národní
úložiště
šedé
literatury

**Habit Formation in Consumption**

Havránek, Tomáš ; Rusnák, Marek ; Sokolova, Anna

2015

Dostupný z http://www.nusl.cz/ntk/nusl-188286

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 17.07.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní nusl.cz .

WORKING PAPER SERIES 3

2015

Tomáš Havránek, Marek Rusnák, Anna Sokolova

# Habit Formation in Consumption: A Meta-Analysis

CZECH NATIONAL BANK

# WORKING PAPER SERIES

**Habit Formation in Consumption: A Meta-Analysis**

Tomáš Havránek
Marek Rusnák
Anna Sokolova

3/2015

**CNB WORKING PAPER SERIES**

The Working Paper Series of the Czech National Bank (CNB) is intended to disseminate the results of the CNB's research projects as well as the other research activities of both the staff of the CNB and collaborating outside contributors, including invited speakers. The Series aims to present original research contributions relevant to central banks. It is refereed internationally. The referee process is managed by the CNB Research Department. The working papers are circulated to stimulate discussion. The views expressed are those of the authors and do not necessarily reflect the official views of the CNB.

Distributed by the Czech National Bank. Available at http://www.cnb.cz.

Reviewed by: Federica Teppa (De Nederlandsche Bank)
Stefan Zimmermann (University of Vienna)
Simona Malovaná (Czech National Bank)

Project Coordinator: Jan Babecký

© Czech National Bank, May 2015
Tomáš Havránek, Marek Rusnák, Anna Sokolova

# Habit Formation in Consumption: A Meta-Analysis

Tomáš Havránek, Marek Rusnák, and Anna Sokolova *


## Abstract

We examine 567 estimates of habit formation from 69 studies published in peer-reviewed journals. In contrast to previous results for most fields of empirical economics, we find no publication bias in the literature. The median estimated strength of habit formation equals 0.4, but the estimates vary widely both within and across studies. We use Bayesian model averaging to assign a pattern to this variance while taking into account model uncertainty. Studies using micro data report consistently smaller estimates than macro studies: 0.1 vs. 0.6 on average. The difference remains large when we control for 21 other study aspects, such as data frequency, geographical coverage, variable definition, estimation approach, and publication characteristics. We also find that estimates of external habit formation tend to be substantially larger than those of internal habits, that evidence for habits weakens when researchers use higher data frequencies, and that estimates differ systematically across countries.

## Abstrakt

V tomto článku analyzujeme 567 odhadů tvorby zvyků ve spotřebě zveřejněných v 69 studiích publikovaných v recenzovaných časopisech. Na rozdíl od předchozích výsledků pro většinu oborů empirické ekonomie v této literatuře nenalézáme žádné stopy publikační selektivity. Mediánový odhad síly tvorby zvyků dosahuje hodnoty 0,4, ale jednotlivé odhady se značně liší mezi studiemi i v rámci studií. K vysvětlení rozdílů mezi odhady využíváme metodu bayesovského modelového průměrování, které bere v úvahu modelovou nejistotu. Studie, které používají mikroekonomická data, prezentují konzistentně nižší odhady než studie, které se spoléhají na makroekonomická data: v průměru 0,1 oproti 0,6. Tento rozdíl zůstává velký, i když bereme v úvahu 21 dodatečných aspektů studií, jako jsou frekvence použitých dat, geografické pokrytí, definice proměnných, metoda odhadu a publikační charakteristiky. Naše výsledky dále naznačují, že odhady externí tvorby zvyků jsou obvykle výrazně vyšší než odhady interních zvyků, že odhadnutá tvorba zvyků bývá nižší na vyšších frekvencích dat a že odhady se systematicky liší mezi zeměmi.



**JEL Codes:** C83, D12, E21.
**Keywords:** Bayesian model averaging, consumption, habit formation, meta-analysis, publication bias.



*Tomáš Havránek, Economic Research Department, Czech National Bank, and Institute of Economic Studies, Charles University, Prague, tomas.havranek@cnb.cz
Marek Rusnák, Financial Stability Department, Czech National Bank, and Institute of Economic Studies, Charles University, Prague, marek.rusnak@cnb.cz
Anna Sokolova, National Research University—Higher School of Economics, Moscow, asokolova@hse.ru

Corresponding author: Tomáš Havránek. We thank Jan Babecký, Kamil Galuščák, Zuzana Iršová, Simona Malovaná, Tom Stanley, Federica Teppa, and Stefan Zimmermann for comments on earlier versions of the manuscript and acknowledge support from the Czech National Bank (project B3/13). Havránek and Rusnák acknowledge support from the Czech Science Foundation (grant #P402/12/G097). Sokolova acknowledges support from the Czech Science Foundation (grant #15-02411S). An online appendix with data and code is available at meta-analysis.cz/habits. The views expressed here are ours and not necessarily those of the Czech National Bank.

## Nontechnical Summary

Habit formation in consumption is a key component of the modern structural models used by central banks around the world to evaluate the effects of various policy measures. As shown by Fuhrer (2000), the observed inflation dynamics are consistent with a large habit formation coefficient. Furthermore, habit formation helps explain various empirical regularities: the risk-free rate puzzle (Campbell and Cochrane, 1999), the equity premium puzzle (Abel, 1990), and the happiness puzzle (Choudhary et al., 2012).

Habits in consumption can assume two forms: internal and external. Internal habit formation arises when a consumer becomes accustomed to a certain level of consumption, comparing current consumption with consumption in the previous period. In other words, the consumer's utility is no longer a function of current consumption, but one of consumption growth, with past consumption reducing present utility: more food today makes the consumer hungrier tomorrow. In contrast, external habit formation describes "keeping up with the Joneses": the consumer's utility depends on the difference between her consumption and the consumption of a reference group (such as people in the town where she lives).

Dozens of researchers have attempted to estimate the strength of habit formation, but their results vary widely and it is not clear what values should be used for the calibration of stylized models (Zimmermann, 2014). In this paper we collect the published estimates and perform a quantitative review of the literature. We find that the average reported estimate is close to 0.4, which is consistent with moderate habit formation, but does not suffice to explain some of the major puzzles in economics, such as the equity premium puzzle. Remarkably, the literature does not seem to be plagued with publication bias. Our results suggest that micro estimates of habit formation tend to be substantially smaller than macro estimates—by about 0.5.

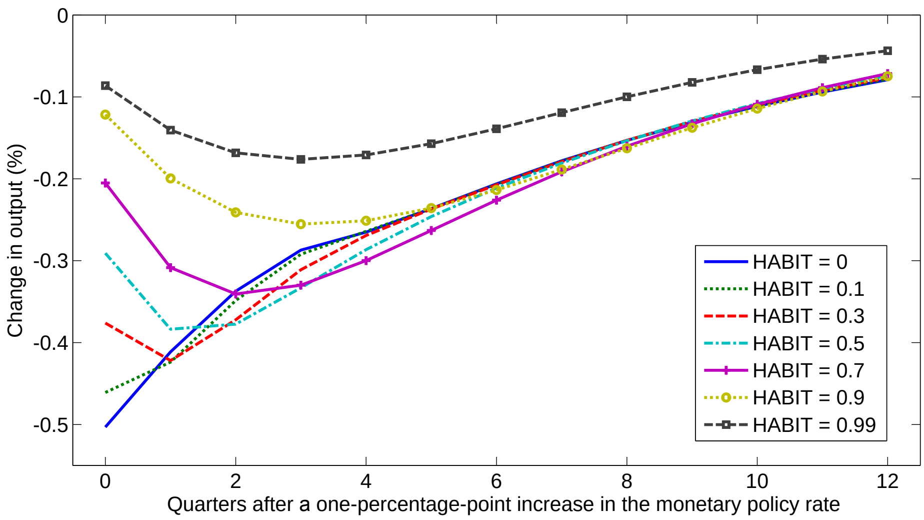
Moreover, studies examining internal habit formation report estimates that are 0.2 smaller on average than studies exploring external habits. When a researcher opts for monthly data frequency, she is more likely to find no evidence of habit formation (the difference is about 0.3), because at high frequencies more consumption goods are durable: for instance, clothing is usually durable at monthly frequency (most people do not buy new shoes every month), but not at annual frequency. We also find more evidence of habit formation in the US and EU than in Japan, a phenomenon which might be connected with cultural differences.

# 1. Introduction

The concept of habit formation in consumption is crucial for the explanation of various stylized facts in macroeconomics and finance. These stylized facts include the equity premium puzzle (Abel, 1990; Constantinides, 1990), the excess volatility of the current account (Gruber, 2004), the risk-free rate and predictability of excess returns puzzles (Campbell and Cochrane, 1999), the positive effect of growth on saving (Carroll et al., 2000), inflation dynamics (Fuhrer, 2000), and the happiness puzzle (Choudhary et al., 2012). Consequently, habit formation has become a key ingredient of the dynamic stochastic general equilibrium (DSGE) models employed by many central banks as a supportive tool for monetary-policy decisions.

The size of the parameter specifying the strength of habit formation shapes the quantitative predictions of DSGE models. Figure 1 shows how the impulse response of output to a monetary policy shock changes in the popular model by Smets and Wouters (2007) when we assume different values of habit formation: the modeled behavior of the economy within one year after the shock depends heavily on the assumed strength of habits.

***Figure 1: The Importance of Habit Formation for DSGE Models***




***Notes:*** The figure shows simulated impulse responses of GDP to a one-percentage-point increase in the monetary policy rate. We use a calibrated version of the model developed by Smets and Wouters (2007) and vary the value of the habit formation parameter while leaving all other parameters calibrated at the posterior values from Smets and Wouters (2007). For the simulations we use Matlab code from The Macroeconomic Model Data Base (Wieland et al., 2012).

Dozens of papers have estimated the habit formation parameter, but their results vary widely. The variance can be partially attributed to differences in the definition of habits: some authors assume internal habit formation (past consumption decreases present utility), while others estimate external habits ("keeping up with the Joneses"). Another important factor is the data used in the estimation—some studies analyze Euler equations for aggregate consumption (Fuhrer, 2000; Carroll et al., 2011; Everaert and Pozzi, 2014), some employ micro panel data sets (Dynan, 2000; Collado and Browning, 2007; Alessie and Teppa, 2010), and others use DSGE models (Christiano et al., 2005; Smets and Wouters, 2007), often employing prior values for the habit formation parameter. A brief look at the results of some of the seminal studies in each category suggests that the estimates are all over the place: Fuhrer (2000) asserts that habit formation is crucial for his model to fit the data and obtains estimates that lie within the range 0.8–0.9. In contrast, Dynan (2000) uses panel household data and finds no evidence of habit formation. Christiano et al. (2005) estimates the same parameter using a DSGE model and obtains a value in the range 0.5–0.7.

The lack of consensus on the value of the habit formation parameter calls for a quantitative synthesis tracing the differences in results to differences in study design. To our knowledge, this paper is the only quantitative synthesis—or meta-analysis—of habit formation. Meta-analysis is the quantitative method of research review frequently used in medical research, and has recently become used by economists as well (Stanley, 2001). In economics the method has been applied to a wide range of topics: the effect of the minimum wage on unemployment (Card and Krueger, 1995), returns from education (Ashenfelter et al., 1999), the effect of distance on trade (Disdier and Head, 2008), the intertemporal elasticity of substitution in labor supply (Chetty et al., 2011), the impact of FDI on domestic firms' productivity (Havranek and Irsova, 2011), and the effectiveness of development aid (Doucouliagos and Paldam, 2011), among others.

We attempt to gather all published estimates of habit formation, their publication characteristics, and 22 aspects related to study design, such as estimation techniques used, variable definition, data characteristics, and geographical coverage. We investigate whether these aspects systematically affect the value of the reported habit formation parameter. An obstacle to meta-analysis in economics is model uncertainty, as we do not know *a priori* which of the many potential study characteristics should be included in the baseline model. To address this problem we employ Bayesian model averaging (BMA; Raftery et al., 1997; Moral-Benito, 2015)—a method that estimates many regressions consisting of subsets of the potential explanatory variables and weights them by model fit and model complexity.

Our results show that the difference between micro estimates (think Dynan, 2000) and macro estimates (think Fuhrer, 2000) remains large even after controlling for other aspects of study design. This finding resonates with Chetty et al. (2011), who report similar divergence between micro and macro estimates in the literature estimating the intertemporal elasticity of labor supply. Our results also indicate that estimates of external habit formation are, on average, much larger than those of internal habits. In contrast, the definition of consumption used by researchers does not seem to affect their results much: studies using total non-durable consumption, food expenditures, or measures that include durable consumption come up with estimates that are roughly the same. Estimates obtained using US and EU data tend to be substantially larger than those reported for Japan and other countries. Furthermore, the frequency of the data used in the estimation matters: estimates from studies employing monthly data tend to be substantially smaller than those obtained with lower frequencies. Several additional aspects of study design, such as estimation methods, systematically affect the value of the estimates reported, while publication characteristics (the number of citations or the impact factor of the journal where the study was published) are not much correlated with the results.

The remainder of the paper is structured as follows. Section 2 describes the approach we use to collect estimates of habit formation and presents the summary statistics for our data set. Section 3 tests for publication bias in the literature. Section 4 investigates the sources of heterogeneity in the estimated habit formation parameters. Section 5 concludes. Appendix A provides the correlation matrix of the variables used, diagnostics of the Bayesian model averaging exercise, and additional robustness checks. Appendix B shows the list of studies included in our data set. An online appendix with data and code is available at `meta-analysis.cz/habits`.

## 2. The Data Set of Habit Formation Estimates

Modeling habit formation usually involves the following utility function:

$$\sum_t \beta^t u(c_{i,t} - \gamma h_{i,t}), \tag{2.1}$$

where $\beta$ is a discount factor, $u(\cdot)$ denotes the instantaneous utility function, $c_{i,t}$ is the consumption of individual $i$ in period $t$, $h_{i,t}$ is the reference habit stock, and $\gamma \in [0,1)$ captures the strength of habit formation (when $\gamma = 0$, we obtain the standard time-separable utility function). Papers that explore internal habits assume $h_{i,t} = c_{i,t-1}$: lagged consumption decreases current utility. Under internal habits, therefore, utility is determined by consumption growth, not just the level of current consumption. Papers studying external habits ("catching up with the Joneses," Abel, 1990) assume that utility is determined by the difference between the current consumption of an individual and the consumption of the corresponding reference group (for instance, the city where the consumer lives). External habits can be modeled by defining $h_{i,t} = \tilde{c}_{t-1}$, where $\tilde{c}_{t-1}$ denotes aggregate consumption in the preceding period. Instead of using consumption directly, some papers use the variable "habit stock" defined by an autoregressive process (for example, Fuhrer, 2000). Finally, a few studies model habits using a multiplicative rather than an additive specification; for example, Andrés et al. (2009) and Bjornland et al. (2011).

To obtain estimates of $\gamma$, researchers often evaluate a linear approximation of the consumption Euler equation. For example, to estimate internal habits they assess:

$$\Delta C_{i,t} = \gamma \Delta C_{i,t-1} + \sum_j \beta_j X_{j,i,t} + \varepsilon_{i,t}, \tag{2.2}$$

where $\Delta C_{i,t}$ is the change in the logarithm of consumption between periods $t$ and $t-1$ and $X_{i,t}$ represents a set of controls typically consisting of the real interest rate (to account for intertemporal substitution), income (to allow for rule-of-thumb consumers or liquidity constraints), and, for micro studies, demographic variables reflecting taste shifters (such as age, marital status, or the number of children in the household).

Several studies obtain estimates of habit formation by using household-level micro data (for example, Dynan, 2000; Guariglia, 2002; Alessie and Teppa, 2010). Micro studies can explore the heterogeneity across individuals, but often only have data covering short periods of time, and only on a fraction of consumption (such as food expenditures). Therefore, the more voluminous stream of empirical literature on habit formation makes use of aggregate consumption data, which are readily available. The macro literature is diverse, employing various data sets and approaches to estimation, as we discuss below. These papers obtain estimates of the habit formation parameter while studying issues like sticky consumption growth (Carroll et al., 2011), habit persistence in current account data (Gruber, 2004; Kano, 2009), predictability of aggregate consumption growth (Everaert and Pozzi, 2014), or inflation dynamics (Fuhrer, 2000). Many estimates of the habit formation parameter come from dynamic stochastic general equilibrium models. Those estimates are obtained by minimizing the distance between the model predictions and the empirical impulse response function (Christiano et al., 2005), by maximizing the likelihood of the state space representation of the model (Bouakez et al., 2005), or by using Bayesian methods (Smets and Wouters, 2007).

The first step of any meta-analysis is to gather the empirical studies on the topic, usually referred to as "primary studies." To collect primary studies, meta-analyses in economics often employ the RePEc or EconLit databases. We use Google Scholar because it provides powerful full-text search, whereas RePEc and EconLit only allow searching through abstracts and keywords related to the studies, thereby making it harder to devise an exhaustive search query. We first collect papers that contain the exact phrases "habit formation" or "habit persistence" and, at the same time, feature occurrences or synonyms of the following words: consumption, estimate, regression, and empirical. After reading the abstracts of the studies returned by our search query we download those that show any promise of containing empirical estimates of the habit formation parameter. In the next step we extend our search to the references of these studies and add the last study on 31 January 2014.

We apply the following three inclusion criteria. First, the study must provide an empirical estimate of the habit formation parameter. Second, the study must include an estimate of the standard error (or a statistic from which the standard error can be computed). We need standard errors to be able to test for potential publication bias.[1] Finally, the third inclusion criterion is that the study must be published in a peer-reviewed journal. Meta-analyses differ in their treatment of unpublished results—sometimes they include unpublished papers as well, especially when the resulting data set would otherwise be small. Since there are many published studies estimating the habit formation parameter, we prefer to focus on studies that have been subjected to a peer-review process. We find 69 studies that comply with our selection criteria, and we list them in Appendix B.

Each primary study typically reports several estimates, and the median number of estimates per study is four. It is hard to pin down each study's representative estimate, because the authors themselves rarely say explicitly which one they prefer. Therefore, we collect all estimates reported in each study. This approach results in an unbalanced data set, as some studies report many more estimates than others—nevertheless, it allows us to exploit the differences in data and method choices both within and across individual studies. Wherever possible, we include study fixed effects to filter out the effects of study-level characteristics that are otherwise unobservable. All studies combined provide us with 567 estimates of the habit formation parameter, and for each of them we collect 22 variables reflecting the context in which researchers obtain the estimates.

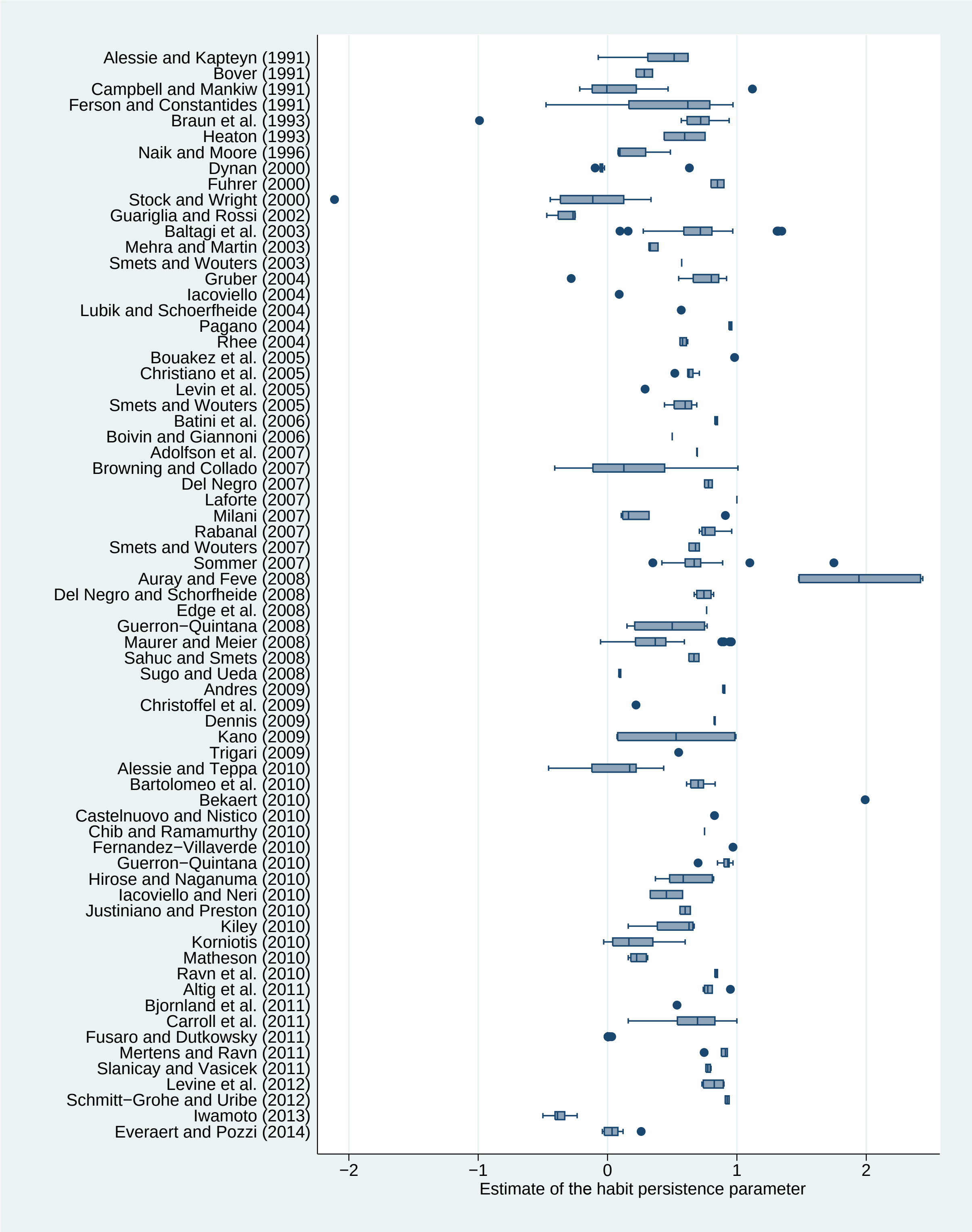
Figure 2 shows a box plot of the estimates that we include in our data set. Three features of the data stand out. First, most studies tend to report estimates lying between 0 and 1; that is, estimates that are consistent with the habit formation hypothesis (estimates above 1 are inconsistent with theory, while negative estimates reject habit formation in favor of durability of the consumption good under investigation). Second, even in the 0–1 range the estimates differ substantially within and between studies, with values around 0.5 being the most common. Third, estimates rejecting habit formation are not rare, and appear on both sides of the distribution. In the literature we generally encounter estimates lying between $-2$ and 2.

Figure 3 presents a histogram of the estimated parameters, providing additional insights. First, the distribution of the estimates is far from normal, and both the lower and upper boundaries of the range 0–1, consistent with habit formation, seem to affect the probability of an estimate being reported.[2] Second, while not normal, the distribution of estimates is relatively symmetric, as both the lower and the upper tails are cut off, and the mean estimate virtually equals the median. Third, studies published in the top five general interest journals tend to report slightly smaller estimates of the habit formation parameter than other studies. Fourth, the histogram has multiple peaks, suggesting heterogeneity generated by different estimation methods, which we investigate in detail in Section 4.

[1] Some studies provide estimates both with and without standard errors. In these cases we collect all estimates from the studies and use the estimates without standard errors in regressions that do not control for publication bias. Our results are robust to excluding all estimates for which standard errors are not reported.

[2] This result may also reflect the constraints that researchers use in the process of estimation, but a large majority of estimates are unconstrained.

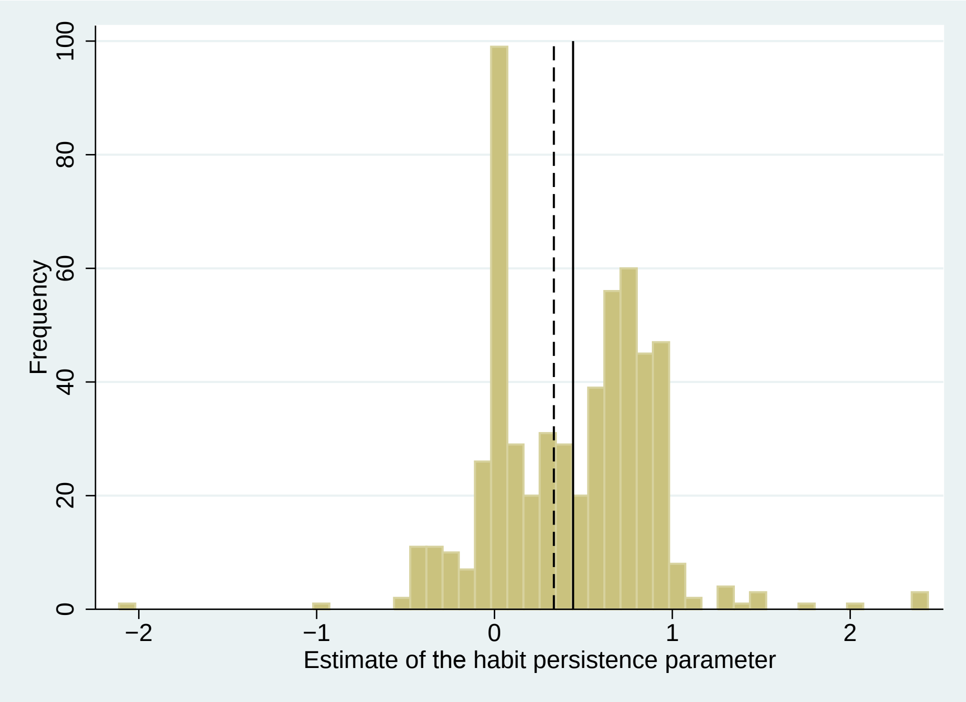
***Figure 2: Estimated Habit Formation Parameters Vary Considerably***




***Notes:*** The figure shows a box plot of the estimates of the habit formation parameter reported in individual studies. Full references for the individual studies used in the meta-analysis are available in Appendix B.

To shed some light on the sources of heterogeneity, we compute average and median values for different groups of estimates and display them in Table 1. The overall mean of the reported estimates

***Figure 3: Studies in Top Journals Report Slightly Smaller Estimates***




***Notes:*** The figure shows a histogram of the estimates of the habit formation parameter reported in the individual studies. The solid vertical line denotes the median of all the estimates. The dashed line denotes the median of estimates reported in studies published in the top five general interest journals: *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*.

***Table 1: Habit Formation Estimates for Different Data and Methods***

| | Unweighted | | | | Weighted | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | Median | 5% | 95% | Mean | Median | 5% | 95% | No of est. |
| All estimates | 0.42 | 0.44 | -0.27 | 0.97 | 0.57 | 0.63 | -0.12 | 0.99 | 567 |
| Micro studies | 0.10 | 0.00 | -0.39 | 0.62 | 0.13 | 0.09 | -0.41 | 0.62 | 190 |
| Macro studies | 0.58 | 0.67 | -0.08 | 1.00 | 0.65 | 0.71 | 0.05 | 1.00 | 377 |
| Internal | 0.27 | 0.09 | -0.38 | 0.94 | 0.42 | 0.50 | -0.33 | 0.98 | 344 |
| External | 0.66 | 0.67 | 0.16 | 1.00 | 0.73 | 0.71 | 0.16 | 1.49 | 223 |
| Micro - internal | 0.03 | 0.00 | -0.40 | 0.60 | 0.10 | 0.08 | -0.41 | 0.62 | 154 |
| Micro - external | 0.40 | 0.37 | 0.06 | 0.96 | 0.40 | 0.37 | 0.06 | 0.96 | 36 |
| Macro - internal | 0.46 | 0.62 | -0.28 | 0.97 | 0.54 | 0.69 | -0.08 | 0.98 | 190 |
| Macro - external | 0.71 | 0.71 | 0.21 | 1.00 | 0.74 | 0.73 | 0.16 | 1.49 | 187 |
| Macro - non DSGE | 0.54 | 0.63 | -0.18 | 1.12 | 0.55 | 0.60 | -0.12 | 1.48 | 248 |
| Macro - DSGE | 0.67 | 0.73 | 0.16 | 0.97 | 0.70 | 0.73 | 0.16 | 1.00 | 129 |

*Notes*: 5% and 95% denote the corresponding percentiles. Weighted = summary statistics based on the observations weighted by the inverse of the number of estimates reported per individual study. In such case each study receives the same weight in the computation of the summary statistics.

is approximately 0.4. Studies using micro data deliver much smaller estimates on average—about 0.1. By contrast, macro studies tend to generate larger estimates: around 0.6. Among the macro approaches to assessing habit formation, DSGE studies tend to yield slightly larger estimates. The nature of the habit formation process matters, too. Estimates of internal habit formation average 0.3, while estimates of external habits tend to be more than twice as large at around 0.7. The difference between estimates of external and internal habits remains substantial even when we calculate the means separately for macro and micro studies. For macro data, estimates of external habits are still larger—this finding contrasts with the argument of Carroll et al. (1997), who suggest that estimates of external and internal habits are empirically indistinguishable when using macro data. These conclusions remain intact even when we weight the estimates by the inverse of the number of estimates reported in each study, thereby giving each study the same weight regardless of the number of estimates the study produces.

***Table 2: Habit Formation Differs Across Countries***

| | Unweighted | | | | Weighted | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | Median | 5% | 95% | Mean | Median | 5% | 95% | No of est. |
| *All estimates* | | | | | | | | | |
| US | 0.42 | 0.39 | -0.05 | 0.97 | 0.64 | 0.70 | 0.00 | 1.00 | 353 |
| EU countries | 0.51 | 0.63 | -0.27 | 1.00 | 0.45 | 0.59 | -0.28 | 0.92 | 146 |
| Japan | 0.06 | -0.24 | -0.46 | 0.94 | 0.30 | 0.10 | -0.41 | 0.96 | 26 |
| Other countries | 0.34 | 0.30 | -0.03 | 0.78 | 0.36 | 0.31 | -0.03 | 0.98 | 42 |
| *Micro estimates* | | | | | | | | | |
| US | 0.12 | 0.00 | -0.05 | 0.63 | 0.18 | 0.10 | -0.06 | 0.59 | 133 |
| EU countries | 0.10 | 0.07 | -0.46 | 0.99 | 0.08 | 0.03 | -0.46 | 0.62 | 36 |
| Japan | -0.37 | -0.39 | -0.50 | -0.23 | -0.37 | -0.39 | -0.50 | -0.23 | 14 |
| Other countries | 0.59 | 0.58 | 0.56 | 0.62 | 0.59 | 0.58 | 0.56 | 0.62 | 7 |
| *Macro estimates* | | | | | | | | | |
| US | 0.60 | 0.69 | -0.08 | 1.00 | 0.70 | 0.75 | 0.09 | 1.10 | 220 |
| EU countries | 0.64 | 0.70 | -0.08 | 1.12 | 0.60 | 0.69 | -0.05 | 0.94 | 110 |
| Japan | 0.57 | 0.68 | 0.02 | 0.96 | 0.55 | 0.73 | 0.09 | 0.96 | 12 |
| Other countries | 0.29 | 0.24 | -0.04 | 0.93 | 0.30 | 0.21 | -0.04 | 0.98 | 35 |

*Notes*: 5% and 95% denote the corresponding percentiles. Weighted = summary statistics based on the observations weighted by the inverse of the number of estimates reported per individual study. In such case each study receives the same weight in the computation of the summary statistics.

Most estimates in our data set are obtained using US data (63%). All studies combined provide results for 17 countries, arguably contributing to the heterogeneity we observe, but the number of countries is not large enough to connect the differences in estimates to the structural differences among the economies. Nevertheless, in Table 2 we compare group averages for the US, Japan, countries belonging to the EU, and the rest of the countries (other OECD economies, such as Australia, Canada, New Zealand, and Korea) and notice several regularities. The estimates of habit formation for the US and EU tend to be larger on average than the estimates for Japan and other countries. This result holds even when we separate macro and micro estimates—the only exception is the group "other countries" for micro data, which, however, only includes seven observations. It is not clear how to interpret these differences: cross-country papers focusing on habit formation are rare, and the prominent study of this category, Carroll et al. (2011), finds homogeneous coefficients for a number of countries in our sample. One conclusion we feel confident to make is that the available empirical literature is inconsistent with the hypothesis of habit formation in consumption for

Japan (not covered by Carroll et al., 2011), in sharp contrast to the US and countries of the European Union.

## 3. Publication Bias

The mean and median reported estimates may represent a biased reflection of the underlying research results if some estimates are more likely to be selected for publication than others. For this reason, most meta-analyses test—and, if necessary, correct—for publication bias. Brodeur et al. (2013) collect 50,000 p-values reported in economics and document widespread publication bias. A recent survey among the members of the European Economic Association, Necker (2014), reveals that a third of economists in Europe admit that they have engaged in presenting empirical findings selectively so they confirm their arguments and in searching for control variables until they get a desired result. Doucouliagos and Stanley (2013) survey meta-analyses conducted in economics and find that most fields suffer from the bias, as editors, referees, or authors themselves prefer statistically significant results that have an intuitive sign.

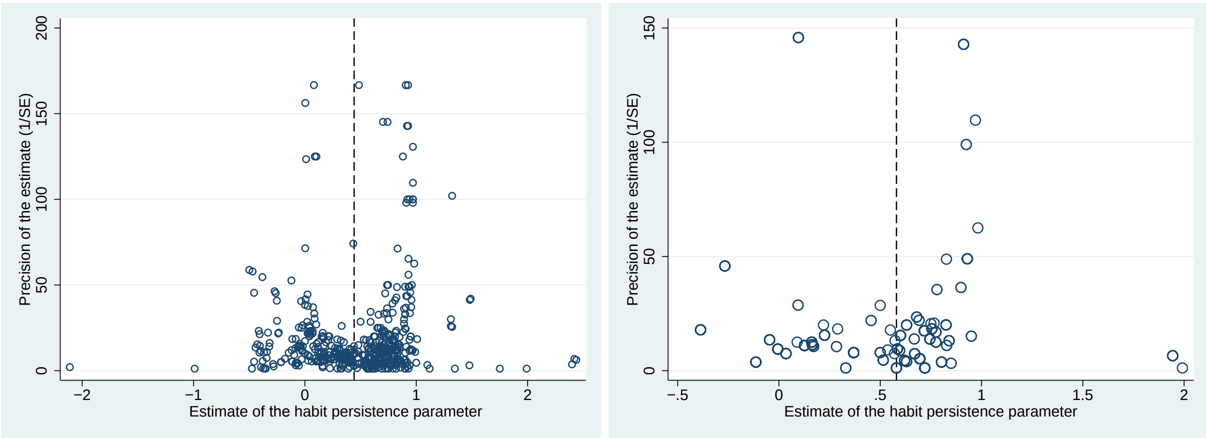
For example, Havranek (2015) finds strong publication bias in the literature that uses consumption Euler equations to estimate the elasticity of intertemporal substitution (often the same specification used to estimate habit formation). Most economists believe that the elasticity of substitution should be positive because negative elasticity implies a convex utility function. Therefore, negative estimates of the elasticity are rarely reported in the literature, as are estimates that are statistically insignificant. The under-reporting of negative estimates and estimates that are positive but small and imprecise biases the means upward because it is not matched by corresponding under-reporting of large imprecise estimates.

The empirical literature on habit formation differs from studies estimating the elasticity of intertemporal substitution in two major ways. First, negative estimates of the habit formation parameter allow for intuitive interpretation: although inconsistent with habit formation, they may result from durability of the consumption measure used in the estimation—and may thus be more publishable than negative estimates of the elasticity of intertemporal substitution. Second, unlike large estimates of the elasticity, estimates of the habit formation parameter that exceed 1 are implausible because they imply non-stationary consumption growth. Figure 3, discussed in Section 2, suggests that the most common estimates lie close to the midpoint between the lower and upper boundaries of the 0–1 interval (consistent with habit formation), and that when an estimate surpasses either limit, its probability of being reported drops—in other words, both very small and very large estimates are sometimes discarded by the researchers. This relative symmetry in decision rules on discarding implausible estimates implies that even if there is publication selection in the literature on habit formation, it does not necessarily lead to publication bias.

To test for potential publication bias researchers often evaluate the so-called funnel plots (Egger et al., 1997). A funnel plot is a scatter plot of the estimates (on the horizontal axis) against the inverse of their standard errors, the estimates' precision (on the vertical axis). In the absence of publication bias the scatter plot forms an inverted funnel: the most precise estimates lie close to each other, while the less precise ones are more dispersed. The funnel plot should be symmetric because most estimation methods presuppose that the ratio of the estimate to its standard error exhibits a symmetric distribution. In other words, all imprecise estimates, small and large, should have the same probability of being reported. If some estimates are reported less often because of their magnitude, the funnel will become asymmetric; if statistically insignificant estimates get under-reported, the funnel will become hollow.

***Figure 4: Funnel Plots Suggest Slight Publication Bias***

(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 estimates of the habit formation parameter. The dashed vertical lines denote the mean of all the estimates in panel (a) and the mean of the median estimates reported in the studies in panel (b). Multiple peaks of the funnel suggest heterogeneity. For ease of exposition we exclude estimates with extreme precision values from the figure, but we use all the estimates in the statistical tests.

The vast majority of the estimates in our sample are obtained via estimation methods presupposing that the estimates have a $t$-distribution (such as GMM, TSLS, or OLS). These methods do not place explicit constraints on the estimates that force them to lie between 0 and 1; therefore, the estimates can lie outside the $(0,1)$ interval even if the habit formation hypothesis holds, given sufficient imprecision in estimation. Figure 4 presents funnel plots for the estimates of the habit formation parameter. The left panel depicts all estimates, while the right panel plots median estimates reported in the studies against their precision. The plots show signs of asymmetry, and both 0 and 1 seem to be the boundaries that affect the probability of estimates being reported. The upper limit seems to be slightly more important than the lower one. An explanation of this result is that while negative estimates can be interpreted as evidence of durability, estimates larger than 1 are inconsistent with theory and are thus harder to justify. Researchers may consider these large estimates as evidence of model misspecification and adjust their models accordingly to produce more intuitive results.

Compared with funnel plots reported in other economic meta-analyses, the funnel plot for habit formation estimates seems to be less skewed—thus, the publication bias in this literature might be partially offset by the discarding of negative estimates. In what follows we test funnel asymmetry formally. We assess the extent of the bias and uncover the underlying mean estimate of habit formation. Our specification is based on Card and Krueger (1995) and Stanley (2008):

$$HABIT_{ij} = \alpha_0 + \beta \cdot SE(HABIT_{ij}) + \varepsilon_{ij}, \tag{3.3}$$

where $HABIT_{ij}$ is the $i$-th estimate from $j$-th study, $SE(HABIT_{ij})$ is the reported standard error of this estimate, and $\varepsilon_{ij}$ is the disturbance term. As we have mentioned, most empirical methods estimating habit formation are based on the assumption that the ratio of the estimate to the standard error is $t$-distributed. This property implies that the numerator and the denominator of the ratio should be statistically independent quantities. Correlation between the two variables arises because of publication bias: suppose that researchers would only like to report estimates that are positive and statistically significant. Given the particular data and estimation technique (and thus given the standard error), they would need to search for a specification that delivers a point estimate of habit formation large enough to offset the standard error and show significance. Therefore, coefficient $\beta$ in regression (3.3), capturing the relation between estimates and their standard errors, indicates the magnitude of publication bias. $\alpha_0$ is the mean estimate of the habit formation parameter conditional on the standard error approaching zero: it shows the mean reported habit formation parameter corrected for publication bias.

***Table 3: Funnel Asymmetry Tests Indicate No Publication Bias***

| | Baseline | Instrument | Study | Precision | Median |
|---|---|---|---|---|---|
| SE (publication bias) | -0.130 | -0.241 | -0.0521 | 0.177*** | 0.395 |
| | (0.272) | (0.898) | (0.247) | (0.0294) | (0.247) |
| Constant (effect beyond bias) | 0.446*** | 0.462*** | 0.575*** | 0.00164*** | 0.523*** |
| | (0.0378) | (0.125) | (0.0324) | (0.0000384) | (0.0575) |
| Observations | 558 | 558 | 558 | 558 | 69 |

***Notes:*** The table presents the results of regression $HABIT_{ij} = \alpha_0 + \beta \cdot SE(HABIT_{ij}) + \varepsilon_{ij}$. $HABIT_{ij}$ and $SE(HABIT_{ij})$ are the $i$-th estimates of the habit formation parameter and their standard errors reported in the $j$-th studies. The standard errors of the regression parameters are clustered at study level. All estimations except for the last include study fixed effects. *Instruments*: we use the inverse of the square root of the number of observations in the individual study as an instrument for the standard error of the estimate of the habit formation parameter. *Study*: we weight the estimates by the inverse of the number of estimates reported in the study. *Precision*: we weight the estimates by the inverse of the reported estimate's standard error. *Median*: we estimate the equation by including the median estimate of the habit formation parameter and the median standard error of the estimated habit formation parameter reported in the individual studies.

The majority of the estimates in our sample are obtained using techniques that yield standard errors directly—for those estimates we simply collect the published statistics. Several macro studies, however, use Bayesian methods to estimate the coefficient and employ an asymmetric prior distribution for the habit formation parameter. We approximate the standard errors of these estimates with the standard deviations reported for the posterior mean estimates of the parameter. This simplification *per se* might introduce a slight correlation between the estimates and their standard errors, but our results do not change qualitatively when we exclude the Bayesian estimates. Furthermore, while several macro studies report very small standard errors (especially studies that place explicit constraints on the habit formation parameter), others report standard errors that are many orders of magnitude greater. To account for these outliers we winsorize the data on standard errors at 5% on both sides of the distribution. Our main results are not sensitive to the choice of the fraction of data to be censored at each tail (censoring at 0.5% delivers largely similar results). The results are also robust to dropping the observations from the 5% tails on each side of the distribution.

Table 3 presents the estimates of regression (3.3); these results can also be interpreted as a test of funnel plot asymmetry. We consider several versions of the test. First, we estimate an OLS regression with study fixed effects and standard errors clustered at the study level. We include fixed effects to filter out unobservable study-specific factors that influence the reported estimates. Second, to address the potential endogeneity problem in meta-analysis we estimate the regression using the instrumental variable technique, while also including study fixed effects. Some method choices are likely to affect both the estimate and its standard error in the same direction, thus creating correlation between the disturbance term $\varepsilon_{ij}$ and $SE(HABIT_{ij})$ and resulting in an inconsistent estimate of $\beta$. As an instrument, we use the inverse of the square root of the number of observations used in each primary study: this variable is roughly proportional to the standard error, but not likely to be correlated with the method choice. Third, we estimate the regression by weighting each estimate by the inverse of the number of estimates reported in the corresponding study, thereby giving each study an equal weight in the regression. Fourth, we weight the estimates by their precision to remove heteroskedasticity. Finally, we exploit between- (instead of within-) study variation in the data using the median estimates and median standard errors reported in the primary studies.

The results can be summarized as follows. Four methods out of five yield insignificant estimates of $\beta$ (the magnitude of publication bias) and significant estimates of $\alpha_0$ (the underlying mean habit formation parameter corrected for publication bias). We estimate the mean corrected habit formation to be around 0.5, close to the sample mean and median reported in the previous section. These results suggest that publication selection does not create a substantial bias in the reported habit formation parameters.

In contrast, the precision-weighted specification delivers a statistically significant estimate of publication bias and a much smaller underlying mean for habit formation. While precision-weighting removes heteroskedasticity, it is highly sensitive to small values of the standard error. We have noted that some studies in our sample place explicit constraints on the habit formation parameter. These studies are likely to obtain tiny standard errors, thus gaining large weights in the precision-weighted estimation. Moreover, this estimation yields a positive estimate of $\beta$, suggesting an upward publication bias, which is at odds with the intuition suggested by Figure 4. According to the guidelines by Doucouliagos and Stanley (2013), the estimate of $\beta$ around 0.177 can be classified as "little to modest" publication bias, and would have to be more than five times larger to be classified differently. Finally, the results of the precision-weighted specification do not hold if we employ instrumental variable estimation, using the inverse of the square root of the number of observations as an instrument for the standard error (this specification is not reported). Therefore, we argue that precision-weighted estimation overstates the effect of publication bias.

To sum up, while we find some indications of publication selection related to the 0 and 1 thresholds that define the range consistent with habit formation, we find little evidence of any systematic bias resulting from this selection. Our findings suggest that the effects of selection against negative estimates and selection against estimates larger than 1 cancel each other out, rendering the mean estimate reported in the habit formation literature unbiased.

# 4. Why Do Estimates of Habit Formation Vary?

## 4.1 Explanatory Variables

We have noted that the estimates of habit formation differ substantially within and between studies. In this section we attempt to relate the differences in the estimates to differences in the design of primary studies. To this end we collect 22 variables that reflect the data characteristics of each study, its geographical coverage, the variable definitions and estimation technique that the study employs, and the study's publication characteristics (for example the number of citations). We cannot hope that these 22 variables will explain all differences across estimates—the set of potential explanatory variables is unlimited—but we believe that our selection reflects the most common choices faced by researchers who estimate habit formation.

**Data characteristics** For each study we collect the number of observations and average year of the data used. We specify whether the study employs micro or aggregate data, and whether it estimates a regression-type model or a dynamic stochastic general equilibrium model: DSGE studies estimating the habit formation parameter often use Bayesian methods, and their results are affected by the prior values of the parameter that the researchers employ. We also account for the frequency of the data used. Bansal et al. (2012) argue that studies estimating consumption Euler equations should account for the difference between the econometrician's sampling frequency and consumers' decision frequency; the authors estimate the latter to be approximately monthly. Habit formation estimates are likely to be affected by the data frequency because at sufficiently high frequencies every consumption good displays durability, rendering the habit formation parameter negative: a full meal makes people saturated for the next few hours. Most studies employ quarterly data; for those using monthly and annual data we include controls.

**Countries examined** Although habit formation is supposed to be a so-called deep parameter, differences in structural characteristics of economies (such as culture) might cause the parameter to vary across countries. Havranek et al. (2015) find substantial cross-country heterogeneity in the elasticity of intertemporal substitution in consumption associated with cross-country differences in income and stock market participation. Since the number of countries investigated by the studies in our sample is small, we only use regional dummy variables. We include dummies for the US data, for the data on Japan, and for data from countries that are members of the European Union. The remaining studies estimate the habit formation parameter for other non-European OECD countries.

**Variable definitions** In Section 2 we show that the mean estimates of internal and external habit formation parameters differ. To see whether the difference holds even after we control for other aspects of data and methodology, we create a dummy variable attributed to the type of habits under investigation. Estimates may also differ depending on the consumption good used in the estimation. Studies that include durable goods should obtain lower estimates of the habit formation parameter, while estimates based on food consumption may be biased if food is non-separable from other consumption goods (Attanasio and Weber, 1995). We distinguish three categories of consumption proxies: food consumption, total non-durable consumption, and measures that include durable consumption; the use of non-durable consumption is our reference group.

**Estimation approach** It is common wisdom in empirical economics that different estimation approaches often yield different results. We want to find out whether the use of a particular method is associated with systematic differences in the reported habit formation parameter. Most studies estimate habit formation by using GMM; some assume homoskedasticity and employ TSLS. Many studies that estimate DSGE models use Bayesian techniques, while other DSGE studies use the minimum distance method, matching empirical impulse response functions to those generated within the models. Some studies employ maximum-likelihood-based methods, and a few panel studies use fixed effects estimation that does not account for Nickell (1981) bias or random effects estimation assumptions of which are unlikely to hold in consumption Euler equations. A small fraction of studies estimate habit formation with OLS—we use this estimation approach as the reference group.

**Publication characteristics** Finally, we control for the publication characteristics of individual studies. We include the year of publication to capture methodological advances that are otherwise hard to codify or that have not been employed by a sufficient number of studies yet. To account for approximate study quality beyond the observed differences in data and methodology, we include the number of citations, the recursive impact factor of the journal that published the study, and a dummy variable for studies published in top journals. We collect the data on the impact factor from the RePEc: unlike other databases, RePEc covers virtually all economics journals and provides a discounted recursive impact factor well-suited for comparison of outlets in economics.

Table 4 describes the 22 explanatory variables mentioned above, listing their means, standard deviations, and means weighted by the inverse of the number of estimates reported in individual studies. The correlation matrix of all the explanatory variables is presented in Figure A1 in Appendix A and shows that the variables reflect different aspects of the studies. The largest correlation appears between micro data and the number of observations: micro-level studies tend to have more observations available than macro studies. Furthermore, Bayesian techniques are often employed within the framework of DSGE models, many micro papers use GMM, and newly published studies tend to use fresher data, which is also intuitive.

***Table 4: Description and Summary Statistics of Regression Variables***

| Variable | Description | Mean | Std. dev. | WM |
|---|---|---|---|---|
| Habit | The estimate of the habit formation parameter (response variable). | 0.42 | 0.45 | 0.57 |
| SE | The standard error of the estimate of the habit formation parameter. | 0.14 | 0.20 | 0.13 |
| *Data characteristics* | | | | |
| No of obs. | The logarithm of the number of observations. | 6.21 | 1.83 | 5.51 |
| Average year | The midpoint of the sample used for the estimation of habit formation (the base is the sample minimum: 1932). | 54.5 | 11.3 | 53.6 |
| Micro | = 1 if micro data are used for the estimation. | 0.34 | 0.47 | 0.16 |
| DSGE | = 1 if the estimation uses a dynamic stochastic general equilibrium model. | 0.23 | 0.42 | 0.55 |
| Monthly | = 1 if the frequency of the data used for the estimation is monthly. | 0.16 | 0.37 | 0.04 |
| Annual | = 1 if the frequency of the data used for the estimation is annual. | 0.33 | 0.47 | 0.20 |
| *Countries examined* | | | | |
| US | = 1 if habit formation is estimated for the US. | 0.62 | 0.49 | 0.68 |
| EU | = 1 if habit formation is estimated for a country belonging to the EU. | 0.26 | 0.44 | 0.20 |
| Japan | = 1 if habit formation is estimated for Japan. | 0.05 | 0.21 | 0.05 |
| *Variable definition* | | | | |
| External | = 1 if external habit formation is estimated. | 0.39 | 0.49 | 0.48 |
| Durable | = 1 if durable consumption goods are included in the measure of consumption. | 0.78 | 0.41 | 0.81 |
| Food | = 1 if food expenditures are used as a proxy for consumption. | 0.11 | 0.32 | 0.07 |
| *Estimation approach* | | | | |
| GMM | = 1 if the general method of moments is employed for the estimation. | 0.45 | 0.50 | 0.24 |
| TSLS | = 1 if the two-step-least-squares method is employed for the estimation. | 0.15 | 0.36 | 0.07 |
| Bayes | = 1 if the estimation uses Bayesian inference. | 0.20 | 0.40 | 0.43 |
| Minimum distance | = 1 if the minimum distance method is employed for the estimation. | 0.05 | 0.22 | 0.09 |
| ML | = 1 if the maximum likelihood method is employed for the estimation. | 0.02 | 0.16 | 0.09 |
| Panel | = 1 if a panel technique (fixed effects, random effects) is employed for the estimation. | 0.06 | 0.23 | 0.03 |
| *Publication characteristics* | | | | |
| Publication year | The year in which the study was published (the base is the sample minimum: 1991). | 15.0 | 6.4 | 15.1 |

Continued on next page

***Table 4: Description and Summary Statistics of Regression Variables (continued)***

| Variable | Description | Mean | Std. dev. | WM |
|---|---|---|---|---|
| Citations | The logarithm of the mean number of Google Scholar citations received per year since the study was published (collected in August 2014). | 0.55 | 0.32 | 0.62 |
| Top journal | = 1 if the study was published in one of the top five journals in economics. | 0.07 | 0.25 | 0.12 |
| Impact | The recursive discounted RePEc impact factor of the outlet (collected in August 2014). | 0.73 | 0.66 | 0.89 |

*Notes*: The variables are collected from published studies estimating the habit formation parameter. The following journals are considered top journals in economics: American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, and Review of Economic Studies. WM = mean weighted by the inverse of the number of estimates reported in a study.

## 4.2 Estimation and Results

To investigate the influence of study design on the estimated habit formation parameter, we consider the following regression:

$$HABIT_{ij} = \alpha_0 + \sum_{k=1}^{22} \beta_k X_{k,ij} + \varepsilon_{ij}, \tag{4.4}$$

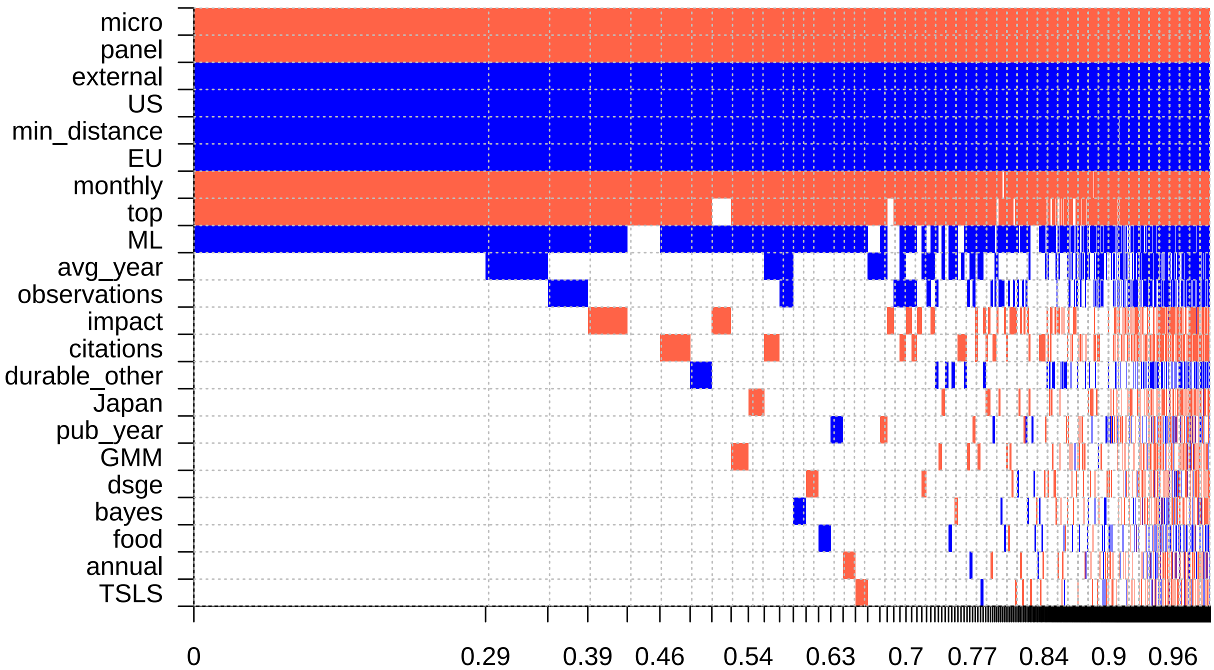
where $X_{k,ij}$ denotes the value of a $k$-th explanatory variable for an $i$-th estimate from a $j$-th study. We believe that each variable in our set can contribute to explaining the heterogeneity among the estimates. But including all 22 variables in the regression would inflate the standard errors and yield inefficient estimates, because some of the variables are likely to prove redundant. The theory does not give us enough guidance to determine the exact subset of the 22 variables that should be included in the final regression. Sequential $t$-testing (sometimes called the "general-to-specific approach"), which is often used to decide which variables belong to the underlying model, is not statistically valid and gives rise to the possibility of excluding relevant variables. The large number of potential variables thus brings about problems related to model uncertainty that could result in severely erroneous inference. To address these issues, we employ the Bayesian Model Averaging technique (BMA)—a method that does not require selecting one individual specification.

Inference in BMA is based on a weighted average of individual regressions that include different combinations of explanatory variables; the weights reflect the posterior model probabilities (PMPs) of the corresponding individual specifications. PMPs can be thought of as a Bayesian analogy of the adjusted R-squared or information criteria used in frequentist econometrics. Researchers typically want to check the robustness of their results by estimating several regressions that include different combinations of explanatory variables; BMA generalizes this approach. Our intention here is to explain the basics of the BMA method and the terms needed for inference, not to give an exhaustive introduction to the BMA procedure; readers interested in such a treatment should consult Koop (2003) for an introduction and Moral-Benito (2015) for a survey of BMA applications in economics.

All of the computations are performed using the R package `BMS` for Bayesian model averaging available at http://bms.zeugner.eu. Estimating all $2^{22}$ possible specifications is computationally too demanding—therefore, we approximate the whole model space by using the Model Composition Markov Chain Monte Carlo algorithm (Madigan and York, 1995), which only traverses the most important part of the model space: that is, the models with high posterior model probabilities. Such a simplification is commonly applied in applications of BMA (see, for example, Feldkircher and Zeugner, 2009).

For the BMA estimation we have to choose priors for the parameters and model space. We follow Eicher et al. (2011), who recommend using the unit information prior for the parameters and the uniform model prior for the model space because these priors perform well in predictive exercises. Our prior setting can be interpreted as follows: the unit information prior provides the same amount of information as one observation of data, while the uniform model prior means that each model has the same prior probability (thereby giving higher prior probabilities to medium model *sizes*). As a robustness check, we also study alternative prior setups. To this end, we employ the benchmark $g$-priors for parameters suggested by Fernandez et al. (2001) along with the beta-binomial model prior for the model space, which gives each model size equal prior probability (Ley and Steel, 2009); we also use the data-dependent hyper-g prior suggested by Feldkircher and Zeugner (2012), which should be less sensitive to noise in the data.

***Figure 5: Model Inclusion in Bayesian Model Averaging***




***Notes:*** Response variable: the estimate of the habit formation parameter. Columns denote individual models; variables are sorted by posterior inclusion probability in descending order. Blue color (darker in greyscale) = the variable is included and the estimated sign is positive. Red color (lighter in greyscale) = the variable is included and the estimated sign is negative. No color = the variable is not included in the model. The horizontal axis measures the cumulative posterior model probabilities. Numerical results of the BMA exercise are reported in Table 5. A detailed description of all variables is available in Table 4.

Figure 5 presents the results of the Bayesian model averaging exercise. The variables are sorted from top to bottom rows by posterior inclusion probability (which can be thought of as a Bayesian analogy of statistical significance), while the columns denote individual models. The color of the cell reflects the sign of the corresponding regression coefficient: negative signs are depicted in red (lighter in greyscale), positive in blue (darker in greyscale); a white cell means that the variable is not included in the given model. The width of the columns is proportional to the posterior model probability (that is, how well the model fits the data relative to its size). We can see that the model that includes all 22 variables is only one of many specifications estimated by BMA. The figure suggests that the most important variables in explaining the heterogeneity among the estimates are *micro*, *panel*, *external*, *US*, *minimum distance*, *EU*, *monthly*, *top*, and *ML*. The regression signs for these variables are stable regardless of whether other control variables are included.

***Table 5: Explaining the Differences in the Estimates of Habit Formation***

| Response variable: | Bayesian model averaging | | | Frequentist check (OLS) | | |
|---|---|---|---|---|---|---|
| Estimate of habit formation | Post. mean | Post. std. dev. | PIP | Coef. | Std. er. | p-value |
| *Data characteristics* | | | | | | |
| No of obs. | 0.005 | 0.015 | 0.154 | | | |
| Average year | 0.001 | 0.002 | 0.224 | | | |
| Micro | -0.361 | 0.074 | 1.000 | -0.327 | 0.071 | 0.000 |
| DSGE | -0.001 | 0.012 | 0.048 | | | |
| Monthly | -0.275 | 0.068 | 0.993 | -0.268 | 0.068 | 0.000 |
| Annual | 0.000 | 0.009 | 0.042 | | | |
| *Countries examined* | | | | | | |
| US | 0.295 | 0.057 | 0.999 | 0.285 | 0.106 | 0.007 |
| EU | 0.236 | 0.053 | 0.996 | 0.237 | 0.118 | 0.044 |
| Japan | -0.005 | 0.030 | 0.057 | | | |
| *Variable definition* | | | | | | |
| External | 0.175 | 0.038 | 0.999 | 0.179 | 0.053 | 0.001 |
| Durable | 0.003 | 0.014 | 0.068 | | | |
| Food | 0.001 | 0.015 | 0.042 | | | |
| *Estimation approach* | | | | | | |
| GMM | -0.002 | 0.014 | 0.055 | | | |
| TSLS | 0.000 | 0.010 | 0.042 | | | |
| Bayes | 0.000 | 0.011 | 0.046 | | | |
| Minimum distance | 0.298 | 0.073 | 0.996 | 0.312 | 0.230 | 0.175 |
| ML | 0.253 | 0.134 | 0.859 | 0.303 | 0.109 | 0.005 |
| Panel | -0.446 | 0.073 | 1.000 | -0.430 | 0.048 | 0.000 |
| *Publication characteristics* | | | | | | |
| Publication year | 0.000 | 0.001 | 0.056 | | | |
| Citations | -0.011 | 0.035 | 0.126 | | | |
| Top journal | -0.242 | 0.091 | 0.937 | -0.275 | 0.130 | 0.035 |
| Impact | -0.010 | 0.028 | 0.150 | | | |
| Constant | 0.226 | NA | 1.000 | 0.285 | 0.106 | 0.007 |
| Studies | 69 | | | 69 | | |
| Observations | 567 | | | 567 | | |

***Notes:*** PIP = posterior inclusion probability. In the frequentist check we only include method characteristics with PIP > 0.5. Standard errors in the frequentist check are clustered at the study level. More details on the BMA estimation are available in Table A1 and Figure A2.

Table 5 presents the numerical results of Bayesian model averaging. In BMA the key statistic is the posterior inclusion probability (PIP), which reflects the importance of each variable. For a given variable, the PIP is calculated by summing the posterior model probabilities of all models in which the variable is included. According to the rule of thumb proposed by Jeffreys (1961) and refined by Kass and Raftery (1995), the significance of each regressor is weak, positive, strong, or decisive if the PIP lies between 0.5–0.75, 0.75–0.95, 0.95–0.99, or 0.99–1, respectively. In the right-hand part of the table we provide a simple frequentist check of our BMA exercise: we use OLS with clustered standard errors to estimate a regression that only includes variables that have at least a weak effect on the reported habit formation (that is, those with PIP$>0.5$). With one exception, the OLS estimation matches the BMA results.

Only two aspects of the data characteristics seem to have a systematic effect on the reported habit formation parameter: the choice between micro and macro data, and the frequency of the data, with both variables showing decisive PIPs. Micro-level studies tend to report smaller estimates of habit formation (by about 0.4), which corroborates the conclusion drawn from the summary statistics in Section 2. Moreover, studies using monthly data report estimates that tend to be smaller by 0.3—an intuitive result, since at higher frequencies a large fraction of the consumption bundle becomes durable. For example, clothing expenditure will probably show durability at monthly frequency, but not at annual frequency. Therefore, researchers can expect to get more evidence for habit formation when they move to lower frequencies. By contrast, the number of observations used in the estimation is not correlated with the magnitude of the reported habit formation parameter, and there is no apparent time trend in the results.

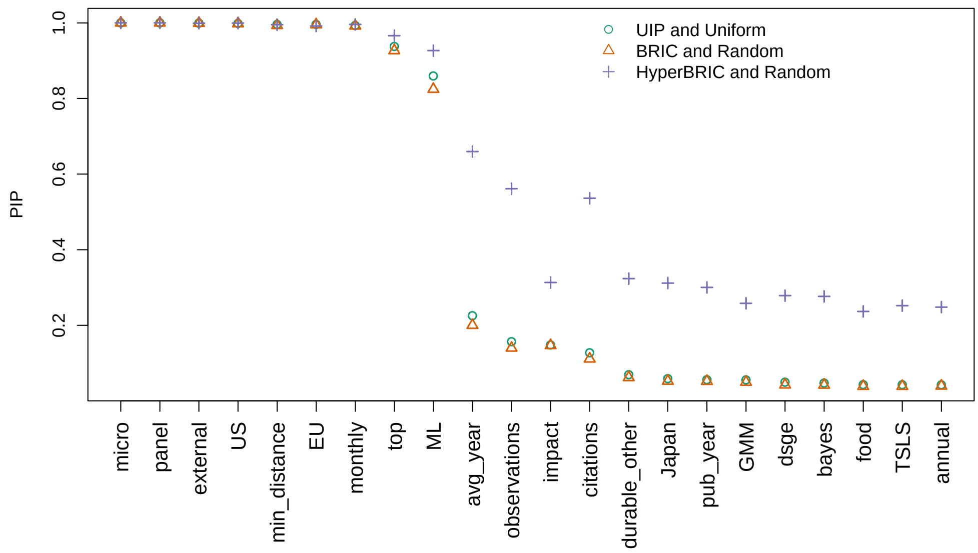
We find evidence of country heterogeneity in the estimates of habit formation. The parameters estimated for the US and EU tend to be 0.2–0.3 larger than those reported for other countries (and Japan in particular). To our knowledge, the only study that discusses cross-country differences in habit formation is Carroll et al. (2011), who find little heterogeneity across countries, but do not consider Japan. The cross-country differences in habit formation might reflect cultural differences—nevertheless, the specifics of the data may play a role, too. For instance, Carroll et al. (2011) mention several problems with Japanese data on consumption related to adjustments in the Japanese national accounts methodology.

Our results suggest that the estimates of external habit formation remain substantially larger than the estimates of internal habits (by about 0.2), even if we control for all other aspects of study design. Thus, the major driver of the observed habits in consumption seems to be "keeping up with the Joneses." We also find that the definition of consumption used for the estimation has little influence on the results, which is surprising but can be explained by the fact that the choice of the proxy for consumption is related to the choice between macro or micro data. Most micro studies only have data on food consumption, while many macro studies include durables and use total consumption as their benchmark.

Furthermore, we find that some estimation techniques deliver results that are systematically different from those obtained via other methods. The minimum distance and the maximum likelihood methods tend to yield larger estimates, while the use of simple panel data techniques without instruments results in estimates that are substantially smaller. Therefore, it is important to take into account the endogeneity created by including a lagged value of the dependent variable among the explanatory variables. Finally, we find that publication characteristics are not very important for the reported habit formation parameters, with the exception of publication in top journals, which is associated with reporting smaller estimates. The latter effect may arise partially because studies published in the best journals often use micro data.

In Figure 6 we report the posterior inclusion probabilities that would result from BMA estimations with alternative priors. The main results are very similar to the baseline case; the only difference worth mentioning is that with the random model prior—that is, if we give each model size the same prior probability—the average year of the data seems to have weak (instead of no) effect on the reported habit formation parameter. In all other respects the three estimated models yield results that are remarkably consistent in qualitative terms.

***Figure 6: Posterior Inclusion Probabilities Across Different Prior Settings***




***Notes:*** UIP and Uniform = priors according to Eicher et al. (2011), who recommend using the unit information prior for the parameters and the uniform model prior for model size, since these priors perform well in predictive exercises. BRIC and Random = we use the benchmark $g$-priors for parameters suggested by Fernandez et al. (2001) with the beta-binomial model prior for the model space, which means that each model size has equal prior probability (Ley and Steel, 2009). HyperBRIC and Random = we use the data-dependent hyper-$g$ prior suggested by Feldkircher and Zeugner (2012), which should be less sensitive to the presence of noise in the data.

We perform four further robustness checks. First, we include all of the explanatory variables in the regression and estimate the model using simple OLS with standard errors clustered at study level. Second, we use study fixed effects, in which case variables with no variation at study level are excluded. We present the results of the two robustness checks in Table A2 in the Appendix and conclude that our main findings are not sensitive to these changes in model specification. Third, we re-run the BMA exercise and the frequentist check for regressions weighted by precision of the estimates $(1/SE(HABIT))$. The precision-weighted results can be found in Table A3 in the Appendix, and the finding concerning the divergence between micro and macro estimates is robust even to this change in specification. Fourth, we run the BMA exercise on the sub-samples of estimates corresponding to internal habits (Table A4), external habits (Table A5), and micro data (Table A6). In this final robustness check we lose a lot of degrees of freedom, especially when macro estimates are excluded, but our main results prove to be insensitive to analyzing these more homogeneous sub-samples separately.

## 5. Concluding Remarks

We collect estimates of the habit formation parameter from studies published in peer-reviewed journals and provide the mean value for the entire sample, as well as for subsets of estimates featuring different aspects of study design. Namely, we calculate and compare mean estimates obtained in studies that use household level (micro) data and studies that employ aggregate (macro) data, studies investigating internal and external habit formation, and studies that assess habit formation for the US, countries of the European Union, and Japan.

We find that the mean value of the habit formation parameter reported in the literature is 0.4. The mean estimate reported in studies using micro data is 0.1, while for macro studies the mean equals

0.6. These values are not large enough to explain some of the best-known empirical puzzles in macroeconomics: for example, Constantinides (1990) shows that to explain the equity premium puzzle the habit formation parameter must exceed 0.8. The difference between micro and macro studies remains large and statistically significant even when we control for other aspects of study design. This divergence arises because micro and macro studies focus on different sources of variation in consumption: micro estimates exploit variation at the level of individual households, but often lack information on consumption patterns over longer time horizons (and typically only use a fraction of consumption, such as food expenditures). By contrast, macro estimates make use of consumption variation over time, while neglecting demographic characteristics. Reconciling the differences between micro and macro estimates constitutes an important challenge for future research in this area.

We also investigate whether the literature on habit formation suffers from publication bias. While our data set provides some evidence for publication selection against results inconsistent with the hypothesis of habit formation, we find no resulting publication bias: the effects of the under-reporting of very small and very large estimates cancel each other out, leaving the mean estimate unbiased. Furthermore, we attempt to connect the differences in estimates to differences in the data used, publication characteristics, and estimation methods. Our results suggest that the frequency of the data matters—estimates obtained employing monthly frequency tend to be substantially smaller than when quarterly and annual frequencies are used. The finding is intuitive, since at higher frequencies more consumption goods are likely to display durability. Our results also highlight the importance of estimation methods: we find that ignoring endogeneity yields smaller estimates. Finally, unlike Carroll et al. (2011), we find substantial cross-country heterogeneity in habit formation, with the US and EU displaying stronger habit formation than Japan and other countries.

## References

ABEL, A. B. (1990): "Asset Prices under Habit Formation and Catching Up with the Joneses." *American Economic Review,* 80(2):38–42.

ALESSIE, R. AND F. TEPPA (2010): "Saving and habit formation: evidence from Dutch panel data." *Empirical Economics,* 38(2):385–407.

ANDRÉS, J., J. DAVID LÓPEZ-SALIDO, AND E. NELSON (2009): "Money and the natural rate of interest: Structural estimates for the United States and the euro area." *Journal of Economic Dynamics and Control,* 33(3):758–776.

ASHENFELTER, O., C. HARMON, AND H. OOSTERBEEK (1999): "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias." *Labour Economics,* 6 (4):453–470.

ATTANASIO, O. P. AND G. WEBER (1995): "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy,* 103(6):1121–57.

BANSAL, R., D. KIKU, AND A. YARON (2012): "Risks For the Long Run: Estimation with Time Aggregation." NBER Working Papers 18305, National Bureau of Economic Research, Inc

BJORNLAND, H., K. LEITEMO, AND J. MAIH (2011): "Estimating the natural rates in a simple New Keynesian framework." *Empirical Economics,* 40(3):755–777.

BOUAKEZ, H., E. CARDIA, AND F. J. RUGE-MURCIA (2005): "Habit formation and the persistence of monetary shocks." *Journal of Monetary Economics,* 52(6):1073–1088.

BRODEUR, A., M. LÉ, M. SANGNIER, AND Y. ZYLBERBERG (2013): "Star Wars: The Empirics Strike Back." IZA Discussion Papers 7268, Institute for the Study of Labor (IZA)

CAMPBELL, J. Y. AND J. COCHRANE (1999): "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy,* 107(2):205–251.

CARD, D. AND A. B. KRUEGER (1995): "Time-Series Minimum-Wage Studies: A Meta-analysis." *American Economic Review,* 85(2):238–43.

CARROLL, C. D., J. OVERLAND, AND D. N. WEIL (1997): "Comparison Utility in a Growth Model." *Journal of Economic Growth,* 2(4):339–67.

CARROLL, C. D., J. OVERLAND, AND D. N. WEIL (2000): "Saving and Growth with Habit Formation." *American Economic Review,* 90(3):341–355.

CARROLL, C. D., J. SLACALEK, AND M. SOMMER (2011): "International Evidence on Sticky Consumption Growth." *The Review of Economics and Statistics,* 93(4):1135–1145.

CHETTY, R., A. GUREN, D. MANOLI, AND 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):471–75.

CHOUDHARY, A., P. LEVINE, P. MCADAM, AND P. WELZ (2012): "The happiness puzzle: analytical aspects of the Easterlin paradox." *Oxford Economic Papers,* 64(1):27–42.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy,* 113(1): 1–45.

COLLADO, M. D. AND M. BROWNING (2007): "Habits and heterogeneity in demands: a panel data analysis." *Journal of Applied Econometrics,* 22(3):625–640.

CONSTANTINIDES, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy,* 98(3):519–43.

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

DOUCOULIAGOS, C. AND T. STANLEY (2013): "Are All Economic Facts Greatly Exaggerated? Theory Competition And Selectivity." *Journal of Economic Surveys,* 27(2):316–339.

DOUCOULIAGOS, H. AND M. PALDAM (2011): "The ineffectiveness of development aid on growth: An update." *European Journal of Political Economy,* 27(2):399–404.

DYNAN, K. E. (2000): "Habit Formation in Consumer Preferences: Evidence from Panel Data." *American Economic Review,* 90(3):391–406.

EGGER, M., G. SMITH, M. SCHNEIDER, AND C. MINDER (1997): "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *British Medical Journal,* 316(7129):469–471.

EICHER, T. S., C. PAPAGEORGIOU, AND A. E. RAFTERY (2011): "Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants." *Journal of Applied Econometrics,* 26(1):30–55.

EVERAERT, G. AND L. POZZI (2014): "The Predictability Of Aggregate Consumption Growth In OECD Countries: A Panel Data Analysis." *Journal of Applied Econometrics,* 29(3): 431–453.

FELDKIRCHER, M. AND S. ZEUGNER (2009): "Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging." IMF Working Papers 09/202, International Monetary Fund

FELDKIRCHER, M. AND S. ZEUGNER (2012): "The impact of data revisions on the robustness of growth determinants—A note on Determinants of economic growth: Will data tell?." *Journal of Applied Econometrics,* 27(4):686–694.

FERNANDEZ, C., E. LEY, AND M. F. J. STEEL (2001): "Benchmark priors for Bayesian model averaging." *Journal of Econometrics,* 100(2):381–427.

FUHRER, J. C. (2000): "Habit Formation in Consumption and Its Implications for Monetary-Policy Models." *American Economic Review,* 90(3):367–390.

GRUBER, J. W. (2004): "A present value test of habits and the current account." *Journal of Monetary Economics,* 51(7):1495–1507.

GUARIGLIA, A. (2002): "Consumption, habit formation, and precautionary saving: evidence from the British Household Panel Survey." *Oxford Economic Papers,* 54(1):1–19.

HAVRANEK, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." *Journal of the European Economic Association,* (forthcoming.

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

HAVRANEK, T., R. HORVATH, Z. IRSOVA, AND M. RUSNAK (2015): "Cross-Country Heterogeneity in Intertemporal Substitution." *Journal of International Economics,* 96(1):100–118.

JEFFREYS, H. (1961): *Theory of Probability.* Clarendon Press, Oxford.

KANO, T. (2009): "Habit formation and the present-value model of the current account: Yet another suspect." *Journal of International Economics,* 78(1):72–85.

KASS, R. AND A. RAFTERY (1995): "Bayes Factors." *Journal of the American Statistical Association,* 90(430):773–795.

KOOP, G. (2003): *Bayesian Econometrics.* John Wiley & Sons.

LEY, E. AND M. F. STEEL (2009): "On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression." *Journal of Applied Econometrics,* 24(4):651–674.

MADIGAN, D. AND J. YORK (1995): "Bayesian graphical models for discrete data." *International Statistical Review,* 63(2):215–232.

MORAL-BENITO, E. (2015): "Model Averaging In Economics: An Overview." *Journal of Economic Surveys,* 29(1):46–75.

NECKER, S. (2014): "Scientific misbehavior in economics." *Research Policy,* 43(10):1747–1759.

NICKELL, S. J. (1981): "Biases in Dynamic Models with Fixed Effects." *Econometrica,* 49(6): 1417–26.

RAFTERY, A., D. MADIGAN, AND J. HOETING (1997): "Bayesian Model Averaging for Linear Regression Models." *Journal of the American Statistical Association,* 92(437):179–191.

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

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

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):103–127.

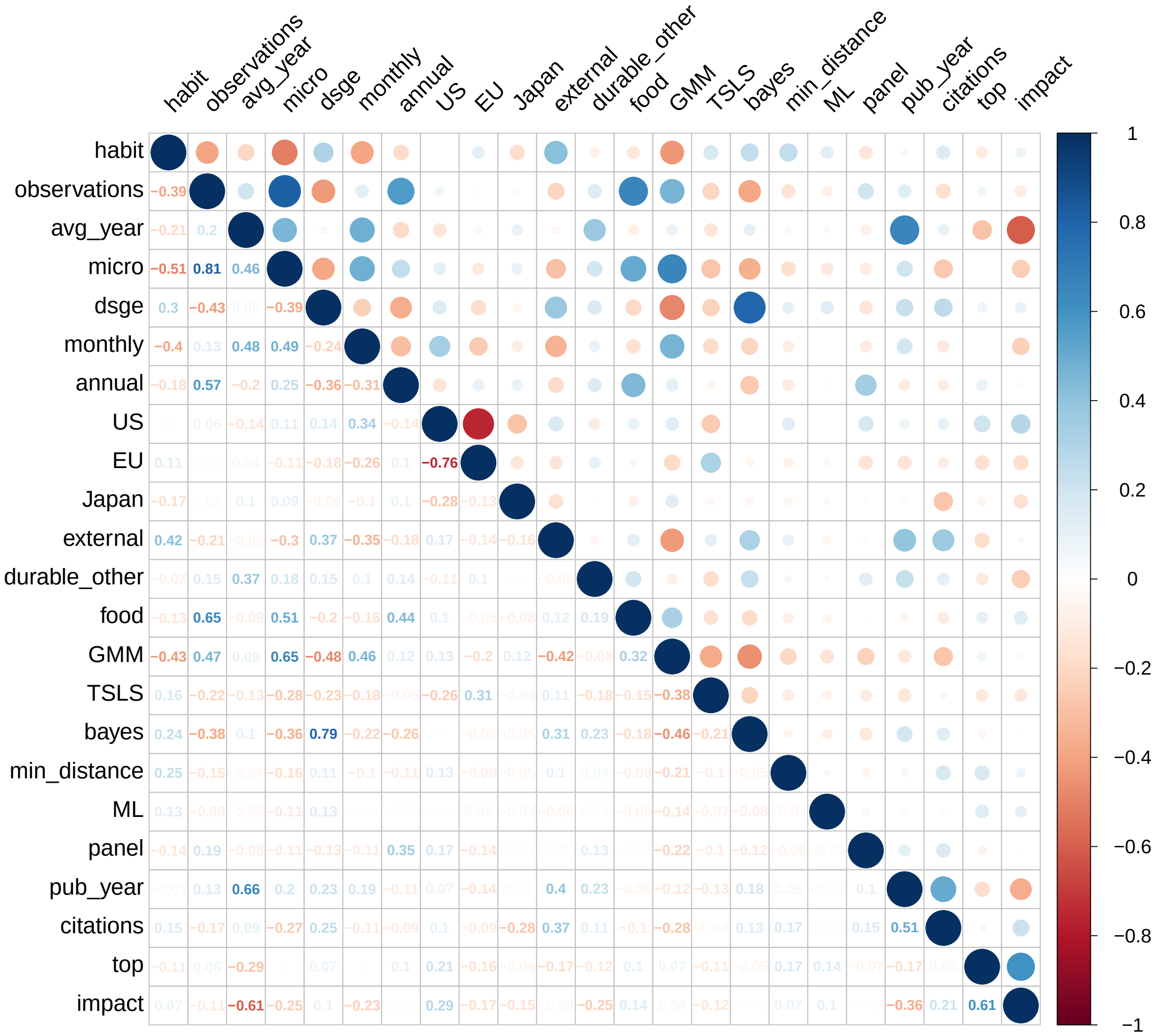
WIELAND, V., T. CWIK, G. J. MÜLLER, S. SCHMIDT, AND M. WOLTERS (2012): "A new comparative approach to macroeconomic modeling and policy analysis." *Journal of Economic Behavior & Organization,* 83(3):523–541.

ZIMMERMANN, S. (2014): "Relative Consumption: The Strength of Internal and External Habits." Working paper, University of Vienna

# Appendix A: Supplementary Statistics and Robustness Checks

## A.1 Correlation of the Variables

***Figure A1: Correlation Matrix***




***Notes:*** A description of the variables is available in Table 4.

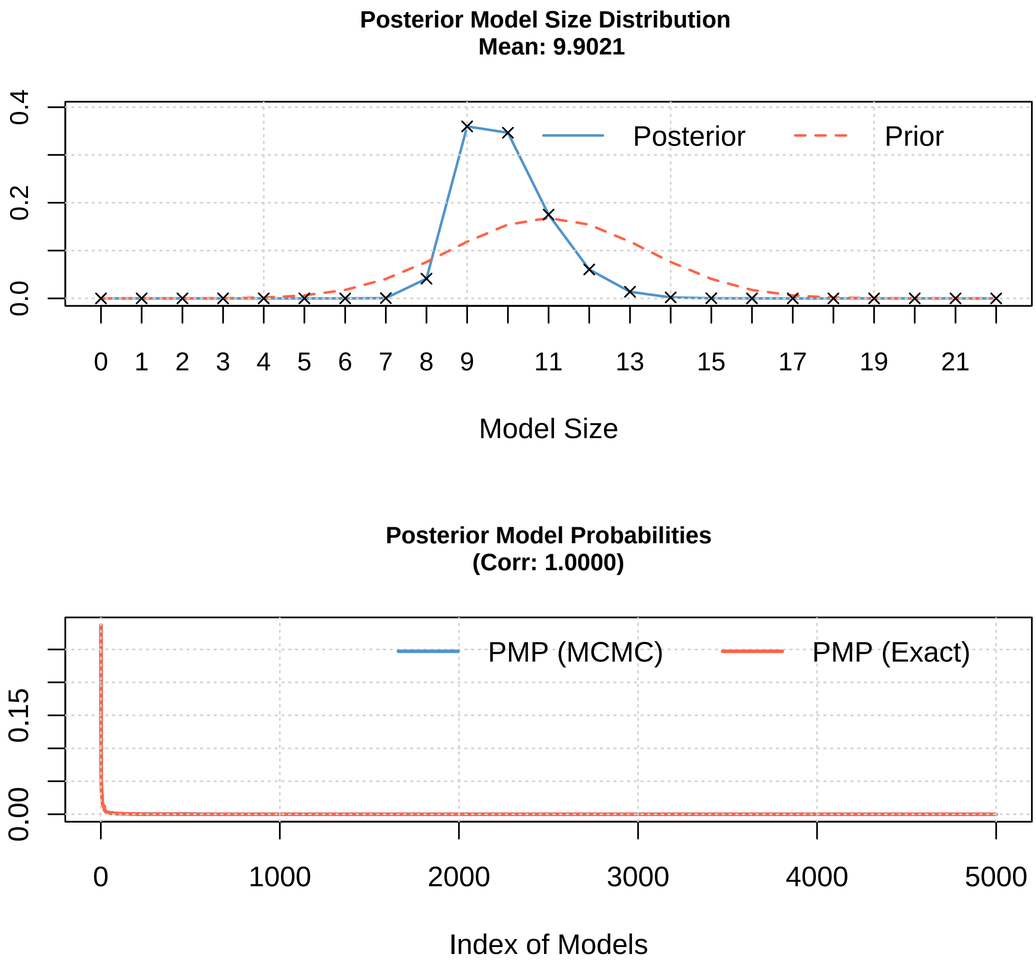
## A.2 Diagnostics of BMA

***Table A1: Summary of BMA Estimation***

| | | | |
|---|---|---|---|
| *Mean no. regressors* 9.9021 | *Draws* $3 \cdot 10^6$ | *Burn-ins* $1 \cdot 10^6$ | *Time* 20.81537 minutes |
| *No. models visited* $460,729$ | *Modelspace* $4.2 \cdot 10^6$ | *Visited* 11% | *Topmodels* 100% |
| *Corr PMP* 1.0000 | *No. Obs.* 567 | *Model Prior* uniform | *g-Prior* UIP |
| *Shrinkage-Stats* Av$=0.9982$ | | | |

***Notes:*** In this specification we employ the priors suggested by Eicher et al. (2011) based on the predictive performance: the uniform model prior (each model has the same prior probability) and the unit information prior (the prior provides the same amount of information as one observation of the data).

***Figure A2: Model Size and Convergence***



## A.3 Robustness Checks

***Table A2: Explaining the Differences in the Estimates of Habit Formation (Frequentist Methods)***

| Response variable: | OLS | | Study fixed effects | |
|---|---|---|---|---|
| Estimate of habit formation | Coef. | Std. er. | Coef. | Std. er. |
| *Data characteristics* | | | | |
| No of obs. | 0.0467 | 0.032 | -0.00558 | 0.019 |
| Average year | 0.00547 | 0.005 | -0.00287 | 0.002 |
| Micro | -0.561*** | 0.169 | | |
| DSGE | -0.0503 | 0.141 | | |
| Monthly | -0.322*** | 0.115 | -0.322*** | 0.044 |
| Annual | -0.0370 | 0.086 | -0.104*** | 0.038 |
| *Countries examined* | | | | |
| US | 0.294* | 0.153 | 0.197 | 0.131 |
| EU | 0.165 | 0.146 | 0.133 | 0.096 |
| Japan | -0.108 | 0.244 | 0.218*** | 0.069 |
| *Variable definition* | | | | |
| External | 0.188*** | 0.068 | 0.313*** | 0.031 |
| Durable | 0.0650 | 0.101 | -0.144 | 0.187 |
| Food | -0.0367 | 0.150 | 0.183 | 0.203 |
| *Estimation approach* | | | | |
| GMM | -0.0593 | 0.082 | 0.00210 | 0.062 |
| TSLS | -0.0744 | 0.092 | 0.0612 | 0.040 |
| Bayes | -0.0804 | 0.105 | -0.0357 | 0.029 |
| Minimum distance | 0.224 | 0.200 | -0.0748*** | 0.002 |
| ML | 0.209 | 0.147 | -0.0805*** | 0.030 |
| Panel | -0.594*** | 0.146 | -0.229* | 0.118 |
| *Publication characteristics* | | | | |
| Publication year | -0.00121 | 0.009 | | |
| Citations | -0.139 | 0.167 | | |
| Top journal | -0.211** | 0.105 | | |
| Impact | -0.00194 | 0.064 | | |
| Constant | -0.0637 | 0.346 | 0.516** | 0.248 |
| Studies | 69 | | 69 | |
| Observations | 567 | | 567 | |

***Notes:*** Standard errors are clustered at the study level. The fixed-effects specification does not include explanatory variables that are constant within studies. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

***Table A3: Explaining the Differences in the Estimates of Habit Formation (Precision Weighting)***

| Response variable: | Bayesian model averaging | | | Frequentist check (OLS) | | |
|---|---|---|---|---|---|---|
| Estimate of habit formation | Post. mean | Post. std. dev. | PIP | Coef. | Std. er. | p-value |
| Precision (1/SE) | 0.761 | 0.108 | 1.000 | 0.822 | 0.017 | 0.000 |
| *Data characteristics* | | | | | | |
| No of obs. | 0.000 | 0.000 | 0.043 | | | |
| Average year | 0.001 | 0.002 | 0.133 | | | |
| Micro | -0.617 | 0.080 | 1.000 | -0.618 | 0.110 | 0.000 |
| DSGE | -0.336 | 0.074 | 1.000 | -0.308 | 0.008 | 0.000 |
| Monthly | -0.500 | 0.114 | 0.986 | -0.535 | 0.108 | 0.000 |
| Annual | 0.001 | 0.016 | 0.044 | | | |
| *Countries examined* | | | | | | |
| US | 0.301 | 0.056 | 1.000 | 0.332 | 0.003 | 0.000 |
| EU | 0.000 | 0.006 | 0.042 | | | |
| Japan | -0.005 | 0.029 | 0.059 | | | |
| *Variable definition* | | | | | | |
| External | 0.060 | 0.109 | 0.296 | | | |
| Durable | 0.004 | 0.033 | 0.086 | | | |
| Food | 0.007 | 0.044 | 0.057 | | | |
| *Estimation approach* | | | | | | |
| GMM | -0.014 | 0.051 | 0.112 | | | |
| TSLS | 0.014 | 0.056 | 0.095 | | | |
| Bayes | 0.129 | 0.055 | 0.885 | 0.156 | 0.012 | 0.000 |
| Minimum distance | 0.004 | 0.039 | 0.045 | | | |
| ML | -0.007 | 0.033 | 0.109 | | | |
| Panel | -0.475 | 0.085 | 1.000 | -0.489 | 0.135 | 0.000 |
| *Publication characteristics* | | | | | | |
| Publication year | 0.001 | 0.003 | 0.135 | | | |
| Citations | 0.004 | 0.036 | 0.090 | | | |
| Top journal | -0.007 | 0.053 | 0.060 | | | |
| Impact | 0.035 | 0.061 | 0.336 | | | |
| Constant | -3.719 | NA | 1.000 | -3.650 | 0.649 | 0.000 |
| Studies | 69 | | | 69 | | |
| Observations | 558 | | | 558 | | |

***Notes:*** PIP = posterior inclusion probability. In the frequentist check we only include method characteristics with PIP $>$ 0.5. Standard errors in the frequentist check are clustered at the study level.

***Table A4: Explaining the Differences in the Estimates of Habit Formation (Internal Habits)***

| Response variable: | Bayesian model averaging | | | Frequentist check (OLS) | | |
|---|---|---|---|---|---|---|
| Estimate of habit formation | Post. mean | Post. std. dev. | PIP | Coef. | Std. er. | p-value |
| *Data characteristics* | | | | | | |
| No of obs. | 0.000 | 0.006 | 0.054 | | | |
| Average year | 0.019 | 0.004 | 1.000 | 0.018 | 0.003 | 0.000 |
| Micro | -0.506 | 0.064 | 1.000 | -0.504 | 0.079 | 0.000 |
| DSGE | 0.021 | 0.060 | 0.163 | | | |
| Monthly | -0.349 | 0.084 | 1.000 | -0.296 | 0.092 | 0.001 |
| Annual | 0.000 | 0.012 | 0.052 | | | |
| *Countries examined* | | | | | | |
| US | 0.096 | 0.126 | 0.482 | | | |
| EU | 0.038 | 0.089 | 0.232 | | | |
| Japan | -0.269 | 0.146 | 0.842 | -0.320 | 0.153 | 0.036 |
| *Variable definition* | | | | | | |
| Durable | -0.028 | 0.056 | 0.259 | | | |
| Food | -0.004 | 0.027 | 0.065 | | | |
| *Estimation approach* | | | | | | |
| GMM | -0.001 | 0.016 | 0.060 | | | |
| TSLS | -0.005 | 0.027 | 0.081 | | | |
| Bayes | 0.000 | 0.020 | 0.056 | | | |
| Minimum distance | 0.005 | 0.035 | 0.063 | | | |
| ML | 0.124 | 0.149 | 0.482 | | | |
| Panel | -0.340 | 0.120 | 0.964 | -0.333 | 0.085 | 0.000 |
| *Publication characteristics* | | | | | | |
| Publication year | -0.017 | 0.007 | 0.919 | -0.018 | 0.008 | 0.019 |
| Citations | -0.045 | 0.100 | 0.222 | | | |
| Top journal | 0.000 | 0.019 | 0.053 | | | |
| Impact | 0.000 | 0.010 | 0.057 | | | |
| Constant | -0.217 | NA | 1.000 | -0.120 | 0.107 | 0.261 |
| Studies | 37 | | | 37 | | |
| Observations | 344 | | | 344 | | |

***Notes:*** PIP = posterior inclusion probability. In the frequentist check we only include method characteristics with PIP > 0.5. Standard errors in the frequentist check are clustered at the study level. Estimates of external habits are excluded from this specification.

***Table A5: Explaining the Differences in the Estimates of Habit Formation (External Habits)***

| Response variable: | Bayesian model averaging | | | Frequentist check (OLS) | | |
|---|---|---|---|---|---|---|
| Estimate of habit formation | Post. mean | Post. std. dev. | PIP | Coef. | Std. er. | p-value |
| *Data characteristics* | | | | | | |
| No of obs. | 0.089 | 0.122 | 0.439 | | | |
| Average year | -0.012 | 0.008 | 0.785 | -0.015 | 0.005 | 0.004 |
| Micro | -1.173 | 0.580 | 0.990 | -0.823 | 0.188 | 0.000 |
| DSGE | -0.052 | 0.081 | 0.359 | | | |
| *Countries examined* | | | | | | |
| US | 0.100 | 0.119 | 0.518 | 0.122 | 0.083 | 0.139 |
| EU | 0.030 | 0.076 | 0.210 | | | |
| Japan | 0.004 | 0.074 | 0.064 | | | |
| *Variable definition* | | | | | | |
| Durable | 0.347 | 0.081 | 1.000 | 0.373 | 0.192 | 0.052 |
| *Estimation approach* | | | | | | |
| GMM | -0.020 | 0.073 | 0.128 | | | |
| TSLS | 0.009 | 0.037 | 0.120 | | | |
| Bayes | -0.133 | 0.135 | 0.580 | -0.248 | 0.138 | 0.072 |
| Minimum distance | 0.261 | 0.138 | 0.863 | 0.227 | 0.272 | 0.404 |
| ML | -0.021 | 0.097 | 0.115 | | | |
| Panel | -0.834 | 0.291 | 0.999 | -0.668 | 0.148 | 0.000 |
| *Publication characteristics* | | | | | | |
| Publication year | 0.010 | 0.016 | 0.330 | | | |
| Citations | -0.464 | 0.249 | 0.898 | -0.394 | 0.247 | 0.111 |
| Top journal | -0.224 | 0.294 | 0.444 | | | |
| Impact | 0.291 | 0.210 | 0.756 | 0.210 | 0.151 | 0.164 |
| Constant | 0.692 | NA | 1.000 | 1.444 | 0.348 | 0.000 |
| Studies | 34 | | | 34 | | |
| Observations | 223 | | | 223 | | |

***Notes:*** PIP = posterior inclusion probability. In the frequentist check we only include method characteristics with PIP $> 0.5$. Standard errors in the frequentist check are clustered at the study level. Estimates of internal habits are excluded from this specification, and some variables are dropped because of collinearity concerns.

***Table A6: Explaining the Differences in the Estimates of Habit Formation (Micro Studies)***

| Response variable: | Bayesian model averaging | | | Frequentist check (OLS) | | |
|---|---|---|---|---|---|---|
| Estimate of habit formation | Post. mean | Post. std. dev. | PIP | Coef. | Std. er. | p-value |
| *Data characteristics* | | | | | | |
| No of obs. | 0.001 | 0.010 | 0.096 | | | |
| Monthly | 0.205 | 0.191 | 0.671 | 0.149 | 0.128 | 0.246 |
| *Countries examined* | | | | | | |
| US | -0.103 | 0.138 | 0.485 | | | |
| EU | 0.005 | 0.029 | 0.103 | | | |
| *Variable definition* | | | | | | |
| External | 0.583 | 0.119 | 1.000 | 0.576 | 0.095 | 0.000 |
| Durable | -0.116 | 0.103 | 0.645 | -0.204 | 0.080 | 0.011 |
| Food | 0.002 | 0.037 | 0.122 | | | |
| *Estimation approach* | | | | | | |
| GMM | 0.077 | 0.114 | 0.393 | | | |
| TSLS | 0.002 | 0.052 | 0.070 | | | |
| *Publication characteristics* | | | | | | |
| Publication year | -0.035 | 0.008 | 1.000 | -0.035 | 0.013 | 0.007 |
| Impact | -0.089 | 0.103 | 0.507 | -0.146 | 0.117 | 0.212 |
| Constant | 0.629 | NA | 1.000 | 0.771 | 0.274 | 0.005 |
| Studies | 11 | | | 11 | | |
| Observations | 190 | | | 190 | | |

***Notes:*** PIP = posterior inclusion probability. In the frequentist check we only include method characteristics with PIP $>$ 0.5. Standard errors in the frequentist check are clustered at the study level. Macro-level estimates of habit formation are excluded from this specification, and some variables are dropped because of collinearity concerns.

## Appendix B: Studies Used in the Meta-Analysis


ADOLFSON, M., S. LASEEN, J. LINDE, & M. VILLANI (2007): "Bayesian estimation of an open economy DSGE model with incomplete pass-through." *Journal of International Economics* **72(2)**: pp. 481–511.

ALESSIE, R. & A. KAPTEYN (1991): "Habit Formation, Interdependent References and Demographic Effects in the Almost Ideal Demand System." *Economic Journal* **101(406)**: pp. 404–19.

ALESSIE, R. & F. TEPPA (2010): "Saving and habit formation: evidence from Dutch panel data." *Empirical Economics* **38(2)**: pp. 385–407.

ALTIG, D., L. CHRISTIANO, M. EICHENBAUM, & J. LINDE (2011): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." *Review of Economic Dynamics* **14(2)**: pp. 225–247.

ANDRÉS, J., J. DAVID LÓPEZ-SALIDO, & E. NELSON (2009): "Money and the natural rate of interest: Structural estimates for the United States and the euro area." *Journal of Economic Dynamics and Control* **33(3)**: pp. 758–776.

AURAY, S. & P. FEVE (2008): "On Sunspots, Habits and Monetary Facts." *Macroeconomic Dynamics* **12(01)**: pp. 72–96.

BALTAGI, B. H., G. BRESSON, J. M. GRIFFIN, & A. PIROTTE (2003): "Homogeneous, heterogeneous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption." *Empirical Economics* **28(4)**: pp. 795–811.

BARTOLOMEO, G. D., L. ROSSI, & M. TANCIONI (2010): "Monetary policy, rule-of-thumb consumers and external habits: a G7 comparison." *Applied Economics* **43(21)**: pp. 2721–2738.

BATINI, N., A. JUSTINIANO, P. LEVINE, & J. PEARLMAN (2006): "Robust inflation-forecast-based rules to shield against indeterminacy." *Journal of Economic Dynamics and Control* **30(9-10)**: pp. 1491–1526.

BEKAERT, G., S. CHO, & A. MORENO (2010): "New Keynesian Macroeconomics and the Term Structure." *Journal of Money, Credit and Banking* **42(1)**: pp. 33–62.

BJORNLAND, H., K. LEITEMO, & J. MAIH (2011): "Estimating the natural rates in a simple New Keynesian framework." *Empirical Economics* **40(3)**: pp. 755–777.

BOIVIN, J. & M. P. GIANNONI (2006): "Has Monetary Policy Become More Effective?" *The Review of Economics and Statistics* **88(3)**: pp. 445–462.

BOUAKEZ, H., E. CARDIA, & F. J. RUGE-MURCIA (2005): "Habit formation and the persistence of monetary shocks." *Journal of Monetary Economics* **52(6)**: pp. 1073–1088.

BOVER, O. (1991): "Relaxing Intertemporal Separability: A Rational Habits Model of Labor Supply Estimated from Panel Data." *Journal of Labor Economics* **9(1)**: pp. 85–100.

BRAUN, P. A., G. M. CONSTANTINIDES, & W. E. FERSON (1993): "Time nonseparability in aggregate consumption: International evidence." *European Economic Review* **37(5)**: pp. 897–920.

CAMPBELL, J. Y. & N. G. MANKIW (1991): "The response of consumption to income: A cross-country investigation." *European Economic Review* **35(4)**: pp. 723–756.

CARROLL, C. D., J. SLACALEK, & M. SOMMER (2011): "International Evidence on Sticky Consumption Growth." *The Review of Economics and Statistics* **93(4)**: pp. 1135–1145.

CASTELNUOVO, E. & S. NISTICO (2010): "Stock market conditions and monetary policy in a DSGE model for the U.S." *Journal of Economic Dynamics and Control* **34(9)**: pp. 1700–1731.

CHIB, S. & S. RAMAMURTHY (2010): "Tailored randomized block MCMC methods with application to DSGE models." *Journal of Econometrics* **155(1)**: pp. 19–38.

CHRISTIANO, L. J., M. EICHENBAUM, & C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* **113(1)**: pp. 1–45.

CHRISTOFFEL, K., K. KUESTER, & T. LINZERT (2009): "The role of labor markets for euro area monetary policy." *European Economic Review* **53(8)**: pp. 908–936.

COLLADO, M. D. & M. BROWNING (2007): "Habits and heterogeneity in demands: a panel data analysis." *Journal of Applied Econometrics* **22(3)**: pp. 625–640.

DEL NEGRO, M. & F. SCHORFHEIDE (2008): "Forming priors for DSGE models (and how it affects the assessment of nominal rigidities)." *Journal of Monetary Economics* **55(7)**: pp. 1191–1208.

DEL NEGRO, M., F. SCHORFHEIDE, F. SMETS, & R. WOUTERS (2007): "On the Fit of New Keynesian Models." *Journal of Business & Economic Statistics* **25**: pp. 123–143.

DENNIS, R. (2009): "Consumption Habits in a New Keynesian Business Cycle Model." *Journal of Money, Credit and Banking* **41(5)**: pp. 1015–1030.

DYNAN, K. E. (2000): "Habit Formation in Consumer Preferences: Evidence from Panel Data." *American Economic Review* **90(3)**: pp. 391–406.

EDGE, R. M., M. T. KILEY, & J.-P. LAFORTE (2008): "Natural rate measures in an estimated DSGE model of the U.S. economy." *Journal of Economic Dynamics and Control* **32(8)**: pp. 2512–2535.

EVERAERT, G. & L. POZZI (2014): "The Predictability Of Aggregate Consumption Growth In OECD Countries: A Panel Data Analysis." *Journal of Applied Econometrics* **29(3)**: pp. 431–453.

FERNÁNDEZ-VILLAVERDE, J. (2010): "The econometrics of DSGE models." *SERIEs - Journal of the Spanish Economic Association* **1(1)**: pp. 3–49.

FERSON, W. E. & G. M. CONSTANTINIDES (1991): "Habit persistence and durability in aggregate consumption: Empirical tests." *Journal of Financial Economics* **29(2)**: pp. 199–240.

FUHRER, J. C. (2000): "Habit Formation in Consumption and Its Implications for Monetary-Policy Models." *American Economic Review* **90(3)**: pp. 367–390.

FUSARO, M. A. & D. H. DUTKOWSKY (2011): "What explains consumption in the very short-run? Evidence from checking account data." *Journal of Macroeconomics* **33(4)**: pp. 542–552.

GRUBER, J. W. (2004): "A present value test of habits and the current account." *Journal of Monetary Economics* **51(7)**: pp. 1495–1507.

GUARIGLIA, A. (2002): "Consumption, habit formation, and precautionary saving: evidence from the British Household Panel Survey." *Oxford Economic Papers* **54(1)**: pp. 1–19.

GUERRON-QUINTANA, P. A. (2008): "Refinements on macroeconomic modeling: The role of non-separability and heterogeneous labor supply." *Journal of Economic Dynamics and Control* **32(11)**: pp. 3613–3630.

GUERRON-QUINTANA, P. A. (2010): "What you match does matter: the effects of data on DSGE estimation." *Journal of Applied Econometrics* **25(5)**: pp. 774–804.

HEATON, J. (1993): "The Interaction between Time-Nonseparable Preferences and Time Aggregation." *Econometrica* **61(2)**: pp. 353–385.

HIROSE, Y. & S. NAGANUMA (2010): "Structural Estimation Of The Output Gap: A Bayesian Dsge Approach." *Economic Inquiry* **48(4)**: pp. 864–879.

IACOVIELLO, M. (2004): "Consumption, house prices, and collateral constraints: a structural econometric analysis." *Journal of Housing Economics* **13(4)**: pp. 304–320.

IACOVIELLO, M. & S. NERI (2010): "Housing Market Spillovers: Evidence from an Estimated DSGE Model." *American Economic Journal: Macroeconomics* **2(2)**: pp. 125–64.

IWAMOTO, K. (2013): "Habit formation in household consumption: evidence from Japanese panel data." *Economics Bulletin* **33(1)**: pp. 323–333.

JUSTINIANO, A. & B. PRESTON (2010): "Can structural small open-economy models account for the influence of foreign disturbances?" *Journal of International Economics* **81(1)**: pp. 61–74.

KANO, T. (2009): "Habit formation and the present-value model of the current account: Yet another suspect." *Journal of International Economics* **78(1)**: pp. 72–85.

KILEY, M. T. (2010): "Habit Persistence, Nonseparability between Consumption and Leisure, or Rule-of-Thumb Consumers: Which Accounts for the Predictability of Consumption Growth?" *The Review of Economics and Statistics* **92(3)**: pp. 679–683.

KORNIOTIS, G. M. (2010): "Estimating Panel Models With Internal and External Habit Formation." *Journal of Business & Economic Statistics* **28(1)**: pp. 145–158.

LAFORTE, J.-P. (2007): "Pricing Models: A Bayesian DSGE Approach for the U.S. Economy." *Journal of Money, Credit and Banking* **39(s1)**: pp. 127–154.

LEVIN, A. T., A. ONATSKI, J. WILLIAMS, & N. M. WILLIAMS (2005): "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models." In "NBER Macroeconomics Annual 2005, Volume 20," NBER Chapters, pp. 229–312. National Bureau of Economic Research, Inc.

LEVINE, P., J. PEARLMAN, G. PERENDIA, & B. YANG (2012): "Endogenous Persistence in an estimated DSGE Model Under Imperfect Information." *Economic Journal* **122(565)**: pp. 1287–1312.

LUBIK, T. A. & F. SCHORFHEIDE (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *American Economic Review* **94(1)**: pp. 190–217.

MATHESON, T. (2010): "Assessing the fit of small open economy DSGEs." *Journal of Macroeconomics* **32(3)**: pp. 906–920.

MAURER, J. & A. MEIER (2008): "Smooth it Like the 'Joneses'? Estimating Peer-Group Effects in Intertemporal Consumption Choice." *Economic Journal* **118(527)**: pp. 454–476.

MEHRA, Y. P. & E. W. MARTIN (2003): "Why does consumer sentiment predict household spending?" *Economic Quarterly* **89(4)**: pp. 51–67.

MERTENS, K. & M. O. RAVN (2011): "Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks." *Review of Economic Dynamics* **14(1)**: pp. 27–54.

MILANI, F. (2007): "Expectations, learning and macroeconomic persistence." *Journal of Monetary Economics* **54(7)**: pp. 2065–2082.

NAIK, N. Y. & M. J. MOORE (1996): "Habit Formation and Intertemporal Substitution in Individual Food Consumption." *The Review of Economics and Statistics* **78(2)**: pp. 321–28.

PAGANO, P. (2004): "Habit persistence and the marginal propensity to consume in Japan." *Journal of the Japanese and International Economies* **18(3)**: pp. 316–329.

RABANAL, P. (2007): "Does inflation increase after a monetary policy tightening? Answers based on an estimated DSGE model." *Journal of Economic Dynamics and Control* **31(3)**: pp. 906–937.

RAVN, M. O., S. SCHMITT-GROHE, M. URIBE, & L. UUSKULA (2010): "Deep habits and the dynamic effects of monetary policy shocks." *Journal of the Japanese and International Economies* **24(2)**: pp. 236–258.

RHEE, W. (2004): "Habit formation and precautionary saving: evidence from the Korean household panel studies." *Journal of Economic Development* **29(2)**: pp. 1–19.

SAHUC, J.-G. & F. SMETS (2008): "Differences in Interest Rate Policy at the ECB and the Fed: An Investigation with a Medium-Scale DSGE Model." *Journal of Money, Credit and Banking* **40(2-3)**: pp. 505–521.

SCHMITT-GROHÉ, S. & M. URIBE (2012): "What's News in Business Cycles." *Econometrica* **80(6)**: pp. 2733–2764.

SLANICAY, M. & O. VAŠÍČEK (2011): "Habit Formation, Price Indexation and Wage Indexation in the DSGE Model: Specification, Estimation and Model Fit." *Review of Economic Perspectives* **11(2)**: pp. 71–91.

SMETS, F. & R. WOUTERS (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* **1(5)**: pp. 1123–1175.

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

SOMMER, M. (2007): "Habit Formation and Aggregate Consumption Dynamics." *The B.E. Journal of Macroeconomics* **7(1)**: pp. 1–25.

STOCK, J. H. & J. WRIGHT (2000): "GMM with Weak Identification." *Econometrica* **68(5)**: pp. 1055–1096.

SUGO, T. & K. UEDA (2008): "Estimating a dynamic stochastic general equilibrium model for Japan." *Journal of the Japanese and International Economies* **22(4)**: pp. 476–502.

TRIGARI, A. (2009): "Equilibrium Unemployment, Job Flows, and Inflation Dynamics." *Journal of Money, Credit and Banking* **41(1)**: pp. 1–33.

WOUTERS, R. & F. SMETS (2005): "Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE Approach." *Journal of Applied Econometrics* **20(2)**: pp. 161–183.

| CNB WORKING PAPER SERIES | | |
|---|---|---|
| 3/2015 | Tomáš Havránek<br>Marek Rusnák<br>Anna Sokolova | *Habit formation in consumption: A meta-analysis* |
| 2/2015 | Tomáš Havránek<br>Diana Žigraiová | *Bank competition and financial stability: Much ado about nothing?* |
| 1/2015 | Tomáš Havránek<br>Zuzana Iršová | *Do borders really slash trade? A meta-analysis* |
| 16/2014 | Mark Joy<br>Marek Rusnák<br>Kateřina Šmídková<br>Bořek Vašíček | *Banking and currency crises: Differential diagnostics for developed countries* |
| 15/2014 | Oxana Babecká Kucharčuková<br>Peter Claeys<br>Bořek Vašíček | *Spillover of the ECB's monetary policy outside the Euro Area: How different is conventional from unconventional policy?* |
| 14/2014 | Branislav Saxa | *Forecasting mortgages: Internet search data as a proxy for mortgage credit demand* |
| 13/2014 | Jan Filáček<br>Jakub Matějů | *Adverse effects of monetary policy signalling* |
| 12/2014 | Jan Brůha<br>Jiří Polanský | *The housing sector over business cycles: Empirical analysis and DSGE modelling* |
| 11/2014 | Tomáš Adam<br>Miroslav Plašil | *The impact of financial variables on Czech macroeconomic developments: An empirical investigation* |
| 10/2014 | Kamil Galuščák<br>Gábor Kátay | *Labour force participation and tax-benefit systems: A cross-country comparative perspective* |
| 9/2014 | Jaromír Tonner<br>Jan Brůha | *The Czech housing market through the lens of a DSGE model containing collateral-constrained households* |
| 8/2014 | Michal Franta<br>David Havrlant<br>Marek Rusnák | *Forecasting Czech GDP using mixed-frequency data models* |
| 7/2014 | Tomáš Adam<br>Soňa Benecká<br>Jakub Matějů | *Risk aversion, financial stress and their non-linear impact on exchange rates* |
| 6/2014 | Tomáš Havránek<br>Roman Horváth<br>Zuzana Iršová<br>Marek Rusnák | *Cross-country heterogeneity in intertemporal substitution* |
| 5/2014 | Ruslan Aliyev<br>Dana Hájková<br>Ivana Kubicová | *The impact of monetary policy on financing of Czech firms* |
| 4/2014 | Jakub Matějů | *Explaining the strength and efficiency of monetary policy transmission: A panel of impulse responses from a Time-varying Parameter Model* |
| 3/2014 | Soňa Benecká<br>Luboš Komárek | *International reserves: Facing model uncertainty* |
| 2/2014 | Kamil Galuščák<br>Petr Hlaváč<br>Petr Jakubík | *Stress testing the private household sector using microdata* |

| | | |
|---|---|---|
| 1/2014 | Martin Pospíšil<br>Jiří Schwarz | *Bankruptcy, investment, and financial constraints: Evidence from a post-transition economy* |

## CNB RESEARCH AND POLICY NOTES

| | | |
|---|---|---|
| 4/2014 | Josef Brechler<br>Václav Hausenblas<br>Zlatuše Komárková<br>Miroslav Plašil | *Similarity and clustering of banks: Application to the credit exposures of the Czech banking sector* |
| 3/2014 | Michal Franta<br>Tomáš Holub<br>Petr Král<br>Ivana Kubicová<br>Kateřina Šmídková<br>Bořek Vašíček | *The exchange rate as an instrument at zero interest rates: The case of the Czech Republic* |
| 2/2014 | František Brázdik<br>Zuzana Humplová<br>František Kopřiva | *Evaluating a structural model forecast: Decomposition approach* |
| 1/2014 | Michal Skořepa<br>Jakub Seidler | *Capital buffers based on banks' domestic systemic importance: Selected issues* |

## CNB ECONOMIC RESEARCH BULLETIN

| | |
|---|---|
| May 2015 | *Forecasting* |
| November 2014 | *Macroprudential research: Selected issues* |
| April 2014 | *Stress-testing analyses of the Czech financial system* |

Czech National Bank
Economic Research Department
Na Příkopě 28, 115 03 Praha 1
Czech Republic
phone: +420 2 244 12 321
fax: +420 2 244 14 278
http://www.cnb.cz
e-mail: research@cnb.cz
ISSN 1803-7070