A Note on Comparing Deep and Aggregate Habit Formation in an Estimated New-Keynesian Model

Gregory E. Givens[a,*]

[a] *Department of Economics, Finance, and Legal Studies, University of Alabama, Tuscaloosa, AL 35487, USA*

This Draft: September 2013

Abstract

Habit formation is a fixture of contemporary new-Keynesian models. The vast majority assume that agents form habits strictly over consumption of an aggregate good, leaving open the question of whether it might be preferable to have them form habits over differentiated products instead–an arrangement known as deep habits. I answer this question by estimating a model that nests both habit concepts as special cases. Estimates reveal that the data favor a specification in which consumption habits are stronger at the product level than at the aggregate level. A mix of significance tests and simulation results indicate that including deep habits greatly improves model fit, most notably with regard to inflation dynamics.

Keywords: Deep Habits, Nominal Rigidities, Inflation Persistence

JEL Classification: E31, E32

*Corresponding author. Tel.: + 205 348 8961.
E-mail address: gegivens@cba.ua.edu (G.E. Givens).

1 Introduction

Habit formation in consumption is a prominent feature of modern new-Keynesian business cycle models (e.g., Fuhrer, 2000; Christiano, Eichenbaum, and Evans, 2005; Smets and Wouters, 2007). A standard assumption in this class of models is that households derive utility by consuming an aggregate good that is comprised of numerous differentiated products. A basic question then is whether habits develop at the level of the aggregate good or at the level of individual good varieties. Until recently this literature has only considered cases where agents become addicted to the overall consumption bundle despite evidence suggesting that shoppers form habits over product categories and even specific brands (e.g., Chintagunta, Kyriazidou, and Perktold, 2001). Motivated by these findings, Ravn, Schmitt-Grohé, and Uribe (2006) propose a "deep habits" model in which habitual consumption develops exclusively on a good-by-good basis. In such an environment the demand function for each product will depend on past sales, causing equilibrium mark-ups of price over marginal cost to be time-varying and to move countercyclically with output. The authors go on to show that by inducing *countercyclical* mark-up behavior, deep habits can account for the observed *procyclical* responses of both consumption and real wages to various demand shocks. In contrast, a traditional habit-persistence model fails to capture this dynamic.

The comparisons between deep and aggregate habit formation made in Ravn *et al.* (2006) take place in an economy with purely flexible product prices. As a result, the model is limited in its ability to match the correlations among nominal and real variables that define postwar US business cycles. New-Keynesian models, on the other hand, are better suited to this task (e.g., Ireland, 2003). A relevant question then is whether incorporating deep habits into these models could improve their fit with the data when compared to standard versions that assume habits thrive only at the composite good level. I try to answer that question here by estimating a small-scale DSGE model with sticky product prices, which I then use as a laboratory for testing the empirical implications of the two habit concepts described above.

The paper begins by presenting a model whose preference structure nests both habit types as special cases. The advantage of employing a nested utility function is that it enables one to consider cases in which only the composite good is habit forming, only individual goods are habit forming, both are addictive with possibly different habit intensities, or neither. Accommodating all of these arrangements during the course of estimation makes it easier to infer from the data which mode of habit persistence is the more empirically compelling.

Three versions of the model are estimated using maximum likelihood. One leaves the utility parameters unconstrained, allowing the data to ascertain the strength of habit formation at both levels. The second permits deep habit formation but sets aggregate habit intensity equal to zero. The third restricts deep habit intensity to zero while leaving the aggregate parameter free. Likelihood ratio tests are used to compare the fit of the constrained models to an unrestricted alternative that allows both habit types to coexist. Estimates reveal that the data favor a model with substantial persistence at the product level along with a modest amount at the composite good level. When examined side-by-side, however, it is clear that deep habits are better at explaining the broad correlations embodied by log likelihood. Indeed, exclusion restrictions on aggregate habits are not easily rejected at standard significance levels but are rejected with high confidence when imposed on deep habits.

Likelihood comparisons, while useful for assessing fit, are uninformative about the specific features of the data best captured by deep habits. Further complicating the analysis is the fact that estimates change from one model to the next, making it difficult to distinguish the effects of habit formation from the effects of changes in parameter values. To isolate the role of the habit mechanism from these other elements, I compare simulations of the restricted model containing only deep habits to those from an identically-parameterized model with deep habits replaced by aggregate habits. Simulations reveal that deep habits are superior because they impart greater persistence on inflation. Correlations between current and lagged inflation, for example, decay more slowly when habits are deeply rooted and are closer to the sample correlations. This dynamic is also reflected in the impulse response functions, which show inflation reacting more sluggishly to demand and supply shocks.

This paper adds to a literature that incorporates deep habits into sticky-price models of the business cycle. Using GMM, Lubik and Teo (2013) estimate a new-Keynesian Phillips curve derived from a model with deep habit preferences. One drawback of single-equation estimation, however, is that it does not account for all of the general equilibrium restrictions on the joint dynamics of the endogenous variables. By contrast, maximum-likelihood imposes all of them by estimating simultaneously the full system of equilibrium difference equations. In comparing deep and aggregate habit formation, a systems-based approach is useful because the two specifications have different testable implications for the co-movement of the endogenous variables. Despite these methodological differences, it is encouraging that the studies report common findings, notably regarding estimates of the deep habit parameter, improvements in fit, and lagged indexation as a trivial source of inflation persistence.

Ravn, Schmitt-Grohé, Uribe, and Uuskula (2010) utilize a complete model for estimation. A key difference from this paper is that parameters are chosen by minimizing a weighted discrepancy between the responses to a monetary shock implied by the model and the empirical ones taken from a VAR. Statistical inference is thus based on limited information contained in the response functions rather than full information provided by the likelihood function.

2 The Model

The economy is inhabited by identical households and a continuum of imperfectly competitive firms that produce differentiated goods and face costs of changing prices.

2.1 Households

Households are indexed by $j \in [0,1]$. Each household j consumes differentiated goods $c_{j,t}(i)$, with varieties indexed by $i \in [0,1]$, and supplies labor $h_{j,t}$. Following Ravn *et al.* (2006), preferences feature *external* habit formation on a good-by-good basis. This so-called "deep habits" specification assumes that period utility depends on a composite good $x_{j,t}$ given by

$$x_{j,t} = \left[\int_0^1 \left(c_{j,t}(i) - b^d c_{t-1}(i) \right)^{1-1/\eta} di \right]^{1/(1-1/\eta)}, \tag{1}$$

where $c_{t-1}(i) \equiv \int_0^1 c_{j,t-1}(i)dj$ denotes the population mean consumption of good i at date $t-1$ and $\eta > 1$ is the substitution elasticity across (habit-adjusted) varieties. Parameter $b^d \in [0,1)$ measures the intensity of habit formation in consumption of each variety.

At the start of date t, household j minimizes $\int_0^1 P_t(i)c_{j,t}(i)di$ subject to the aggregation constraint (1). $P_t(i)$ is the nominal price of good i. In minimizing consumption costs, the household takes $c_{t-1}(i)$ as given. First-order conditions imply demand functions of the form

$$c_{j,t}(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\eta} x_{j,t} + b^d c_{t-1}(i),$$

where $P_t \equiv \left[\int_0^1 P_t(i)^{1-\eta} di \right]^{1/(1-\eta)}$ is the unit price of the final good. Demand for good i has the property of being negatively related to the relative price $P_t(i)/P_t$ and positively related to consumption of the final good $x_{j,t}$. Note, however, that deep habits give rise to an additional component that depends positively on past aggregate sales $c_{t-1}(i)$ when $b^d > 0$.

Household j maximizes the expected lifetime utility function

$$V_{j,0} = E_0 \sum_{t=0}^{\infty} \beta^t a_t \left[\log(x_{j,t} - b^a x_{t-1}) - \frac{h_{j,t}^{1+\chi}}{1+\chi} \right], \tag{2}$$

where E_0 is a date-0 conditional expectations operator, $\beta \in (0,1)$ is a subjective discount factor, and $1/\chi > 0$ is the Frisch elasticity of labor supply. Preference shocks a_t follow the autoregressive process $\log a_t = \rho_a \log a_{t-1} + \varepsilon_{a,t}$, with $\rho_a \in (-1,1)$ and $\varepsilon_{a,t} \sim$ i.i.d. $N(0, \sigma_a^2)$.

The preferences in (2) permit households to form habits directly over the composite good $x_{j,t}$. Here household j values quasi differences between $x_{j,t}$ and $x_{t-1} \equiv \int_0^1 x_{j,t-1} dj$, the population average consumption of the composite good at date $t-1$. Following Abel (1990), x_{t-1} is viewed as an *external* reference in that its evolution is taken as exogenous. Parameter $b^a \in [0,1)$ measures the strength of external habits in consumption of the final good.[1]

The appeal of (2) is that it nests both forms of habit persistence as special cases. When $b^a = 0$ preferences become time separable in $x_{j,t}$, and the model collapses to a strictly deep habits specification. Alternatively, setting $b^d = 0$ causes the deep habits mechanism in (1) to vanish. In this case habits develop only at the level of the final good.[2] Henceforth, I refer to this restricted version as the "aggregate habits" specification.

In each period $t \geq 0$, household j supplies labor to firms at a competitive nominal wage rate W_t. It also has access to riskless one-period bonds $B_{j,t}$ that pay a gross nominal interest rate R_t at date $t+1$. Together with bond wealth and labor income, household j receives an aliquot share of profits from ownership of firms, $\Phi_{j,t}$. The flow budget constraint is

$$P_t x_{j,t} + \varpi_t + B_{j,t} \leq R_{t-1} B_{j,t-1} + W_t h_{j,t} + \Phi_{j,t}, \tag{3}$$

where $\varpi_t \equiv b^d \int_0^1 P_t(i) c_{t-1}(i) di$.[3] Sequences $\{x_{j,t}, h_{j,t}, B_{j,t}\}_{t=0}^{\infty}$ are chosen to maximize $V_{j,0}$ subject to (3) and a borrowing limit, taking as given $\{P_t, \varpi_t, R_t, W_t, \Phi_{j,t}\}_{t=0}^{\infty}$ as well as the initial composite good x_{-1} and bond holdings $B_{j,-1}$. The first-order conditions imply

$$h_{j,t}^{\chi} \left(x_{j,t} - b^a x_{t-1} \right) = w_t \tag{4}$$

and

$$1 = \beta E_t \frac{R_t}{\pi_{t+1}} \frac{a_{t+1} \left(x_{j,t} - b^a x_{t-1} \right)}{a_t \left(x_{j,t+1} - b^a x_t \right)}, \tag{5}$$

where $w_t \equiv W_t / P_t$ is the real wage and $\pi_t \equiv P_t / P_{t-1}$ is the gross inflation rate.

2.2 Firms

Consumption goods are produced by monopolistically competitive firms. Variety i is manufactured from labor according to $y_t(i) = z_t h_t(i)$, where $y_t(i)$ is the output of firm i and $h_t(i)$ is its labor input. Aggregate technology shocks z_t follow an autoregressive process

$\log z_t = (1-\rho_z)\log z + \rho_z \log z_{t-1} + \varepsilon_{z,t}$, with $\rho_z \in [0,1)$, $z > 0$, and $\varepsilon_{z,t} \sim$ i.i.d. $N(0, \sigma_z^2)$.

Firm i selects its price $P_t(i)$ to maximize the present discounted value of nominal profits. Constraining the price-setting decision is the market demand curve

$$c_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\eta} x_t + b^d c_{t-1}(i), \tag{6}$$

obtained by integrating $c_{j,t}(i)$ over all $j \in [0,1]$ households. It is understood that firm i will meet this demand at the posted price, implying $z_t h_t(i) \geq c_t(i)$ for all $t \geq 0$. Following Rotemberg (1982), firms also face costs of adjusting prices of the form $(\alpha/2)\left(P_t(i)/\pi P_{t-1}(i) - 1\right)^2 y_t$. These are expressed in units of aggregate output, $y_t \equiv \int_0^1 y_t(i)di$, and are incurred whenever growth in $P_t(i)$ deviates from the long-run mean inflation rate π.

The Lagrangian of firm i's problem can be written as

$$\begin{aligned}\mathcal{L} = E_0 \sum_{t=0}^{\infty} q_{0,t} \Bigg\{ & P_t(i)c_t(i) - W_t h_t(i) - P_t \frac{\alpha}{2}\left(\frac{P_t(i)}{\pi P_{t-1}(i)} - 1\right)^2 y_t \\ & + P_t \gamma_t(i)\left[z_t h_t(i) - c_t(i)\right] + P_t \nu_t(i) \left[\left(\frac{P_t(i)}{P_t}\right)^{-\eta} x_t + b^d c_{t-1}(i) - c_t(i)\right] \Bigg\},\end{aligned}$$

where $q_{0,t}$ is a stochastic discount factor.[4] In maximizing $\mathcal{L}$, firm i takes $c_{-1}(i)$, $P_{-1}(i)$, and $\{q_{0,t}, W_t, P_t, y_t, z_t, x_t\}_{t=0}^{\infty}$ as given. First-order conditions for $\{h_t(i), c_t(i), P_t(i)\}_{t=0}^{\infty}$ are

$$w_t = \gamma_t(i) z_t, \tag{7}$$

$$\nu_t(i) = \frac{P_t(i)}{P_t} - \gamma_t(i) + b^d E_t \frac{q_{0,t+1}}{q_{0,t}} \pi_{t+1} \nu_{t+1}(i), \tag{8}$$

$$\begin{aligned} c_t(i) = \nu_t(i)\eta \left(\frac{P_t(i)}{P_t}\right)^{-\eta-1} x_t + \alpha \left(\frac{P_t(i)}{\pi P_{t-1}(i)} - 1\right) \frac{P_t y_t}{\pi P_{t-1}(i)} \\ - \alpha E_t \frac{q_{0,t+1}}{q_{0,t}} \pi_{t+1} \left(\frac{P_{t+1}(i)}{\pi P_t(i)} - 1\right) \frac{P_{t+1}(i) P_t y_{t+1}}{\pi P_t(i)^2}. \end{aligned} \tag{9}$$

The multiplier $\gamma_t(i)$ in (7) corresponds to real marginal cost. The multiplier $\nu_t(i)$ in (8) is the shadow value of selling an extra unit of good i in the current period. It is the sum of two parts: the short-run profit from the sale, $P_t(i)/P_t - \gamma_t(i)$, and the discounted value of all future profits that the sale is expected to generate, $b^d E_t \frac{q_{0,t+1}}{q_{0,t}} \pi_{t+1} \nu_{t+1}(i)$.[5] Eq. (9) describes conditions that are satisfied by the firm's optimal price.

2.3 Monetary Policy

Following Ravn *et al.* (2010), the monetary authority sets R_t according to a Taylor rule

$$\log(R_t/R) = \theta_r \log(R_{t-1}/R) + (1-\theta_r)\left[\theta_\pi \log(\pi_t/\pi) + \theta_y \log(y_t/y)\right] + \varepsilon_{r,t},$$

where θ_π and θ_y capture the long-run policy response to fluctuations in gross inflation and aggregate output. Parameter $\theta_r \in [0,1)$ measures the degree of interest rate smoothing.[6] Positive constants R, π, and y denote the steady-state values of the nominal interest rate, inflation, and output. The purely random element of policy is summarized by $\varepsilon_{r,t} \sim$ i.i.d. $N(0, \sigma_r^2)$.

2.4 Competitive Equilibrium

I consider a *symmetric* competitive equilibrium in which households make identical consumption and labor decisions and all firms charge the same price.[7]

Equilibrium requires that both labor and product markets clear at prevailing prices. This is accomplished in the labor market by imposing

$$\int_0^1 h_{j,t} dj = \int_0^1 h_t(i) di \equiv h_t$$

for $t \geq 0$. In product markets, output of the final good must be allocated to total consumption expenditure and to resource costs originating from the adjustment of prices:

$$y_t = c_t + \frac{\alpha}{2}\left(\frac{\pi_t}{\pi} - 1\right)^2 y_t.$$

3 Econometric Strategy

I log-linearize the equilibrium conditions around the deterministic steady state of the model and compute a rational expectations equilibrium using methods developed by Klein (2000). The solution has a state-space representation given by

$$\mathbf{s}_t = \mathbf{\Pi}\mathbf{s}_{t-1} + \mathbf{\Omega}\boldsymbol{\varepsilon}_t, \tag{10}$$

$$\mathbf{f}_t = \mathbf{U}\mathbf{s}_t, \tag{11}$$

where $\mathbf{s}_t$ contains exogenous and endogenous state variables, $\boldsymbol{\varepsilon}_t$ holds the guassian innovations, and $\mathbf{f}_t$ contains the forward-looking variables. Elements of $\mathbf{\Pi}$, $\mathbf{\Omega}$, and $\mathbf{U}$ are functions of the structural parameters governing preferences and technologies.

Models of the form (10) and (11) can be estimated via maximum likelihood using the Kalman filter (e.g., Harvey, 1989). With data on the observables, the filter compiles a history $\{\boldsymbol{\varepsilon}_t\}_{t=1}^{T}$ necessary for calculating sample log likelihood. Since these innovations depend on $\mathbf{\Pi}$, $\mathbf{\Omega}$, and $\mathbf{U}$, structural parameters may be estimated by maximizing the likelihood function.

I estimate the model with US data on consumption, inflation, and a nominal interest rate. The sample is 1965:Q3 to 2012:Q1. Consumption is real personal consumption expenditures (PCE) divided by the noninstitutional population. Inflation is the first-differenced log of the PCE price index. The interest rate is the log of the gross yield on 3-month Treasury bills.[8]

4 Estimation Results

Three parameters are held fixed prior to estimation. The discount factor β is set to 0.9965, which equals the ratio of the sample means of inflation and the nominal interest rate. The substitution elasticity η is set equal to 6. In the absence of deep habits ($b^d = 0$), this value implies an average mark-up of 20 percent and is consistent with Basu and Fernald (1997). When deep habits are present, the mark-up is given by $\mu \equiv \left[1 - (1/\eta)\left(\frac{1-\beta b^d}{1-b^d}\right)\right]^{-1}$ and depends on both η and b^d. Estimates of b^d discussed below put μ in the 21-22 percent range.

Attempts to estimate α returned values that point to extreme levels of price rigidity. As a result, I follow Monacelli (2009) by calibrating α so that the model is consistent with a price-change frequency of one year in a Calvo-Yun framework. Letting $1-\phi$ denote the Calvo reset probability, $\phi = 0.75$ implies an average contract duration of $(1-\phi)^{-1} = 4$ quarters (e.g., Woodford, 2003). To obtain α, I set the key slope coefficient in the linearized version of (9), $(\eta-1)/\alpha$, equal to the Phillips curve slope in the Calvo-Yun model given by $(1-\phi)(1-\beta\phi)/\phi$. This restriction implies that adjustment costs satisfy $\alpha = \phi(\eta-1)/(1-\phi)(1-\beta\phi)$.[9]

Table 1 reports maximum-likelihood estimates and standard errors for the nested habits model and the two restricted models that consider deep and aggregate habits separately.[10]

Looking at the nested model, estimates of b^a and b^d reveal that the data favor a specification in which habits are strongest at the product level. The estimate of b^d is 0.94 and is close to the value of 0.85 reported by Lubik and Teo (2013) and Ravn *et al.* (2010). The estimate of b^a is only 0.61, which is near the range in Christiano *et al.* (2005) (0.65) and Smets and Wouters (2007) (0.71). The standard error for b^d is also an order of magnitude smaller than the one for b^a, indicating far more precision in the estimate of deep habits.[11]

Concerning the Taylor rule, the estimate of θ_r is 0.90, reflecting the Federal Reserve's penchant for adjusting the interest rate gradually in response to shocks. The estimate of θ_π is 1.54, ensuring that policy is stabilizing and satisfies the Taylor principle over the sample period. By contrast, policy does not appear to have reacted strongly to output fluctuations. The estimate of θ_y is 0.07 and not significantly different from zero.

Turning to the shocks, estimates of σ_a, σ_z, and σ_r indicate that innovations to preference shocks are more volatile than technology and monetary policy innovations. Moreover, estimates of ρ_a (-0.30) and ρ_z (0.93) suggest that while technology shocks are highly persistent, preference shocks are not persistent and may even be negatively autocorrelated.[12,13]

Estimates of the deep habits model are obtained by restricting $b^a = 0$. The contribution of aggregate habits can be assessed by comparing these results to the nested model. Fixing $b^a = 0$ evidently has little effect on most of the parameters, notably the degree of deep habits b^d, which is nearly unchanged at 0.94. Obvious exceptions are the persistence and volatility of preference shocks. The estimate of σ_a falls by over half (0.12) while that of ρ_a becomes significantly positive (0.50). Omitting aggregate habits also lowers maximized log likelihood from 2380.41 to 2377.65. The p-value for the likelihood ratio statistic in this case is 0.0189, implying that the exclusion restriction is rejected at the 5% level but not the 1% level.

To evaluate the contribution of deep habits, I report estimates from the aggregate habits model obtained by restricting $b^d = 0$. Eliminating deep habits does not greatly affect estimates of b^a (0.65), but it has a big impact on how one interprets the preference shocks. Estimates of ρ_a (0.95) and σ_a (0.02) suggest that they are highly persistent but not significantly more volatile than technology shocks. Along with these changes, maximized log likelihood drops all the way to 2343.57. The p-value for the likelihood ratio test of $b^d = 0$ is less than 0.0001. Thus compared to previous results concerning aggregate habit formation, the model suffers a greater loss of explanatory power when deep habits are excluded.

5 Examining the Role of Deep Habits

Although likelihood analysis points to deep habits as the more empirically compelling, it is unclear which aspects of the data are driving the improvement in fit. To answer this question and gain insight into the role of habits *per se*, I compare simulations of the deep habits model to those from an identically-parameterized aggregate habits model.[14]

5.1 Volatilities and Correlations

Table 2 and Fig. 1 report standard deviations and autocorrelations of detrended consumption, inflation, and the interest rate from the two models described above. To see which one has superior business cycle properties, I also report the corresponding set of moments from the US data. These moments are computed from an unrestricted VAR(4).[15]

The deep habits model does a better job of accounting for the joint volatility of $\hat{c}_t$, $\hat{\pi}_t$, and $\hat{R}_t$. For each variable the model-implied standard deviation is well-within the 90% confidence interval around the VAR-based estimate, so differences between the two are insignificant at the 10% level. This is not true of the aggregate habits model, which tends to understate consumption volatility but greatly overstate inflation volatility.

Fig. 1 reveals that deep habits are also better at replicating some of the key dynamic interactions reflected in the autocorrelation functions. Nowhere is this more obvious than in the own correlation of inflation. The VAR results show inflation to be highly persistent with a correlation "half-life" of about seven quarters. Under deep habits the correlation between inflation and its own lag still exceeds 0.50 after one year and stays positive for up to three years. By contrast, there is almost no inflation persistence under aggregate habits. The correlation half-life is less than one quarter and turns slightly negative after just one year.[16]

5.2 Impulse Response Analysis

What factors drive the persistence and volatility of inflation observed under deep habits? To shed light on the key mechanisms, I now report impulse responses to a one-standard-deviation increase in the preference shock a_t and the technology shock z_t (both in logs).[17]

Consider the effects of a preference shock in Fig. 2. A positive innovation to a_t lifts consumption in both models because it increases the marginal utility of habit-adjusted consumption. Meanwhile, the shock also boosts labor demand as firms try to satisfy the temporary consumption boom. This raises work hours and the real wage along a fixed labor supply curve since preference shocks do not affect households' marginal rate of substitution in (4). With productivity z_t unchanged, the wage increase implies an equal percentage increase in real marginal cost given by the Lagrange multiplier γ_t in (7).

The adjustment paths described so far are similar for the two habit specifications. Where they depart is in the response of the shadow value of sales. Recall that this quantity, represented by the multiplier ν_t on the consumer demand function, measures the value to the firm of selling an extra unit in the current period. According to the figure, it falls sharply on impact under aggregate habits–by about 53%–compared to only 5% under deep habits.

The manner in which consumption habits affect the shadow value of sales can be seen more clearly by expressing ν_t in terms of the present value of expected future per-unit profits. Iterating (8) forward and imposing the various symmetric equilibrium conditions yields

$$\nu_t = E_t \sum_{j=0}^{\infty} \left(\beta b^d\right)^j \frac{\lambda_{t+j}}{\lambda_t} \left(1 - \gamma_{t+j}\right), \tag{12}$$

where λ_t is the marginal utility of habit-adjusted consumption. Note that when b^d is close to one, ν_t depends heavily on expectations of per-unit profits in the distant future. Because the shocks are transitory, long-run profit forecasts do not deviate much from the steady state. Consequently, even large changes in marginal cost over the short run have only a small percentage effect on the entire present value expression. When $b^d = 0$, as is true of the aggregate habits model, (12) collapses to $\nu_t = 1 - \gamma_t$. The value of selling an extra unit in this case is just current marginal profit since the sale is not expected to induce any future sales. Large shifts in marginal cost will therefore have a comparable percentage effect on ν_t.

Returning to Fig. 1, movements in the value of sales determine how firms react to the rise in marginal cost. With aggregate habits, the big drop in ν_t motivates firms to pass cost increases on to consumers via higher prices. As a result, annualized inflation surges to 9.4% and recedes quickly as the effects on marginal cost subside. With deep habits, the small decline in ν_t encourages firms to shield customers from higher costs by keeping prices low. In this case inflation falls in the impact period but rises gently to 4.3% three quarters later.

The reason why the shadow value of sales is more rigid and thus inflation less volatile and more persistent under deep habits is partly due to the *intertemporal* effect identified by Ravn *et al.* (2006). This effect emerges because firms recognize that current sales affect future consumption demand if $b^d > 0$. When faced with rising demand and cost conditions, firms have a powerful incentive to broaden their market share by holding down prices. The resulting growth in the habit stock enables firms to smooth out price increases over several quarters rather than front-load all of them as seen in the aggregate habits model.

The other channel through which deep habits affect inflation dynamics is the *price-elasticity* effect. As explained by Ravn *et al.* (2006), shocks that lift aggregate spending increase the relative size of the price-elastic component of the firm's demand schedule (6) when $b^d > 0$. This raises the short-run price elasticity of demand, which in a symmetric equilibrium can be expressed as $\epsilon_t \equiv \eta(1 - b^d c_{t-1}/c_t)$.[18] Following an expansionary preference shock, firms will therefore seek to limit any price increase in an effort to preserve market share. The same incentives do not exist in the aggregate habits model. When $b^d = 0$ the demand elasticity is constant and equal to η regardless of the spending level. Fig. 1 affirms this result. Under deep habits ϵ_t climbs by 10.7% on impact. This drives mark-ups even lower and bolsters the inertia already present in inflation from the intertemporal effect.

The joint impact that these two channels have on the inflation process is also evident in the optimal price-setting condition (9) when expressed in log-linear form as

$$\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + (1/\alpha)\left(\hat{c}_t - \hat{\nu}_t - \hat{x}_t\right). \tag{13}$$

Solving (13) forward and recognizing that $\hat{x}_t = \left(\hat{c}_t - b^d \hat{c}_{t-1}\right)/(1-b^d)$ gives

$$\hat{\pi}_t = -(1/\alpha)E_t \sum_{j=0}^{\infty} \beta^j \left(\hat{\nu}_{t+j} + \underbrace{\left(b^d/(1-b^d)\right)\left(\hat{c}_{t+j} - \hat{c}_{t+j-1}\right)}_{\hat{\epsilon}_{t+j}} \right), \tag{14}$$

showing inflation to be a function of the future paths of consumption growth and the shadow value of sales. As described above, the intertemporal effect is best captured by variation in ν_t and the price-elasticity effect by changes in the demand elasticity ϵ_t. A log-linear approximation of the demand elasticity yields $\hat{\epsilon}_t = \left(b^d/(1-b^d)\right)\left(\hat{c}_t - \hat{c}_{t-1}\right)$, which is precisely the second term in the forward solution for $\hat{\pi}_t$. This way of dissecting inflation clarifies how both mechanisms impart inertia. Following a preference shock, intertemporal effects choke off inflation by preventing a collapse in the shadow value of sales. Price-elasticity effects stamp out inflation to the extent that consumption growth in (14) offsets the declines in ν_t.

Fig. 3 displays the responses to a technology shock. A positive innovation to z_t relaxes the lifetime budget constraint which boosts consumption through a wealth effect channel. Meanwhile, the *real* interest rate (not shown) temporarily rises because monetary policy reduces the *nominal* rate by less than the fall in expected inflation. Higher real rates, in turn, mitigate some of the expansionary effects on consumption. Facing only modest growth in consumption, the rise in productivity enables firms to roll back their demand for labor, pushing hours of work, wages and marginal cost lower in periods immediately after the shock.

For reasons apparent in (12), the decline in marginal cost has disparate effects on the shadow value of sales in the two models. According to the figure, the impact-period rise in ν_t under aggregate habits is 12.7% compared to 2.5% under deep habits. It follows that firms are not as eager to slash prices when habits are deeply rooted. Indeed, inflation in this case only drops to 2.3% (down from 3.9%) but plunges to 1.1% under aggregate habits. Thus contrary to a preference shock, intertemporal effects help to discourage the large price cuts (instead of the price hikes) that would otherwise follow an increase in total factor productivity.

Price elasticity effects also influence the path of inflation. Unlike Fig. 2, however, these tend to undermine rather than reinforce the intertemporal effects. Because technology shocks gradually increase consumption, demand elasticities under deep habits rise and then fall in a hump-shaped pattern. During high-elasticity periods, firms have an incentive to lower prices in an effort to capture market share. This intensifies the downward pressure on inflation and counteracts the upward pull being exerted by the intertemporal effects. The result can also be seen in (14). Technology shocks, because they increase ν_t and ϵ_t, amplify the disinflation experienced under deep habits. Yet despite the extra push given by the price-elasticity effect in this case, the intertemporal effect is evidently strong enough to prevent inflation from falling as much as it would if deep habits were absent from the model altogether.

6 Additional Sources of Persistence

The gains in fit observed under deep habits come from its ability to generate greater inflation persistence. One could conclude then that the aggregate habits model is deficient simply because it has no internal persistence mechanisms other than serial correlation inherited from the shocks. But this raises the possibility that the case for deep habits might not be as compelling if evaluated in the context of a richer model with multiple sources of persistence.

To explore this possibility, I modify the original setup by including two ad hoc elements capable of generating persistence irrespective of deep habits. The first one borrows from Ireland (2007) by placing a backward-looking term in the adjustment cost function. Firms now face adjustment costs of the form $(\alpha/2)\left[P_t(i)/\left((\pi_{t-1}^{\varrho}\pi^{1-\varrho}P_{t-1}(i)\right)-1\right]^2 y_t$, where $\varrho \in [0,1]$ measures the degree to which lagged inflation serves as a reference for price setting. If $\varrho = 1$ firms incur costs only to the extent that growth in $P_t(i)$ deviates from π_{t-1}. If $\varrho = 0$ adjustment costs reduce to the benchmark case where steady-state inflation is the reference value. A log-linear approximation of the optimal pricing condition yields

$$(\hat{\pi}_t - \varrho\hat{\pi}_{t-1}) = \beta E_t\left(\hat{\pi}_{t+1} - \varrho\hat{\pi}_t\right) + (1/\alpha)\left(\hat{c}_t - \hat{\nu}_t - \hat{x}_t\right). \tag{15}$$

Inflation now has two sources of persistence. One is inherited from the forcing variable, $\hat{c}_t - \hat{\nu}_t - \hat{x}_t$, and the other is the "built-in" persistence imparted by lagged inflation when $\varrho > 0$. The second modification allows for serial correlation in the policy shock by modeling $\varepsilon_{r,t}$ as an AR(1) process with autoregressive coefficient $\rho_r \in [0,1)$. Due to sticky prices, greater persistence in the nominal interest rate translates into greater persistence in consumption. This in turn strengthens inflation persistence via the forcing process in (15).

Table 3 reports new estimates of the nested habits model along with the two restricted models. The results should help determine whether the original model falsely attributes inflation persistence to deep habits when some of that persistence is actually due to backward-looking components in the pricing equation or serial correlation in the policy shocks.

Changes made in this section have little effect on inferences of the nested model or the deep habits model. None of the estimates are significantly different from their counterparts in Table 1. I also find no evidence of lagged inflation in (15). Estimates of ϱ lie up against the zero bound, so the data prefer to have persistence derive from deep habits rather than backward-looking frictions in price setting. There is some evidence of serial correlation in the policy shock. Estimates of ρ_r span 0.21 to 0.25 and are significantly different from zero.[19]

Estimates of the aggregate habits model are not as robust. Setting $b^d = 0$ drives up the estimate of ρ_r to 0.66 but pushes down θ_r and θ_π to 0.31 and 1.17. Inferences about habit formation are also affected. The estimate of b^a is 0.79, almost 22% higher than the estimate in Table 1. Despite these changes, the aggregate habits model still does not attribute any persistence to lagged inflation. The estimate of ϱ is near zero and statistically insignificant.

7 Conclusion

This paper estimates a new-Keynesian model with habit formation. Central to the model is a utility function that nests aggregate and deep consumption habits as special cases. Maximum likelihood reveals that the data prefer an arrangement in which habits over differentiated products are stronger than habits over the aggregate finished good. Although separate likelihood ratio tests reject the hypothesis that either type should be excluded (at the 5% level), results show that the slump in model fit is far greater when deep habits are missing.

I trace this to the ability of deep habits to shape the dynamics of inflation in a manner consistent with US data. Product-level habits motivate firms to smooth out price changes over time. This feature derives from well-known intertemporal and price-elasticity effects that coalesce with nominal frictions to produce a model capable of imparting substantial inertia on inflation. Simulations indicate that deep habits are critical for matching the volatility and persistence observed in the sample. The same behavior is evident in the impulse responses, which show inflation reacting sluggishly to preference and technology shocks when deep habits are preserved but swiftly and less persistent when replaced by aggregate habits.

Notes

[1]This specification, often referred to as "catching up with the Joneses," is different from *internal* habit formation in which households value consumption relative to their own past consumption. Dennis (2009) studies the empirical consequences of internal vs. external habit formation from a new-Keynesian perspective.

[2]Setting $b^d = 0$ implies $x_{j,t} = \left[\int_0^1 c_{j,t}(i)^{1-1/\eta} di\right]^{1/(1-1/\eta)}$, which is just a standard Dixit-Stiglitz aggregator commonly used in models with imperfectly competitive markets.

[3]Household j's efforts to minimize the period-by-period cost of assembling each unit of $x_{j,t}$ implies that, at the optimum, $\int_0^1 P_t(i)c_{j,t}(i)di = P_t x_{j,t} + b^d \int_0^1 P_t(i)c_{t-1}(i)di$.

[4]In equilibrium the stochastic discount factor satisfies $q_{0,t}P_t = \beta^t a_t/(x_t - b^a x_{t-1})$.

[5]Due to habit formation, selling a unit of good i in the current period raises sales by b^d units in the following period, the present discounted value of which equals $b^d E_t \frac{q_{0,t+1}}{q_{0,t}} \pi_{t+1}\nu_{t+1}(i)$.

[6]The policy coefficients $\{\theta_r, \theta_\pi, \theta_y\}$ are jointly restricted to guarantee a locally unique rational expectations equilibrium. See Zubairy (2013) for a discussion of how deep habits modify the local determinacy conditions of an otherwise standard new-Keynesian model.

[7]Subscript j and function argument i can therefore be dropped from variables appearing in the optimality conditions.

[8]I de-mean the inflation and interest rate series prior to estimation. Consumption data exhibits a secular trend, so I regress the log of per capita consumption against a constant and a linear time trend. Least squares residuals are used for estimation.

[9]In a separate appendix to the paper (available at http://mycba.ua.edu/∼gegivens/research) I examine how robust the parameter estimates are to setting $(1-\phi)^{-1} = 3$ quarters. Also reported in the appendix are estimates of the deep habits model for varying degrees of price fixity ranging from 4 months to 12 months.

[10]Standard errors correspond to the square roots of the diagonals of the inverse Hessian matrix.

[11]The appendix contains a brief discussion of identification issues concerning estimates of b^a and b^d.

[12]The chi-square statistic from a likelihood ratio test of the hypothesis $\rho_a = 0$ has a p-value of 0.0084.

[13]In the appendix I provide some intuition for why preference shocks in the nested habits model lack persistence but are highly volatile, which is somewhat atypical of most empirical DSGE models. I also re-estimate the nested habits model under the restriction $\rho_a \in [0, 1)$ and comment briefly on the findings.

[14]The analysis here compares simulations of the deep habits model using the estimates in the second column of Table 1 to simulations of an aggregate habits model that uses the exact same parameter values (e.g., $b^d = b^a = 0.9414$). Estimates of the nested habits model in the first column of Table 1 are not used.

[15]The appendix reports additional standard deviation and autocorrelation results for the nested and aggregate habits models using the estimates that appear in the first and third columns of Table 1, respectively.

[16]Confidence intervals for the autocovariance functions are obtained using Monte Carlo methods as follows. First, I take the joint distribution of the VAR coefficient estimates and the residual covariance matrix to be asymptotically normal with mean given by the sample estimates and covariance given by the sample covariance matrix of those estimates. Second, I draw 10,000 random vectors from this multivariate normal distribution and compute the corresponding autocovariance functions for each draw. Third, I rank the autocovariances for each variable pair and for each lag in descending order. The 90% confidence intervals are bounded by the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles of the ordered autocovariances.

[17]Absent here are the responses to a policy shock $\varepsilon_{r,t}$. Since they originate on the demand side, monetary shocks produce dynamics that are similar to preference shocks. Moreover, Ravn *et al.* (2010) only consider monetary shocks in their analysis of deep habits. Emphasizing preference and technology shocks therefore shifts the orientation of this paper towards findings that have not received as much attention in the literature.

[18]The price elasticity of demand for good i is $\epsilon_t(i) \equiv -\frac{(P_t(i)/P_t)}{c_t(i)} \frac{\partial c_t(i)}{\partial (P_t(i)/P_t)} = \eta \left(c_t(i) - b^d c_{t-1}(i)\right)/c_t(i)$.

[19]The chi-square statistic from a likelihood ratio test of the hypothesis that $\rho_r = 0$ has a p-value of 0.0010 in the nested habits model and 0.0035 in the deep habits model.

References

Abel, Andrew B. "Asset Prices under Habit Formation and Catching Up with the Joneses." *The American Economic Review Papers and Proceedings*, May 1990, 80(2), pp. 38-42.

Basu, Susanto and Fernald, John G. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy*, April 1997, 105(2), pp. 249-83.

Chintagunta, Pradeep; Kyriazidou, Ekaterini and Perktold, Josef. "Panel Data Analysis of Household Brand Choices." *Journal of Econometrics*, July 2001, 103(1-2), pp. 111-53.

Christiano, Lawrence J.; Eichenbaum, Martin and Evans, Charles L. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, January 2005, 113(1), pp. 1-45.

Dennis, Richard. "Consumption Habits in a New Keynesian Business Cycle Model." *Journal of Money, Credit and Banking*, August 2009, 41(5), pp. 1015-30.

Fuhrer, Jeffrey C. "Habit Formation in Consumption and its Implications for Monetary-Policy Models." *The American Economic Review*, June 2000, 90(3), pp. 367-90.

Harvey, Andrew C. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge, UK: Cambridge University Press, 1989.

Ireland, Peter N. "Endogenous Money or Sticky Prices?" *Journal of Monetary Economics*, November 2003, 50(8), pp. 1623-48.

__________. "Changes in the Federal Reserve's Inflation Target: Causes and Consequences." *Journal of Money, Credit and Banking*, November 2007, 39(8), pp. 1851-82.

Klein, Paul. "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model." *Journal of Economic Dynamics and Control*, September 2000, 24(10), pp. 1405-23.

Lubik, Thomas A. and Teo, Wing Leong. "Deep Habits in the New Keynesian Phillips Curve." forthcoming in *Journal of Money, Credit and Banking*, 2013.

Monacelli, Tommaso. "New Keynesian Models, Durable Goods, and Collateral Constraints." *Journal of Monetary Economics*, March 2009, 56(2), pp. 242-54.

Ravn, Morten; Schmitt-Grohé, Stephanie and Uribe, Martin. "Deep Habits." *The Review of Economic Studies*, January 2006, 73(1), pp. 195-218.

Ravn, Morten O.; Schmitt-Grohé, Stephanie; Uribe, Martin and Uuskula, Lenno. "Deep Habits and the Dynamic Effects of Monetary Policy Shocks." *Journal of the Japanese and International Economies*, June 2010, 24(2), pp. 236-58.

Rotemberg, Julio J. "Sticky Prices in the United States." *Journal of Political Economy*, December 1982, 90(6), pp. 1187-1211.

Smets, Frank and Wouters, Raf. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *The American Economic Review*, June 2007, 97(3), pp. 586-606.

Woodford, Michael. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton and Oxford: Princeton University Press, 2003.

Zubairy, Sarah. "Interest Rate Rules and Equilibrium Stability under Deep Habits." forthcoming in *Macroeconomic Dynamics*, 2013.

Table 1
Parameter Estimates (1965:Q3 - 2012:Q1)

Model Parameter	Parameter Description	Nested Habits	Deep Habits	Aggregate Habits
σ_a	*preference shock*	0.3006 (0.0575)	0.1229 (0.0170)	0.0224 (0.0065)
σ_z	*technology shock*	0.0114 (0.0026)	0.0162 (0.0059)	0.0164 (0.0034)
σ_r	*policy shock*	0.0017 (0.0001)	0.0017 (0.0001)	0.0020 (0.0001)
ρ_a	*AR preference shock*	-0.2999 (0.0922)	0.5011 (0.0719)	0.9478 (0.0233)
ρ_z	*AR technology shock*	0.9333 (0.0295)	0.8983 (0.0449)	0.9965 (0.0048)
θ_r	*interest rate smoothing*	0.9026 (0.0183)	0.9073 (0.0188)	0.7956 (0.0261)
θ_π	*inflation response*	1.5406 (0.3031)	1.4814 (0.3331)	1.4687 (0.1825)
θ_y	*output response*	0.0685 (0.0427)	0.0894 (0.0486)	-0.0155 (0.0131)
b^a	*aggregate habit*	0.6111 (0.0546)	*0*	0.6461 (0.0674)
b^d	*deep habit*	0.9438 (0.0069)	0.9414 (0.0077)	*0*
χ	*Frisch elasticity*	2.0135 (0.6603)	1.0655 (0.6334)	1.5749 (0.7689)
α	*price adjustment cost*	*59.3778*	*59.3778*	*59.3778*
β	*discount factor*	*0.9965*	*0.9965*	*0.9965*
η	*substitution elasticity*	*6*	*6*	*6*
μ	*markup*	1.2142 (0.0019)	1.2136 (0.0019)	1.2000 (0.0000)
$\ln \mathcal{L}$	*log likelihood*	2380.4089	2377.6548	2343.5735
p-value	*likelihood ratio test*	–	0.0189	0.0000

Notes: The table reports maximum-likelihood estimates of the nested model, the deep habits model ($b^a = 0$), and the aggregate habits model ($b^d = 0$). Standard errors are in parentheses. Italicized numbers denote values that are imposed prior to estimation.

Table 2
Standard Deviations

Model	SD($\hat{c}_t$)		SD($\hat{\pi}_t$)		SD($\hat{R}_t$)	
Deep Habits	0.0320		0.0075		0.0070	
Aggregate Habits	0.0275		0.0172		0.0075	
VAR(4)	0.0376	[0.0296, 0.0841]	0.0066	[0.0056, 0.0136]	0.0076	[0.0061, 0.0182]

Notes: Simulations of the deep and aggregate habits models use the same parameter values. Numbers in squared brackets correspond to 90% confidence intervals for the standard deviations implied by an unconstrained VAR(4) on $\hat{c}_t$, $\hat{\pi}_t$, and $\hat{R}_t$.

Table 3
Parameter Estimates (1965:Q3 - 2012:Q1)

Model Parameter	Parameter Description	Nested Habits	Deep Habits	Aggregate Habits
σ_a	*preference shock*	0.2795 (0.0542)	0.1181 (0.0181)	0.0120 (0.0073)
σ_z	*technology shock*	0.0119 (0.0032)	0.0172 (0.0096)	0.0316 (0.0140)
σ_r	*policy shock*	0.0017 (0.0001)	0.0017 (0.0001)	0.0031 (0.0006)
ρ_a	*AR preference shock*	-0.3007 (0.0926)	0.4919 (0.0748)	0.9651 (0.0182)
ρ_z	*AR technology shock*	0.9273 (0.0345)	0.8866 (0.0663)	0.9956 (0.0060)
ρ_r	*AR policy shock*	0.2488 (0.0756)	0.2143 (0.0726)	0.6626 (0.0503)
θ_r	*interest rate smoothing*	0.8790 (0.0243)	0.8911 (0.0243)	0.3062 (0.1166)
θ_π	*inflation response*	1.3909 (0.3014)	1.3251 (0.3895)	1.1683 (0.1436)
θ_y	*output response*	0.0513 (0.0419)	0.0823 (0.0495)	-0.0245 (0.0164)
b^a	*aggregate habit*	0.6080 (0.0543)	*0*	0.7865 (0.0861)
b^d	*deep habit*	0.9400 (0.0079)	0.9390 (0.0090)	*0*
χ	*Frisch elasticity*	1.6564 (0.6672)	0.8568 (0.7132)	0.6279 (0.4068)
ϱ	*lagged inflation*	6.9e-8 (0.0884)	3.8e-8 (0.0938)	1.2e-7 (0.0463)
α	*price adjustment cost*	*59.3778*	*59.3778*	*59.3778*
β	*discount factor*	*0.9965*	*0.9965*	*0.9965*
η	*substitution elasticity*	*6*	*6*	*6*
μ	*markup*	1.2133 (0.0019)	1.2130 (0.0021)	1.2000 (0.0000)
$\ln \mathcal{L}$	*log likelihood*	2385.8152	2381.9230	2355.2368
p-value	*likelihood ratio test*	–	0.0053	0.0000

Notes: The table reports maximum-likelihood estimates of the nested model, the deep habits model ($b^a = 0$), and the aggregate habits model ($b^d = 0$). Standard errors are in parentheses. Italicized numbers denote values that are imposed prior to estimation.

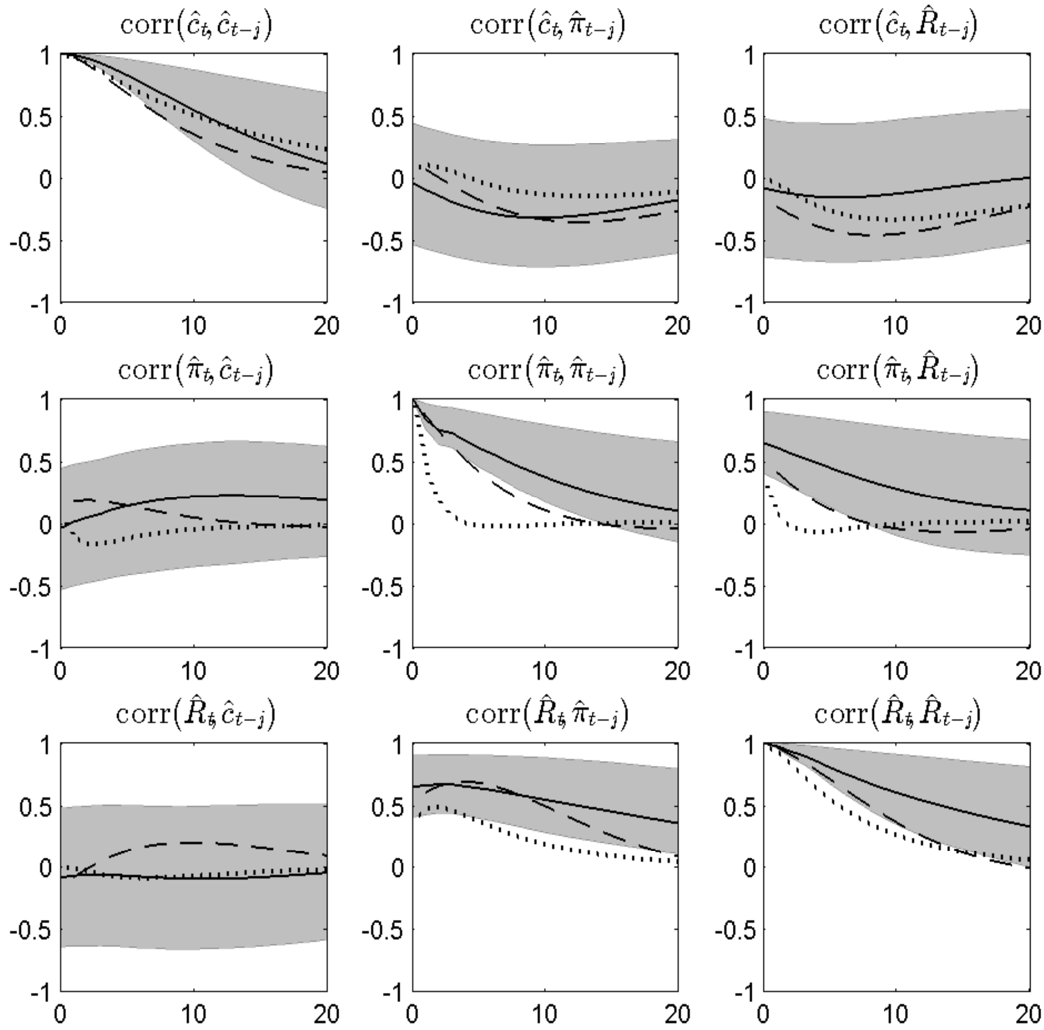

Fig. 1. The autocorrelation function for consumption $\hat{c}_t$, inflation $\hat{\pi}_t$, and the interest rate $\hat{R}_t$ is drawn for the US data (solid line), the estimated deep habits model (dashed line), and an aggregate habits model (dotted line) that uses the same parameter values. Correlations for the US data are obtained from a VAR(4), and the shaded areas correspond to 90% confidence bands.

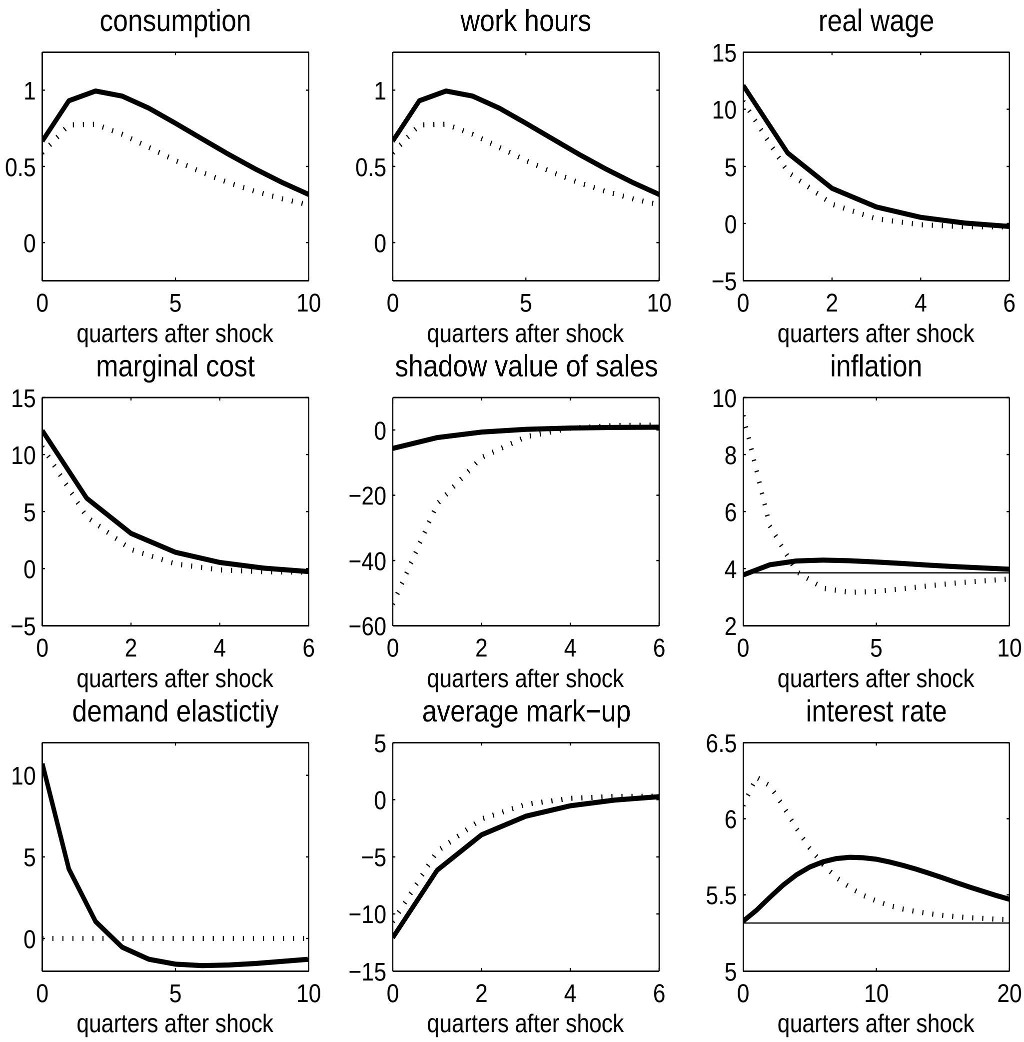

Fig. 2. Impulse responses to a **preference** shock are drawn for the estimated deep habits model (solid line) and an aggregate habits model (dotted line) that uses the same parameter values. Inflation and the nominal interest rate are measured in annualized percentage points. All other variables are expressed as percent deviations from the steady state.

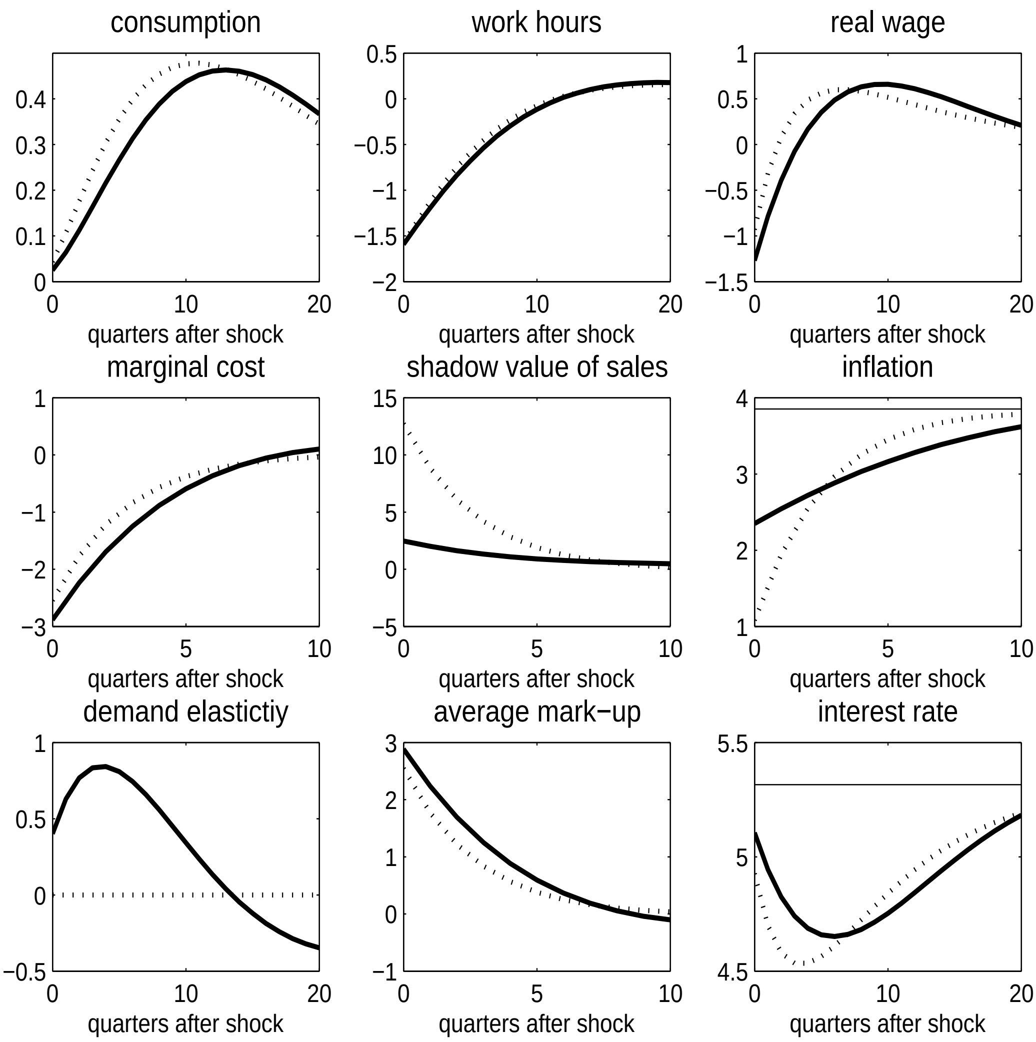

Fig. 3. Impulse responses to a **technology** shock are drawn for the estimated deep habits model (solid line) and an aggregate habits model (dotted line) that uses the same parameter values. Inflation and the nominal interest rate are measured in annualized percentage points. All other variables are expressed as percent deviations from the steady state.