiased. The funnel plot can thus serve as a quick check of unintentional publication selection.

Figure 3: Publication bias in the estimates of the bias?

(a) Funnel plot (b) Funnel asymmetry test



Notes: Panel (a) shows the estimates of publication bias reported in the paper. The funnel should be approximately symmetrical around the most precise estimates because of the assumptions I made when estimating the bias. The dashed line denotes the average estimate. The funnel asymmetry test shown in panel (b) follows from rotating the funnel plot and dividing the values on the new vertical axis by the standard error to correct for heteroskedasticity. The intercept should be zero and the *t*-statistic should increase with precision.

References

Ai, H. (2010): "Information Quality and Long-Run Risk: Asset Pricing Implications." *Journal of Finance* **65(4)**: pp. 1333–1367.

Ashenfelter, O. & M. Greenstone (2004): "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *American Economic Review* **94(2)**: pp. 454–460.

Ashenfelter, O., C. Harmon, & H. Oosterbeek (1999): "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias." *Labour Economics* **6(4)**: pp. 453–470.

Attanasio, O. P. & M. Browning (1995): "Consumption over the Life Cycle and over the Business Cycle." *American Economic Review* **85(5)**: pp. 1118–37.

Attanasio, O. P. & H. Low (2004): "Estimating Euler Equations." *Review of Economic Dynamics* **7(2)**: pp. 405–435.

Attanasio, O. P. & G. Weber (1989): "Intertemporal Substitution, Risk Aversion and the Euler Equation for Consumption." *Economic Journal* **99(395)**: pp. 59–73.

Attanasio, O. P. & G. Weber (1993): "Consumption Growth, the Interest Rate and Aggregation." *Review of Economic Studies* **60(3)**: pp. 631–49.

Attanasio, O. P. & G. Weber (1995): "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* **103(6)**:

pp. 1121–57.

Bansal, R., D. Kiku, & A. Yaron (2007): "Risks for the Long Run: Estimation and Inference." *Working paper*, Duke University and University of Pennsylvania.

Bansal, R. & A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance* **59(4)**: pp. 1481–1509.

Barro, R. J. (2009): "Rare Disasters, Asset Prices, and Welfare Costs." *American Economic Review* **99(1)**: pp. 243–64.

Barsky, R. B., M. S. Kimball, F. T. Juster, & M. D. Shapiro (1997): "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *The Quarterly Journal of Economics* **112(2)**: pp. 537–79.

Beeler, J. & J. Y. Campbell (2009): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment." *NBER Working Papers 14788*, National Bureau of Economic Research, Inc.

Browning, M., L. P. Hansen, & J. J. Heckman (1999): "Micro Data and General Equilibrium Models." In J. B. Taylor & M. Woodford (editors), "Handbook of Macroeconomics," volume 1 of *Handbook of Macroeconomics*, chapter 8, pp. 543–633. Elsevier.

Browning, M. & A. Lusardi (1996): "Household Saving: Micro Theories and Micro Facts." *Journal of Economic Literature* **34(4)**: pp. 1797–1855.

Campbell, J. Y. (1999): "Asset Prices, Consumption, and the Business Cycle." In J. B. Taylor & M. Woodford (editors), "Handbook of Macroeconomics," volume 1 of *Handbook of Macroeconomics*, chapter 19, pp. 1231–1303. Elsevier.

Campbell, J. Y. & N. G. Mankiw (1989): "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence." In "NBER Macroeconomics Annual 1989, Volume 4," NBER Chapters, pp. 185–246. National Bureau of Economic Research, Inc.

Card, D., J. Kluve, & A. Weber (2010): "Active Labour Market Policy Evaluations: A Meta-Analysis." *Economic Journal* **120(548)**: pp. F452–F477.

Card, D. & A. B. Krueger (1995): "Time-Series Minimum-Wage Studies: A Meta-Analysis." *American Economic Review* **85(2)**: pp. 238–43.

Carroll, C. D. (2001): "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation)." *Advances in Macroeconomics* **1**: p. 6.

Chari, V. V., P. J. Kehoe, & E. R. McGrattan (2002): "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?" *Review of Economic Studies* **69(3)**: pp. 533–63.

Chetty, R., A. Guren, D. Manoli, & A. Weber (2011): "Are Micro and Macro Labor Supply

Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *American Economic Review* **101(3)**: pp. 471–75.

Colacito, R. & M. M. Croce (2011): "Risks for the Long Run and the Real Exchange Rate." *Journal of Political Economy* **119(1)**: pp. 153–181.

DeLong, J. B. & K. Lang (1992): "Are All Economic Hypotheses False?" *Journal of Political Economy* **100(6)**: pp. 1257–72.

Disdier, A.-C. & K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics* **90(1)**: pp. 37–48.

Doucouliagos, H. & T. D. Stanley (2013): "Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

Egger, M., G. D. Smith, M. Scheider, & C. Minder (1997): "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *British Medical Journal* **316**: pp. 629–634.

Görg, H. & E. Strobl (2001): "Multinational Companies and Productivity Spillovers: A Meta-analysis." *The Economic Journal* **111(475)**: pp. F723–39.

Gruber, J. (2006): "A Tax-Based Estimate of the Elasticity of Intertemporal Substitution." *NBER Working Papers 11945*, National Bureau of Economic Research, Inc.

Guvenen, F. (2006): "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective." *Journal of Monetary Economics* **53(7)**: pp. 1451–1472.

Hall, R. E. (1988): "Intertemporal Substitution in Consumption." *Journal of Political Economy* **96(2)**: pp. 339–57.

Hansen, L. P. & K. J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica* **50(5)**: pp. 1269–86.

Havranek, T., R. Horvath, Z. Irsova, & M. Rusnak (2013): "Cross-Country Heterogeneity in Intertemporal Substitution." *William Davidson Institute Working Paper 1056*, University of Michigan.

Havranek, T. & Z. Irsova (2011): "Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is." *Journal of International Economics* **85(2)**: pp. 234–244.

Hedges, L. V. (1992): "Modeling Publication Selection Effects in Meta-Analysis." *Statistical Science* **7(2)**: pp. 246–255.

House, C. L. & M. D. Shapiro (2006): "Phased-In Tax Cuts and Economic Activity." *American Economic Review* **96(5)**: pp. 1835–1849.

Jeanne, O. & R. Ranciere (2011): "The Optimal Level of International Reserves For Emerging Market

Countries: A New Formula and Some Applications." *Economic Journal* **121(555)**: pp. 905–930.

Jin, K. (2012): "Industrial Structure and Capital Flows." *American Economic Review* **102(5)**: pp. 2111–46.

Lawrance, E. C. (1991): "Poverty and the Rate of Time Preference: Evidence from Panel Data." *Journal of Political Economy* **99(1)**: pp. 54–77.

Mulligan, C. B. (2002): "Capital, Interest, and Aggregate Intertemporal Substitution." *NBER Working Papers 9373*, National Bureau of Economic Research, Inc.

Ogaki, M. & C. M. Reinhart (1998): "Measuring Intertemporal Substitution: The Role of Durable Goods." *Journal of Political Economy* **106(5)**: pp. 1078–1098.

Piazzesi, M., M. Schneider, & S. Tuzel (2007): "Housing, Consumption and Asset Pricing." *Journal of Financial Economics* **83(3)**: pp. 531–569.

Rudebusch, G. D. & E. T. Swanson (2012): "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks." *American Economic Journal: Macroeconomics* **4(1)**: pp. 105–43.

Runkle, D. E. (1991): "Liquidity Constraints and the Permanent-Income Hypothesis: Evidence from Panel Data." *Journal of Monetary Economics* **27(1)**: pp. 73–98.

Rusnak, M., T. Havranek, & R. Horvath (2013): "How to Solve the Price Puzzle? A Meta-Analysis." *Journal of Money, Credit and Banking* **45(1)**: pp. 37–70.

Siegfried, J. J. (2012): "Minutes of the Meeting of the Executive Committee: Chicago, IL, January 5, 2012." *American Economic Review* **102(3)**: pp. 645–52.

Smets, F. & R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* **97(3)**: pp. 586–606.

Smith, V. K. & J.-C. Huang (1995): "Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models." *Journal of Political Economy* **103(1)**: pp. 209–27.

Stanley, T. D. (2001): "Wheat from Chaff: Meta-Analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

Stanley, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

Stanley, T. D. (2008): "Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection." *Oxford Bulletin of Economics and Statistics* **70(1)**: pp. 103–127.

Stanley, T. D. & S. B. Jarrell (1989): "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* **3(2)**: pp. 161–70.

Trabandt, M. & H. Uhlig (2011): "The Laffer Curve Revisited." *Journal of Monetary Economics* **58(4)**: pp. 305–327.

Vissing-Jorgensen, A. (2002): "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *Journal of Political Economy* **110(4)**: pp. 825–853.

A Data Description and Detailed Results

Table A1: Description and summary statistics of explanatory variables

Variable	Description	Mean	Std. dev.
Utility			
Epstein-Zin	=1 if the estimation differentiates between the EIS and the coefficient of relative risk aversion.	0.053	0.224
Habits	=1 if habits in consumption are assumed.	0.040	0.196
Nonsep. durables	=1 if the model allows for nonseparability between durables and nondurables.	0.041	0.199
Nonsep. public	=1 if the model allows for nonseparability between private and public consumption.	0.044	0.206
Nonsep. tradables	=1 if the model allows for nonseparability between tradables and nontradables.	0.046	0.210
Data			
No. of households	The logarithm of the number of cross-sectional units used in the estimation (households, cohorts, countries).	1.103	2.384
No. of years	The logarithm of the number of years of the data period used in the estimation.	3.184	0.570
Average year	The logarithm of the average year of the data period.	7.590	0.006
Micro data	=1 if the coefficient comes from a micro-level estimation.	0.187	0.390
Annual data	=1 if the data frequency is annual.	0.328	0.469
Monthly data	=1 if the data frequency is monthly.	0.097	0.296
Design			
Quasipanel	=1 if quasipanel (synthetic cohort) data are used.	0.053	0.224
Inverse estimation	=1 if the rate of return is the dependent variable in the estimation.	0.317	0.465
Asset holders	=1 if the estimate is related to the rich or asset holders.	0.054	0.226
First lag instrument	=1 if the first lags of variables are included among instruments.	0.305	0.460
No year dummies	=1 if year dummies are omitted in micro studies using the Panel Study of Income Dynamics.	0.030	0.171
Income	=1 if income is included in the specification.	0.241	0.428
Taste shifters	The logarithm of the number of controls for taste shifters.	0.117	0.452
Variable definition			
Total consumption	=1 if total consumption is used in the estimation.	0.203	0.402
Food	=1 if food is used as a proxy for nondurables.	0.059	0.235
Stock return	=1 if the rate of return is measured as stock return.	0.189	0.392
Capital return	=1 if the rate of return is measured as the return on capital.	0.113	0.317
Estimation			
Exact Euler	=1 if the exact Euler equation is estimated.	0.238	0.426
ML	=1 if maximum likelihood methods are used for estimation.	0.049	0.216
TSLS	=1 if two-stage least squares are used for estimation.	0.338	0.473
OLS	=1 if ordinary least squares are used for estimation.	0.104	0.306
Publication			
SE	The reported standard error of the estimate of the EIS.	136.9	3,999
Publication year	The logarithm of the year of publication of the study.	7.601	0.004
Citations	The logarithm of the number of per-year citations of the study in Google Scholar.	2.024	1.256
Top journal	=1 if the study was published in one of the top five journals in economics.	0.207	0.405
Impact	The recursive RePEc impact factor of the outlet.	1.089	1.535

Notes: Collected from published studies estimating the elasticity of intertemporal substitution. The list of studies is available in the online appendix at `meta-analysis.cz/eis`.

Table A2: Explaining the differences in the reported estimates of the EIS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SE	2.465^{***}	1.926^{***}	1.864^{***}	2.109^{***}	1.975^{***}	1.961^{***}	1.809^{***}
	(0.394)	(0.251)	(0.243)	(0.268)	(0.261)	(0.262)	(0.248)
Micro data	0.200^{***}	0.209^{***}	0.269^{***}	0.350^{***}	0.476^{***}	0.502^{***}	0.430^{***}
	(0.0250)	(0.0308)	(0.0495)	(0.0986)	(0.0854)	(0.0865)	(0.106)
Asset holders	0.136^{***}	0.174^{***}	0.195^{***}	0.189^{***}	0.228^{***}	0.236^{***}	0.316^{***}
	(0.0303)	(0.0365)	(0.0626)	(0.0565)	(0.0482)	(0.0460)	(0.0586)
Utility							
Epstein-Zin		-0.0200^{***}	-0.0261^{***}	-0.0207^{***}	-0.0257^{***}	-0.0245^{***}	-0.0754^{*}
		(0.00655)	(0.00699)	(0.00738)	(0.00742)	(0.00623)	(0.0384)
Habits		0.425^{***}	0.398^{***}	0.409^{***}	0.292^{***}	0.328^{***}	0.304^{***}
		(0.0671)	(0.0710)	(0.0362)	(0.0487)	(0.0443)	(0.0472)
Nonsep. durables		0.0320^{***}	0.0123	0.0309^{*}	0.0261^{*}	0.0265	0.0367^{*}
		(0.00324)	(0.0122)	(0.0163)	(0.0156)	(0.0160)	(0.0190)
Nonsep. public		0.0709	0.109	0.117	0.0399	0.0560	0.0959
		(0.0871)	(0.0952)	(0.0916)	(0.0964)	(0.0881)	(0.0921)
Nonsep. tradables		0.358^{***}	0.328^{***}	0.316^{***}	0.187^{***}	0.195^{***}	0.212^{***}
		(0.0456)	(0.0512)	(0.0644)	(0.0593)	(0.0679)	(0.0668)
Data							
No. of households			-0.0114	-0.0254	-0.0447^{***}	-0.0504^{***}	-0.0595^{***}
			(0.0101)	(0.0163)	(0.0142)	(0.0138)	(0.0171)
No. of years			0.00729	0.00317	0.000970	0.000292	-0.00926
			(0.00822)	(0.00639)	(0.00528)	(0.00477)	(0.0114)
Average year			3.626	4.955^{*}	3.470	4.286	6.391^{*}
			(2.823)	(2.895)	(2.430)	(2.958)	(3.568)
Annual data			-0.0260	-0.0149	-0.0207	-0.0237	-0.0142
			(0.0195)	(0.0149)	(0.0143)	(0.0149)	(0.0174)
Monthly data			-0.00511	-0.0251	0.00324	-0.0284	-0.0368
			(0.0107)	(0.0224)	(0.0188)	(0.0521)	(0.0538)
Design							
Quasipanel				-0.0932^{*}	-0.165^{***}	-0.123^{***}	-0.0886^{*}
				(0.0554)	(0.0442)	(0.0403)	(0.0490)
Inverse estimation				0.0392	0.0397	0.0513^{*}	0.0225
				(0.0240)	(0.0275)	(0.0294)	(0.0429)
First lag instrument				-0.00893	0.0133	0.0111	0.0415
				(0.0218)	(0.0204)	(0.0307)	(0.0274)
No year dummies				-0.458^{***}	-0.237	-0.240	-0.222
				(0.161)	(0.219)	(0.210)	(0.218)
Income				-0.0218	-0.0315^{**}	-0.0328	-0.0350
				(0.0172)	(0.0127)	(0.0201)	(0.0230)
Taste shifters				0.0649^{**}	0.0423^{**}	0.0375^{*}	0.0712^{**}
				(0.0251)	(0.0203)	(0.0207)	(0.0310)
Variable definition							
Total consumption					0.102^{***}	0.114^{***}	0.0888^{***}
					(0.0242)	(0.0292)	(0.0337)
Food					-0.120	-0.0827	-0.0689
					(0.187)	(0.181)	(0.193)
Stock return					-0.00760	-0.00659	-0.00444
					(0.0126)	(0.0112)	(0.00960)
Capital return					-0.0431	-0.0476	-0.0472
					(0.0327)	(0.0385)	(0.0411)
Estimation							
Exact Euler						0.0606	0.0477

Continued on the next page

Table A2: Explaining the differences in the reported estimates of the EIS (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
						(0.0384)	(0.0345)
ML						0.0874	0.130
						(0.0770)	(0.0830)
TSLS						0.0641	0.0284
						(0.0459)	(0.0553)
OLS						0.102	0.0887
						(0.0665)	(0.0710)
Publication							
Publication year							-0.609
							(6.062)
Citations							0.0255^{*}
							(0.0150)
Top journal							0.0945
							(0.0765)
Impact							-0.0185
							(0.0272)
Constant	0.0237^{**}	0.00512	-27.52	-37.61^{*}	-26.33	-32.58	-43.89
	(0.0109)	(0.00322)	(21.42)	(21.97)	(18.44)	(22.45)	(43.05)
Observations	2,735	2,735	2,735	2,735	2,735	2,735	2,735
Studies	169	169	169	169	169	169	169

Notes: The table presents the results of regression $EIS_{ij} = EIS_0 + \beta \cdot SE(EIS_{ij}) + \boldsymbol{\gamma}\mathbf{X}_{ij} + u_{ij}$. See notes to Table 3.