Econometrica Supplementary Material

SUPPLEMENT TO "CONSUMPTION COMMITMENTS AND HABIT FORMATION"

(*Econometrica*, Vol. 84, No. 2, March 2016, 855–890)

BY RAJ CHETTY AND ADAM SZEIDL

THIS MATERIAL SUPPLEMENTS THE PAPER "Consumption Commitments and Habit Formation." We provide missing proofs for results stated in the main paper and we explain the numerical methods used to simulate the model.

S.1. PROOFS OF PROPOSITIONS 2 AND 3

PROOF OF PROPOSITION 2: Since the only risky assets for household i are S and S^i, there exists a unique state price density associated with the household-specific private market. The following dynamics for adjustable consumption generates a state price density that prices both risky assets as well as the safe asset:

$$a_t^i = a_0^i \exp\left\{\frac{1}{\gamma}\left(\frac{\pi^2}{2\sigma^2} + \frac{\pi_I^2}{2\sigma_I^2} + r - \rho\right)t + \frac{\pi}{\gamma\sigma}z_t + \frac{\pi_I}{\gamma\sigma_I}z_t^i\right\} \tag{15}$$

and hence must describe the optimal choice of household i. Because $a_0^i = A_0$ for all i, aggregating across i yields, by the strong law of large numbers for a continuum of agents (Sun (1998)),

$$\begin{aligned} A_t &= A_0 \exp\left\{\frac{1}{\gamma}\left(\frac{\pi^2}{2\sigma^2} + \frac{\pi_I^2}{2\sigma_I^2} + r - \rho\right)t + \frac{\pi}{\gamma\sigma}z_t\right\} \int_i \exp\left\{\frac{\pi_I}{\gamma\sigma_I}z_t^i\right\} di \\ &= A_0 \exp\left\{\frac{1}{\gamma}\left(\frac{\pi^2}{2\sigma^2} + \frac{\pi_I^2}{2\sigma_I^2}\left(1 + \frac{1}{\gamma}\right) + r - \rho\right)t + \frac{\pi}{\gamma\sigma}z_t\right\}. \end{aligned}$$

Define a new discount rate $\delta = \rho - (1 + \frac{1}{\gamma})\pi_I^2/(2\sigma_I^2)$. Then the dynamics of aggregate adjustable consumption is given by

$$A_t = A_0 \exp\left\{\frac{1}{\gamma}\left(\frac{\pi^2}{2\sigma^2} + r - \delta\right)t + \frac{\pi}{\gamma\sigma}z_t\right\}.$$

This is exactly the dynamics of adjustable consumption that would obtain for a representative consumer with power utility over A_t and discount rate δ who can invest in the publicly traded risky and safe assets. *Q.E.D.*

PROOF OF PROPOSITION 3: We are interested in characterizing the evolution of the conditional distribution of y_t^i given a realization of the path of A

 DOI: 10.3982/ECTA9390

under Q. Using (15), we obtain

$$d\log a_t^i = \frac{1}{\gamma}\left(\frac{\pi^2}{2\sigma^2} + \frac{\pi_I^2}{2\sigma_I^2} + r - \rho\right)dt + \frac{\pi}{\gamma\sigma}dz + \frac{\pi_I}{\gamma\sigma_I}dz^i$$
$$= \theta\, dt + \frac{\pi}{\gamma\sigma}dz + \frac{\pi_I}{\gamma\sigma_I}d\overline{z}^i,$$

where

$$\theta = \frac{1}{\gamma}\left(\frac{\pi^2}{2\sigma^2} + \frac{\pi_I^2}{2\sigma_I^2} + r - \rho\right) + \frac{\pi_I^2}{\gamma^2\sigma_I^2} \tag{16}$$

is the drift under Q. We first show that $F(y,t)$ is absolutely continuous for all $t > 0$ for almost all realizations of the path of aggregate shocks. We do this assuming that the initial condition is $a_0^i = A_0$ and $x_0^i = X_0$ for all agents i, that is, that the initial distribution $F_0(y)$ is concentrated on a single point. For other initial distributions, the density $f(y,t)$ can simply be computed as an integral of these densities with respect to $F_0(y)$.

Throughout the argument, we work with the probability measure Q. Our proof logic is to fix $t = T$ and the realization of A_t for $t \in [0,T]$, pick a collection of intervals $I \subset [L,U]$, compute an upper bound on the probability that $y_T \in I$, and then establish that the upper bound goes to zero as the total length of these intervals, denoted $|I|$, goes to zero. Our upper bound is obtained by separately bounding the probabilities of two events:

(1) Reaching I through paths that do not involve "too many" adjustments. Let $\widetilde{y}_0^i = y_0$ and

$$d\widetilde{y}_t^i = -\theta\cdot dt - \sigma_A\cdot dz_t - \sigma_I\cdot d\overline{z}_t^i.$$

Given the dynamics of $\log a_t^i$, this specification implies that the evolution of $\widetilde{y}$ is the same as that of y except for the discrete adjustments. In particular, $\widetilde{y}_t^i = y_t^i$ before the first adjustment occurs. More generally, if y^i experiences n_U upward and n_D downward adjustments in the interval $[0,t]$, then $y_t^i = \widetilde{y}_t^i + n_D(M - L) - n_U(U - M)$. Because $\widetilde{y}_t^i$ is a Brownian motion with a drift, its density is bounded from above by some constant which depends on the parameters of the process, which we denote by $K(\mu_a, \sigma_A, \sigma_I, T)$. As a result, for any given $n \geq 1$, the total probability of paths which involve $n_U < n$ upward and $n_D < n$ downward adjustments such that $y_T^i \in I$ is at most $K(\mu_a, \sigma_A, \sigma_I, T)\cdot n^2\cdot |I|$.

(2) The total probability of paths that involve at least n adjustments. Let $\widetilde{y}_0^A = y_0$ and $d\widetilde{y}_t^A = -\theta\cdot dt - \sigma_A\cdot dz_t$ so that $\widetilde{y}_t^A$ represents the aggregate shocks and trend in $\widetilde{y}_t$, and let $\widetilde{y}_0^{I,i} = 0$ and $d\widetilde{y}_t^{I,i} = \sigma_I\cdot d\overline{z}_t^i$ so that $\widetilde{y}_t^{I,i}$ represents the idiosyncratic shocks. Then $\widetilde{y}_t^i = \widetilde{y}_t^A + \widetilde{y}_t^{I,i}$. The path of $\widetilde{y}_t^A$ contains the same information as the path of aggregate shocks A_t, hence we are effectively conditioning on the realization of the path of $\widetilde{y}_t^A$. Set $\Delta_y = \min(U - M, M - L)/2$.

We say that a process u_t moves Δ_y between s and t if $|u_t - u_s| = \Delta_y$. Suppose that $s_1 < s_2$ are two consecutive adjustment dates for household i. Then either $\widetilde{y}_t^A$ or $\widetilde{y}_t^{I,i}$ must move at least Δ_y between s_1 and s_2. Because almost surely the path of $\widetilde{y}_t^A$ is continuous, one can straightforwardly verify that there is an upper bound $K(\widetilde{y}_{[0,T]}^A)$ on the number of non-overlapping time intervals in $[0, T]$ over which y_t^A moves at least Δ_y. For ease of notation, in the rest of this proof we will simply denote $K(\widetilde{y}_{[0,T]}^A) = K$. Then, if household i adjusts at least n times in $[0, T]$, there must exist at least $n - K$ non-overlapping intervals in $[0, T]$ over which $\widetilde{y}_t^{I,i}$ moves at least Δ_y. Assume now that $n > 2K + 1$. At least one of these intervals—denote it by $[s_1, s_2]$—cannot be longer than $T/(n - K)$. Now cover the $[0, T]$ interval with subintervals of length $2T/(n - K)$ starting at zero, and by another set starting at $T/(n - K)$. It is clear that an interval in one of these covers, say $[s_0, s_3]$, must fully contain $[s_1, s_2]$.

The probability that $\widetilde{y}_t^{I,i}$ moves at least Δ_y over $[s_1, s_2]$ is bounded by the probability that the difference between the minimum and the maximum of $\widetilde{y}_t^{I,i}$ in $[s_0, s_3]$ is at least Δ_y. Given that the density of the running maximum of a standard Brownian motion is $(2/(\pi t))^{1/2} e^{-m^2/(2t)}$, this probability is bounded above by a universal constant times $((n - K)/(\pi T \sigma_I^2))^{1/2} \exp[-\Delta^2(n - K)/(2T\sigma_I^2)]$. Because the total number of intervals in the two covers we introduced is at most $2(n - K)$, the probability that $\widetilde{y}_t^{I,i}$ moves at least Δ_y over an interval of length at most $T/(n - K)$ is bounded from above by a constant (which depends on T and σ_I) times $(n - K)^{3/2} \exp[-\Delta^2(n - K)/(2T\sigma_I^2)]$. Recalling the assumption that $n > 2K + 1$, the last expression can be bounded above by a different constant (which depends on T and σ_I^2) times $\exp[-\Delta^2 n/(8T\sigma_I^2)]$.

We now combine these bounds. Given K, which is determined by the path of $\widetilde{y}_t^A$, and maintaining $n > 2K + 1$, the total probability that $y_T^i \in I$ is at most

$$K(\mu_a, \sigma_A, \sigma_I, T) \cdot n^2 \cdot |I| + K\big(\sigma_I^2, T\big) \cdot \exp\big[-\Delta^2 n/\big(8T\sigma_I^2\big)\big].$$

Setting $n = |I|^{-1/4}$, for small enough $|I|$ such that $n > 2K + 1$ is satisfied, the bound becomes

$$K(\mu_a, \sigma_A, \sigma_I, T) \cdot |I|^{1/2} + K\big(\sigma_I^2, T\big) \cdot \exp\big[-\Delta^2 |I|^{-1/4}/\big(8T\sigma_I^2\big)\big],$$

which goes to zero as $|I|$ goes to zero.

We now turn to the stochastic partial differential equation. Proposition 1 in Caballero (1993) derives a stochastic partial differential equation, given the path of aggregate shocks, for the conditional density of a double-barrier Brownian motion with rebirth. Caballero's equation is

$$df(y, t) = \left[\theta \frac{\partial f(y, t)}{\partial y} + \frac{\sigma_T^2}{2} \frac{\partial^2 f(y, t)}{\partial y^2}\right] dt + \sigma_A \frac{\partial f(y, t)}{\partial y}\, dz.$$

Substituting in (16) yields the equation in the text. The boundary conditions follow directly from Caballero's proposition.

To derive the dynamics of aggregate commitments, note that $X_t = \int_L^U e^y \times f(y,t)\,dy \cdot A_t$ and we can use Ito's lemma to write

$$
\begin{aligned}
dX_t = {} & A_t \int_L^U e^y \cdot df(y,t) \cdot dy + dA_t \cdot \int_L^U e^y f(y,t)\,dy \\
& + \left\langle \int_L^U e^y \cdot df(y,t) \cdot dy, dA_t \right\rangle .
\end{aligned}
$$

We now evaluate each term on the right-hand side. The first term is

$$
\begin{aligned}
& A_t \int_L^U e^y \cdot \frac{\partial f(y,t)}{\partial y} \left\{ \left(\mu + \frac{\pi_I^2}{2\gamma^2 \sigma_I^2} \right) dt + \frac{\pi}{\gamma\sigma}\,dz \right\} dy \\
& \quad + F_t \int_L^U e^y \cdot \frac{\partial^2 f(y,t)}{\partial y^2} \frac{\sigma_T^2}{2}\,dt \cdot dy.
\end{aligned}
$$

Integrating by parts, and using the boundary conditions, shows that this term equals

$$
\begin{aligned}
& -X_t \left(\left(\mu + \frac{\pi_I^2}{2\gamma^2 \sigma_I^2} \right) dt + \frac{\pi}{\gamma\sigma}\,dz \right) \\
& \quad + A_t \frac{\sigma_T^2}{2} \cdot \left(f_y(L,t)\left(e^M - e^L\right) + f_y(U,t)\left(e^U - e^M\right) \right) dt + \frac{\sigma_T^2}{2} X_t\,dt.
\end{aligned}
$$

The second term is

$$
X_t \cdot \frac{dA_t}{A_t} = X_t \left(\left(\mu + \frac{\pi^2}{2\gamma^2\sigma^2} \right) dt + \frac{\pi}{\gamma\sigma}\,dz \right),
$$

while the third term is simply $-\pi^2/(\gamma\sigma)^2 X_t\,dt$. Collecting terms gives the result of the proposition. *Q.E.D.*

S.2. PROOFS OF RESULTS LEADING UP TO THEOREM 1

S.2.1. *Proofs of Auxiliary Results About the Commitments Model Including Proof of Proposition 4*

PROOF OF LEMMA 1: We start with the case where w_t is driven by a standard Brownian motion. Let $\zeta_y = \inf\{t \geq 0 : w_t \notin [L,U], w_0 = y\}$. Let $F_w(t) = \Pr[\zeta_y \leq t]$ and $\overline{h}(y,t) = E[e^{w_t} \cdot 1\{\zeta_y > t\}]$ be $h(y,t)$ killed at the boundary. Let

$F_y^{(1)}(t) = F_y(t)$ and $F_y^{(n+1)}(t) = \int_0^t F_{y^*}^{(n)}(t-\tau)\, dF_y(\tau) = \int_0^t F_M(t-\tau)\, dF_y^{(n)}(\tau)$ be the distribution of the $n+1$st exit time. Then

$$
\begin{aligned}
h(y,t) &= \overline{h}(y,t) + \sum_{n=1}^{\infty} \int_0^t \overline{h}(M, t-\tau)\, dF_y^{(n)}(\tau) \\
&= \overline{h}(y,t) + \int_0^t \overline{h}(M, t-\tau)\, dF_y^*(\tau),
\end{aligned}
\tag{17}
$$

where

$$
\begin{aligned}
F_y^*(t) &= \sum_{n=1}^{\infty} F_y^{(n)}(t) = F_y(t) + \int_0^t F_M^*(t-\tau)\, dF_y(\tau) \\
&= F_y(t) + \int_0^t F_M(t-\tau)\, dF_y^*(\tau)
\end{aligned}
\tag{18}
$$

is the expected number of boundary hits until t.

The transition density of the killed diffusion $p(y, y', t) = \Pr[\zeta_y > t, y_t = y']$ can be expressed as an infinite sum of normal densities (Revuz and Yor (1994, p. 106)), and in particular, is infinitely many times differentiable in $[L, U] \times [L, U] \times (0, \infty)$. This implies that $\overline{h}(y,t) = \int e^{y'} p(y, y', t)\, dy'$ is infinitely many times differentiable in $[L, U] \times (0, \infty)$. The density of the first hitting time ζ_y can also be expressed in closed form as an infinite sum (Darling and Siegert (1953)), and is infinitely many times differentiable in y and t over $[L, U] \times (0, \infty)$. This, combined with (18), implies that $F_y^*(t)$ is C^∞ in $[L, U] \times (0, \infty)$. Combining these observations with (17) shows that $h(y,t)$ is also C^∞ in the $[L, U] \times (0, \infty)$ domain.[1]

We next show that h is also smooth when driven by any Brownian motion with drift and variance, and that it is smooth in the other parameters. Changing the clock of y_t scales both the mean and the variance, and is obviously a smooth transformation of $h(y,t)$ as it just scales the time argument. Shifting and rescaling the vertical axis are smooth operations that shift and rescale the triple $[L, M, U]$. Thus we only need to show smoothness in the drift and in M. The drift can be dealt with using the Girsanov theorem, which implies that the density of the killed diffusion under drift can be obtained as $p^{\mu w}(y, y', t) = p(y, y', t) \cdot \exp[\mu_w(y' - y) - \mu_w^2 t/2]$, which is clearly C^∞ in μ_w, and hence so is $\overline{h}(y,t)$. Next, the distribution of the first hitting time is $1 - F_y^{\mu_y}(t) = \int p^{\mu_y}(y, y', t)\, dy'$ which is also smooth. The smoothness of h in μ_y now follows from (17). Smoothness in M follows easily from (17). *Q.E.D.*

[1]Grigorescu and Kang (2002) computed the transition density of y explicitly.

PROOF OF LEMMA 2: We have

$$E_s[\overline{X}_t] = \overline{A}_s \cdot E_s^R[\overline{X}_t/\overline{A}_t] = \overline{A}_s \cdot E_s^{QR}[x_t/a_t]$$
$$= \overline{A}_s \cdot \int_L^U h(t-s, y) f(y, s)\, dy,$$

which is a martingale in s. Computing the Ito differential

$$d_s E_s[\overline{X}_t] = d\overline{A}_s \cdot E_s^{QR}[x_t/a_t] + \overline{A}_s \cdot \int_L^U h(t-s, y) f_y(y, s) \sigma_A\, dz_s \cdot dy,$$

where we used (7) for the evolution of $f(y, s)$ and collected only the dz terms, since the ds terms must cancel by the martingale property. Equivalently,

$$d_s E_s[\overline{X}_t] = d\overline{A}_s \cdot \left(E_s^{QR}[x_t/a_t] + \int_L^U h(t-s, y) f_y(y, s)\, dy \right)$$
$$= d\overline{A}_s \cdot \int_L^U \big(h(u, y) - h_y(u, y)\big) f(y, s)\, dy,$$

where we integrated by parts. This equation shows the existence of ξ as well as the desired representation. *Q.E.D.*

PROOF OF LEMMA 3: Ben-Ari and Pinsky (2009) showed that $y_t = \log[x_t/a_t]$ converges exponentially fast to a unique invariant distribution. It follows from Ben-Ari and Pinsky (2007) that the rate of convergence is uniformly bounded if the drift is from a bounded interval. This implies uniform convergence for all $\sigma_A \in [0, \overline{\sigma}_A]$ through a clock-change argument. Since

$$E_0[\overline{X}_t] = E_0^R[\overline{X}_t/\overline{A}_t] = E_0^{QR}[x_t/a_t],$$

it follows that $E_0[\overline{X}_t]$ converges exponentially fast to the mean $\overline{x}$ of x/a under the invariant distribution, and that this is uniform in σ_A. Recalling that $h(u, y) = E^{QR}[x_u/a_u | x_0/a_0 = e^y]$, we also have $h(u, y)$ converge at the same rate to $\overline{x}$ as $u \to \infty$, uniformly in y and σ_A. Letting $F_t^{QR}[y|y_0]$ denote the cross-sectional distribution of y_t given initial value y_0, fixing some $s < u$, we can write

$$h_{y_0}(u, y_0) = \frac{\partial}{\partial y_0} \int_L^U h(u-s, y)\, dF_t^{QR}[y|y_0]$$
$$= \int_L^U h(u-s, y) \frac{\partial^2 F_t^{QR}[y|y_0]}{\partial y_0\, \partial y}\, dy$$
$$= \int_L^U \big(h(u-s, y) - \overline{x}\big) \frac{\partial^2 F_t^{QR}[y|y_0]}{\partial y_0\, \partial y}\, dy,$$

where at the last step we used that $\partial^2 F_t^{QR}[y|y_0]/\partial y_0\,\partial y$ integrates to zero in y. By the arguments of Lemma 1, $\partial^2 F_t^{QR}[y|y_0]/\partial y_0\,\partial y$ is bounded, while $h(u-s, y)-\overline{x}$ converges exponentially fast to zero; hence so does the integral. *Q.E.D.*

PROOF OF PROPOSITION 4: We show that $\xi(u, f)$ equals the impulse response of Definition 1. Let $\overline{A}_0^*$ be the point at which we want to differentiate $E_0[\overline{X}_t(A_0, F^x(x_0|A_0^*))]$. We can write

$$\begin{aligned} &E_0\big[\overline{X}_t\big(A_0, F^x\big(x_0|A_0^*\big)\big)\big] \\ &\quad = A_0 \cdot E_0^R\big[\overline{X}_t\big(A_0, F^x\big(x_0|A_0^*\big)\big)_t/\overline{A}_t\big] \\ &\quad = A_0 \cdot \int_L^U h\big(t, y - \big(\log A_0 - \log A_0^*\big)\big)\, dF_0(y). \end{aligned}$$

This is because when $\overline{A}_0 = \overline{A}_0^*$, the mass of people at any point y is given by $dF_0(y)$, and the conditional expectation given y is summarized by h. When $\overline{A}_0$ changes, the mass of these people is unaffected, and hence $dF_0(y)$ is unchanged; but—because commitments are held fixed while A_0 changes—their y shifts. Hence we must evaluate h at a point which recognizes this change.

Differentiating this expression in A_0 gives

$$\begin{aligned} \frac{E_0\big[\overline{X}_t\big(A_0, F^x\big(x_0|A_0^*\big)\big)\big]}{\partial A_0} &= \int_L^U h(t, y)\, dF_0(y) - \int_L^U h_y(t, y)\, dF_0(y) \\ &= \int_L^U \big[h(t, y) - h_y(t, y)\big]\, dF_0(y), \end{aligned}$$

which is exactly the definition of ξ given above when $F_0(y)$ has a density. This confirms that the impulse response is well defined, that it is independent of A_0^*, and that the MA representation claimed in the proposition holds. *Q.E.D.*

PROOF OF LEMMA 4: We know that EF converges to F^* uniformly in y. Fix $\varepsilon > 0$ and pick s so that, for all $t > s$, $|EF_t - F^*| < \varepsilon/8$ for all initial conditions and for all σ small enough. Consider the rectangular set $[-\kappa, \kappa] \times [t-s, t]$, and let G_κ denote the event when the realization of $\log \overline{A}_u - \log \overline{A}_{t-s}$ for $u \in [t-s, t]$ is in this set. Let $F(y, t, \overline{A}_{[t-s,t]}, y_s)$ denote the distribution of y_t under Q when started at y_s in s, and when the realization of aggregate shocks is given by $\overline{A}_{[t-s,t]}$. We then have that $\{\sup_{y_t, y_s} |F(y, t, \overline{A}_{[t-s,t]}, y_s) - F(y, t, \overline{A}'_{[t-s,t]}, y_s)| \overline{A}_{[t-s,t]}, \overline{A}'_{[t-s,t]} \in G_\kappa\}$ goes to zero as $\kappa \to 0$: two sufficiently close paths of aggregate consumption generate cross-sectional distributions that are themselves close. This is because the share of people for whom the two aggregate paths result in sufficiently different behavior goes to zero. Take κ small enough so that this quantity is less than $\varepsilon/8$. For any fixed κ, we can pick σ small enough so that $\Pr[\overline{A}_{[t-s,t]} \in G_\kappa] > 1 - \varepsilon/8$. This implies that $|E_s F_t - E[F_t | f(s), G_\kappa]| < \varepsilon/4$. Combining these bounds, for $\overline{A}_{[t-s,t]} \in G_\kappa$ we have

$$\begin{aligned}
&\left|F\left(y, t, \overline{A}_{[t-s,t]}, f(s)\right) - F^*(y)\right| \\
&\quad \le \left|F\left(y, t, \overline{A}_{[t-s,t]}, f(s)\right) - E\left[F_t | f(s), G_\kappa\right]\right| \\
&\qquad + \left|E\left[F_t | f(s), G_\kappa\right] - E_s F_t\right| + \left|E_s F_t - F^*(y)\right| \\
&\quad < \frac{\varepsilon}{8} + \frac{\varepsilon}{4} + \frac{\varepsilon}{8} = \frac{\varepsilon}{2}.
\end{aligned}$$

Using this, we have

$$\begin{aligned}
&\left\| \sup_y \left|F(y, t) - F^*(y)\right| \right\|_p^p \\
&\quad = \Pr[G_\kappa] \cdot E\left[\sup_y \left(F(y, t) - F^*(y)\right)^p \Big| G_\kappa\right] \\
&\qquad + \left(1 - \Pr[G_\kappa]\right) \cdot E\left[\sup_y \left(F(y, t) - F^*(y)\right)^p \Big| \text{not } G_\kappa\right] \\
&\quad \le \left[\left(\frac{\varepsilon}{2}\right)^p + 2^p \frac{\varepsilon}{8}\right] < 2^p \varepsilon.
\end{aligned}$$

Since this is true for all $t > s$, it is also true for the lim sup. But ε was arbitrary, and the bound applies for all σ small enough given ε; hence the desired result follows. *Q.E.D.*

S.2.2. *Proofs of Auxiliary Results About the Habit Model*

PROOF OF LEMMA 5: Starting with the A-weighted habit model, consider the unique solution of the integral equations for ζ and o (see Lew (1972) for existence and uniqueness) and define

$$\widetilde{X}_t = o(t) X_0 + \int_0^t \zeta(t - s) C_s \, ds.$$

We will show that $\widetilde{X}_t = X_t$ for all $t \ge 0$. First note that

$$\begin{aligned}
\widetilde{X}_t &= o(t) X_0 + \int_0^t \zeta(t - s)[A_s + X_s] \, ds \\
&= o(t) X_0 \\
&\quad + \int_0^t \zeta(t - s) A_s + \zeta(t - s)\left[\int_0^s j(s - u) A_u \, du + k(s) X_0\right] ds
\end{aligned}$$

$$= o(t)X_0 + \int_0^t A_s\left[\zeta(t-s) + \int_0^{t-s} j(u)\zeta(t-s-u)\,du\right]ds$$
$$+ X_0\int_0^t \zeta(t-s)k(s)\,ds.$$

Equating coefficients, $X_t = \widetilde{X}_t$ holds if

$$j(t-s) = \zeta(t-s) + \int_0^{t-s} j(u)\zeta(t-s-u)\,du$$

or, with $t - s = u$,

$$\zeta(u) = j(u) - \int_0^u \zeta(v)j(u-v)\,dv$$

and

$$o(u) = k(u) - \int_0^u \zeta(u-v)k(v)\,dv.$$

Substituting in $u = 0$ gives $\zeta(0) = j(0)$ and $o(0) = k(0)$. The integral equation for $\zeta(u)$ then yields a unique solution, which can be used to determine $o(\cdot)$. By the above argument, a pair of functions that solve these equations also give $X_t = \widetilde{X}_t$, which is the desired representation. *Q.E.D.*

PROOF OF LEMMA 6: Detrending both sides and integrating by parts (using that ξ^* is smooth):

$$\overline{X}_t^h = \int_0^t \xi^{*\prime}(t-s)\overline{A}_s\,ds + \left[\overline{x} - \xi^*(t)\right]A_0$$
$$= \left[-\xi^*(t-u)\overline{A}_u\right]_0^t + \int_0^t \xi^*(t-s)\,d\overline{A}_s + \left[\overline{x} - \xi^*(t)\right]A_0$$
$$= \int_0^t \xi^*(t-s)\,d\overline{A}_s + \overline{x}A_0.$$

Q.E.D.

S.2.3. *Proofs of Results Used in Establishing Theorem 1*

PROOF OF LEMMA 7: We proceed by induction on t. Fix some $k > 0$. We show that (i) the desired bound holds when $t \le k$, and (ii) if the bound holds for some t, it also holds for $t + k$. We begin by showing (ii), which is the more difficult part.

We can write

$$\|G_t\|_p \le \left\| \frac{\overline{A}_{t-k}}{\overline{A}_t} \int_{t-k}^{t} g(t-s) \frac{\overline{A}_s}{\overline{A}_{u-k}} \, dz_s \right\|_p + \left\| \frac{\overline{A}_{t-k}}{\overline{A}_t} \right\|_p \cdot \left\| \frac{1}{\overline{A}_{t-k}} \int_0^{t-k} g(t-s) \overline{A}_s \, dz_s \right\|_p,$$

where we used independence of the Brownian increments. Denoting $\overline{g}(u, s) = e^{K_2 k} g(u+k, s)$, we can rewrite the final term in brackets as

$$e^{-K_2 k} \cdot \frac{1}{\overline{A}_{t-k}} \int_0^{t-k} \overline{g}(t-k-s, s) A_s \, dz_s,$$

where $|\overline{g}(u, s)| \le K_1 e^{-K_2 u}$ by construction. By our induction assumption, this term has p-norm bounded by $e^{-K_2 k} \cdot M(p)$. To bound the remaining terms, first observe that by lognormality,

$$\left\| \frac{\overline{A}_{t-k}}{\overline{A}_t} \right\|_p \le K_p(\sigma_A, k)$$

for some $K_p(\sigma_A, k)$ that goes to 1 in σ_A for all k. Next note that

$$\left\| \frac{\overline{A}_{t-k}}{\overline{A}_t} \int_{t-k}^{t} g(t-s, s) \frac{\overline{A}_s}{\overline{A}_{t-k}} \, dz_s \right\|_p \le \left\| \frac{\overline{A}_{t-k}}{\overline{A}_t} \right\|_{2p} \cdot \left\| \int_{t-k}^{t} g(t-s, s) \frac{\overline{A}_s}{\overline{A}_{t-k}} \, dz_s \right\|_{2p}$$

by the Cauchy–Schwarz inequality. Here

$$\left\| \frac{\overline{A}_{t-k}}{\overline{A}_t} \right\|_{2p} \le K_{2p}(\sigma_A, k),$$

where $K_{2p}(\sigma_A, k)$ also goes to 1 in σ_A for all k. Finally, using standard bounds (e.g., Karatzas and Shreve (1988)) for moments of the Ito integral, we obtain

$$\left\| \int_{t-k}^{t} g(t-s, s) \frac{\overline{A}_s}{\overline{A}_{t-k}} \, dz_s \right\|_{2p} \le K_{2p} \left(\int_{t-k}^{t} K_1^2 \left\| \left(\frac{\overline{A}_s}{\overline{A}_{u-k}} \right)^2 \right\|_p ds \right)^{1/2},$$

which is bounded by $K_{2p} K_1 k \cdot K_{2p}(\sigma_A, k)$. Combining terms, we obtain

$$\|G_t\|_p \le K_{2p}^2(\sigma_A, k) \cdot K_{2p} K_1 k + K_p(\sigma_A, k) \cdot e^{-K_2 k} \cdot M(p).$$

It is easy to see that if

$$M(p) = \frac{K_{2p}^2(\sigma_A, k) \cdot K_{2p} K_1 k}{1 - K_p(\sigma_A, k) \cdot e^{-K_2 k}}$$

is positive, then the induction step follows. We can make sure that this is the case by first choosing some $k > 0$, and then picking $\overline{\sigma}_A$ small enough so that, for all $\sigma_A \leq \overline{\sigma}_A$, we have $K_p(\sigma_A, k) < e^{K_2 k/2}$. With this choice of $M(p)$, the induction step follows; and (i) can be verified easily from the argument of the induction step. *Q.E.D.*

PROOF OF LEMMA 8: We verify directly that changing the clock is equivalent to rescaling the relevant parameters in the setup of the problem. Maximizing the consumer's problem in the original model is equivalent to maximizing

$$E \int_0^\infty e^{-\rho t \tau} \left(\frac{a_{\tau t}^{1-\gamma}}{1-\gamma} + \mu \frac{x_{\tau t}^{1-\gamma}}{1-\gamma} \right) dt,$$

which is proportional to the objective function in the model with new parameters. Similarly, the budget constraint of the original model implies

$$\begin{aligned} dw_{\tau t} = {} & \left[\left(\tau r + \alpha_{\tau t} \tau \pi + \alpha_{\tau t}^i \tau \pi_I \right) w_t - \tau c_t \right] dt \\ & + \alpha_{\tau t} w_{\tau t} \sigma \tau^{1/2} \, dz_{\tau t} + \alpha_{\tau t}^i w_{\tau t} \sigma_i \tau^{1/2} \, dz_{\tau t}^i \end{aligned}$$

on all non-adjustment dates due to the scaling invariance of Brownian motion. Finally, on adjustment dates, $dw = \overline{\lambda}_1 x_{t-}/r + \overline{\lambda}_2 x_t / r = \overline{\lambda}_1 \cdot \tau x_{t-}/(\tau r) + \overline{\lambda}_2 \cdot \tau x_t/(\tau r)$. Since the optimal policy is unique, the claim follows. *Q.E.D.*

S.3. PROOFS FOR SECTION 4.1

S.3.1. *Proof of Proposition 5*

(1) Excess smoothness. Using a Taylor expression, we can write

$$\log C_{t_1} - \log C_{t_0} = \frac{A_{t_0}}{C_{t_0}} (\log A_{t_1} - \log A_{t_0}) + \varepsilon_{t_1}, \tag{19}$$

where, because X_t has bounded variation, there exists K_ε such that $E\varepsilon_{t_1}^2 < K_e(t_1 - t_0)^2$. Thus

$$\begin{aligned} \beta_1(t_1) &= \frac{\operatorname{cov}\left(\log(C_{t_1}/C_{t_0}), \log(A_{t_1}/A_{t_0})\right)}{\operatorname{var}\left(\log(A_{t_1}/A_{t_0})\right)} \\ &\leq \frac{A_{t_0}}{C_{t_0}} + \frac{\sigma_A (t_1 - t_0)^{1/2} K_\varepsilon (t_1 - t_0)}{\sigma_A^2 (t_1 - t_0)} = \frac{A_{t_0}}{C_{t_0}} + \frac{(t_1 - t_0)^{1/2} K_\varepsilon}{\sigma_A} \end{aligned}$$

and the right-hand side approaches A_{t_0}/C_{t_0} as $t_1 \to t_0$.

(2) Excess sensitivity. Let $t_1 = t_2$. From the proof of Lemma 3, we know that $\log[A_{t_3}/C_{t_3}]$ converges exponentially fast to an invariant distribution. In particular, $E_{t_1}[\log[A_{t_3}/C_{t_3}]]$ converges exponentially fast to the mean of this invariant distribution, which we denote by $\overline{\xi}$, so that we can write $\log C_{t_3} = \log A_{t_3} + \overline{\xi} + \varepsilon_{t_3}$, where $E_{t_1}[\varepsilon_{t_3}]$ converges to zero at a given exponential rate as $t_3 \to \infty$. Using (19), we can write

$$\begin{aligned}\log C_{t_3} - \log C_{t_1} &= \log A_{t_3} + \overline{\xi} + \varepsilon_{t_3} - \log C_{t_0} \\ &\quad - \frac{A_{t_0}}{C_{t_0}}(\log A_{t_1} - \log A_{t_0}) - \varepsilon_{t_1} \\ &= \frac{X_{t_0}}{C_{t_0}}(\log A_{t_1} - \log A_{t_0}) + (\log A_{t_3} - \log A_{t_1}) \\ &\quad + (\log A_{t_0} + \overline{\xi}) + (\varepsilon_{t_3} - \varepsilon_{t_1}).\end{aligned}$$

To compute β_2, we evaluate the covariance of $\log A_{t_1} - \log A_{t_0}$ with each of the terms in this expression. Because $\log A_t$ is a Brownian motion with drift, the covariance with the term in the second parentheses is zero. Conditional on the history up to t_0, the terms in the third parentheses are constants, hence their covariance is also zero. The terms in the fourth parentheses are error terms: just like in the proof of (1), ε_1 can be made arbitrarily small by choosing t_1 small; and ε_{t_3} is approximately orthogonal to events before t_1 for t_3 large. Thus, for t_1 small and t_3 large, the regression coefficient is determined by the first term, implying that β_2 is approximately $X_{t_0}/C_{t_0} > 0$.

S.3.2. *Modeling Large Shocks*

Our approach is to construct, on a single probability space, a set of "shock" processes for each $t_1 > t_0$, such that the distribution of the process for a given t_1 is identical to the distribution of $\overline{A}_t$ conditional on the shock event $S(t_1, \Delta)$. This construct will allow us to take limits while holding fixed the probability space.

Formally, we introduce the auxiliary process $\widetilde{A}_t$, which agrees with $\overline{A}_t$ for $t \le t_0$, and has the same distribution as $\overline{A}_t$ for $t > t_0$. The idea is that innovations in $\widetilde{A}_t$ will be driving $\overline{A}_t$ after the shock. We also introduce an independent standard Brownian motion $\overline{B}_s$ defined for $s \ge 0$, which will drive the innovations during the shock. We then model the positive shock as a Brownian bridge for $\log \overline{A}_t$ conditioned to start at $\log \overline{A}_{t_0}$ at time t_0, and to reach $\log \overline{A}_{t_0} + \Delta$ at time t_1. We denote this process by $\overline{A}_t(+, t_1, \Delta)$, and construct it as follows: for $t_0 \le t \le t_1$, we let $\log \overline{A}_t(+, t_1, \Delta) = \sigma_A(\overline{B}_{t-t_0} - (t - t_0)\overline{B}_{t_1}) + (t - t_0)\Delta$, and for $t \ge t_1$, we let $d \log \overline{A}_t(+, t_1, \Delta) = d \log \widetilde{A}_t$. Although the expression for

$t_0 \le t \le t_1$ does not make this clear, it is well known that this Brownian bridge is an Ito processes. We construct $\overline{A}_t(-, t_1, \Delta)$ analogously. Given that it is a Brownian bridge between $t_0 \le t \le t_1$, it follows that $\log \overline{A}_t(+, t_1, \Delta)$ has the same distribution as our original process $\log \overline{A}_t$ conditional on $S(+, t_1, \Delta)$.

The formulas for the dynamics of X_t, C_t, X_t^h, and C_t^h, once we replace $\overline{A}_t$ by $\overline{A}_t(+, t_1, \Delta)$ respectively $\overline{A}_t(-, t_1, \Delta)$, directly extend, and generate the distributions of commitments, habit, and consumption conditional on the shock event. To clarify which process we have in mind, we sometimes use notation such as $\overline{X}_t(+, t_1, \Delta)$ to refer to aggregate commitments (during or after a positive shock) on the probability space just constructed. However, when it does not cause confusion we often just write $\overline{X}_t^h$ and say in words that we work with the "shock" processes.

One key feature of this construction is that instead of considering a sequence of non-overlapping events $S(t_1, \Delta)$, we consider a single probability space and a sequence of processes. The advantage is that we can use the L_p norm on this common probability space when we take various limits over $\bar{t}$. In particular, throughout the analysis below, we use L_p (conditional on the history up to t_0) for all $p \ge 1$ as we take the limits $t_1 \to t_0$ and $t_2 \to t_0$.

S.3.3. *Continuity After Large Shocks*

We show that $\overline{X}_t^h$ and $\overline{C}_t^h$ change continuously around t_0 in the limit as $t_1 \to t_0$ and as $t_2 \to t_0$.

LEMMA 9: *We have*

$$\lim_{t_1 \to t_0} \overline{X}_{t_1}^h(+, t_1, \Delta) = \overline{X}_{t_0}^h \quad \textit{and} \quad \lim_{t_1 \to t_0} \overline{X}_{t_1}^h(-, t_1, \Delta) = \overline{X}_{t_0}^h.$$

Moreover, even after taking the limit $t_1 \to t_0$, the dynamics of $\overline{X}_t^h$ are continuous at t_0:

$$\lim_{t_2 \to t_0} \lim_{t_1 \to t_0} \overline{X}_{t_2}^h(+, t_1, \Delta) = \overline{X}_{t_0}^h \quad \textit{and} \quad \lim_{t_2 \to t_0} \lim_{t_1 \to t_0} \overline{X}_{t_2}^h(-, t_1, \Delta) = \overline{X}_{t_0}^h.$$

PROOF: Consider the case when the shock is positive. Suppressing in notation that we work with the "shock" processes, according to the representation in Lemma 7, $\overline{X}_{t_2}^h = \int_0^{t_2} \xi^{*\prime}(t_2 - s)\overline{A}_s(+, t_1, \Delta)\, ds + [\bar{x} - \xi^*(t_2)]A_0$. When $t_2 = t_1$ goes to t_0, this expression converges to $\int_0^{t_0} \xi^{*\prime}(t_2 - s)\overline{A}_s(+, t_1, \Delta)\, ds + [\bar{x} - \xi^*(t_2)]A_0 = \overline{X}_{t_0}^h$, proving, for a positive shock, the first claim. For the second claim, note that as $t_1 \to t_0$, the last term is constant while the first term converges to $\int_{t_0}^{t_2} \xi^{*\prime}(t_2 - s)\widetilde{A}_s \cdot e^{\Delta}\, ds + \int_0^{t_0} \xi^{*\prime}(t_2 - s)\overline{A}_s\, ds$. Here only the first integral depends on t_2, and as $t_2 \to t_0$, it converges to zero. The same logic works when the shock is negative. *Q.E.D.*

LEMMA 10: *We have*

$$\lim_{t_1\to t_0}\log\bigl[\overline{C}^h_{t_1}(+,t_1,\Delta)\bigr]=\log\bigl[e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}\bigr]\quad \textit{and}$$

$$\lim_{t_1\to t_0}\log\bigl[\overline{C}^h_{t_1}(-,t_1,\Delta)\bigr]=\log\bigl[e^{-\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}\bigr].$$

And analogously we have

$$\lim_{t_2\to t_0}\lim_{t_1\to t_0}\log\bigl[\overline{C}^h_{t_2}(+,t_1,\Delta)\bigr]=\log\bigl[e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}\bigr]\quad \textit{and}$$

$$\lim_{t_2\to t_0}\lim_{t_1\to t_0}\log\bigl[\overline{C}^h_{t_2}(-,t_1,\Delta)\bigr]=\log\bigl[e^{-\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}\bigr].$$

PROOF: Suppose the shock is positive. Then, suppressing in notation that we work with the "shock" processes, using the fact that $\log(1+z)\le z$,

$$\begin{aligned}
&\bigl|\log\bigl[\overline{C}^h_{t_2}\bigr]-\log\bigl[e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}\bigr]\bigr|\\
&=\left|\log\left[\frac{\overline{A}_{t_2}+\overline{X}^h_{t_2}}{e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}}\right]\right|\le\max\left[\frac{\overline{A}_{t_2}+\overline{X}^h_{t_2}}{e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}}-1,\ \frac{e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}}{\overline{A}_{t_2}+\overline{X}^h_{t_2}}-1\right]\\
&\le\max\left[\frac{(\overline{A}_{t_2}-e^{\Delta}\overline{A}_{t_0})+(\overline{X}^h_{t_2}-\overline{X}^h_{t_0})}{e^{\Delta}\overline{A}_{t_0}+\overline{X}^h_{t_0}},\ \frac{(e^{\Delta}\overline{A}_{t_0}-\overline{A}_{t_2})+(\overline{X}^h_{t_0}-\overline{X}^h_{t_2})}{\overline{A}_{t_2}+\overline{X}^h_{t_2}}\right]\\
&\le\max\left[\frac{|\overline{A}_{t_2}-e^{\Delta}\overline{A}_{t_0}|+|\overline{X}^h_{t_2}-\overline{X}^h_{t_0}|}{e^{\Delta}\overline{A}_{t_0}},\ \frac{|e^{\Delta}\overline{A}_{t_0}-\overline{A}_{t_2}|+|\overline{X}^h_{t_0}-\overline{X}^h_{t_2}|}{\overline{A}_{t_2}}\right].
\end{aligned}$$

For the first set of limits, we assume $t_1=t_2$ and take them to t_0 simultaneously; for the second set of limits, we first take $t_1\to t_0$ and then take $t_2\to t_0$. In either case, in both terms of the maximum, the numerator converges to zero in L_{2p} while the inverse of the denominator is bounded in L_{2p}. By the Cauchy–Schwarz inequality, the terms themselves converge to zero in L_p, hence so does their maximum. The argument for a negative shock is analogous. *Q.E.D.*

S.3.4. *Notation and Proof Structure*

Bounds

We use the notation that $K(\bar{t},\Delta)$ refers to a family of random variables which are uniformly bounded independently of Δ, in the limit as $t_1\to t_0$, when t_2 and t_3 are appropriately chosen. Formally, we require that there exists a family of constants $K(p)$, such that given p, for any Δ, we can find $t_2(\Delta,p)$ small enough and $t_3(\Delta, p)$ large enough so that $\lim_{t_1\to t_0}\sup\|K(t_1, t_2(\Delta, p), t_3(\Delta, p))\|_p \le K(p)$. Different occurrences of $K(\bar{t}, \Delta)$ may refer to different families of random variables and may have different $K(p)$ values associated with them. For example, Lemma 10 implies that $\log[\overline{C}^h_{t_2}(+, t_1, \Delta)] = \log[e^{\Delta}\overline{A}_{t_0} + \overline{X}^h_{t_0}] + K(\bar{t}, \Delta)$.

Order of Limits

The statement of Proposition 7 assumes that n is large enough; this means that σ_A/σ_I is small enough, while other parameters of the model, as described in Section 3.4, remain bounded. We first analyze the case in which σ_A becomes small, and then establish the result when σ_I becomes sufficiently large using a clock change.

S.3.5. *Long-Term Behavior*

LEMMA 11: *Suppose that n is large enough and σ_A is small enough. Then*

$$\lim_{t_3\to\infty}\lim_{t_1\to t_0}\left[\frac{\overline{X}^h_{t_3}(-, t_1, \Delta)}{\overline{A}_{t_3}(-, t_1, \Delta)} - \frac{\widetilde{X}^h_{t_3}}{\widetilde{A}_{t_3}}\right] = 0.$$

The intuition for the lemma is that X_{t_3} is just a weighted sum of past A_s values, with the weights for the distant past going to zero exponentially fast. Thus, if A_s is multiplied by a constant after date t_0, then for t_3 large enough, most of the terms determining X_{t_3} in this weighted sum will also be multiplied by that constant, and hence X_{t_3}/A_{t_3} will be approximately the same as it would be on the no-shock path. The caveat is that the terms in the weighted average corresponding to the distant past, divided by current A_{t_3}, must not blow up. For this we need that $1/A_{t_3}$ does not become big too quickly relative to the rate with which the weights on the past converge to zero. These weights go to zero at a given exponential rate, so if the variance of the A_t process is not too big, we are fine.

PROOF OF LEMMA 11: Suppressing in notation that we work with the "negative shock" processes, we have

$$\begin{aligned}
&\lim_{t_1\to t_0}\frac{\overline{X}^h_{t_3}}{\overline{A}_{t_3}} \\
&\quad = \lim_{t_1\to t_0}\frac{1}{\overline{A}_{t_3}}\int_0^{t_3}\xi^{*\prime}(t_3 - s)\overline{A}_s\,ds + \left[\bar{x} - \xi^*(t_3)\right]\frac{\overline{A}_0}{\overline{A}_{t_3}} \\
&\quad = \frac{1}{\overline{A}_{t_3}}\int_{t_0}^{t_3}\xi^{*\prime}(t_3 - s)\widetilde{A}_s e^{-\Delta}\,ds
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\overline{A}_{t_3}}\int_0^{t_0}\xi^{*\prime}(t_3-s)\overline{A}_s\,ds+\big[\overline{x}-\xi^*(t_3)\big]\frac{\overline{A}_0}{\overline{A}_{t_3}}\\
&=e^{-\Delta}\frac{\widetilde{X}^h_{t_3}}{\overline{A}_{t_3}}+\big(1-e^{-\Delta}\big)\frac{1}{\overline{A}_{t_3}}\left(\int_0^{t_0}\xi^{*\prime}(t_3-s)\overline{A}_s\,ds+\big[\overline{x}-\xi^*(t_3)\big]\overline{A}_0\right)\\
&=\frac{\widetilde{X}^h_{t_3}}{\widetilde{A}_{t_3}}+\big(1-e^{-\Delta}\big)\frac{1}{\widetilde{A}_{t_3}e^{-\Delta}}\Bigg(\big[-\xi^*(t_3-s)\overline{A}_s\big]_0^{t_0}\\
&\quad+\int_0^{t_0}\xi^*(t_3-s)\,d\overline{A}_s+\big[\overline{x}-\xi^*(t_3)\big]\overline{A}_0\Bigg)\\
&=\frac{\widetilde{X}^h_{t_3}}{\widetilde{A}_{t_3}}+\big(e^{\Delta}-1\big)\frac{1}{\widetilde{A}_{t_3}}\Bigg(-\xi^*(t_3-t_0)\overline{A}_{t_0}+\xi^*(t_3)\overline{A}_0\\
&\quad+\int_0^{t_0}\xi^*(t_3-s)\,d\overline{A}_s+\big[\overline{x}-\xi^*(t_3)\big]\overline{A}_0\Bigg)\\
&=\frac{\widetilde{X}^h_{t_3}}{\widetilde{A}_{t_3}}+\big(e^{\Delta}-1\big)\frac{1}{\widetilde{A}_{t_3}}\left(\int_0^{t_0}\xi^*(t_3-s)\,d\overline{A}_s+\overline{x}\overline{A}_0-\xi^*(t_3-t_0)\overline{A}_{t_0}\right).
\end{aligned}
$$

Here the last term can be written as

$$
\big(e^{\Delta}-1\big)\frac{1}{\widetilde{A}_{t_3}}\left(\int_0^{t_0}\big(\xi^*(t_3-s)-\overline{x}\big)\,d\overline{A}_s+\overline{A}_{t_0}\big[\overline{x}-\xi^*(t_3-t_0)\big]\right).
$$

Because, by Lemma 3, $|\xi^*(t_3-s)-\overline{x}|\le K_1e^{-K_2(t_3-s)}$ for some constants K_1, K_2 independent of n, it follows from Lemma 7 that, for n large enough, the first term here converges to zero as $t_3\to\infty$. Also by Lemma 3, the second term converges to zero as $t_3\to\infty$. *Q.E.D.*

LEMMA 12: *Suppose that n is large enough and σ_A is small enough. There exists a constant K_2 such that the following holds. For any Δ, we can find t_2 and t_3 such that, for all t_1 close enough to t_0,*

$$
\begin{aligned}
&E\big[\log\overline{C}^h_{t_3}-\log\overline{C}^h_{t_2}|S(+,t_1,\Delta)\big]-E\big[\log\overline{C}^h_{t_3}-\log\overline{C}^h_{t_2}|S(-,t_1,\Delta)\big]\\
&\quad\ge\Delta-K_2.
\end{aligned}
$$

PROOF: A key element of the proof is that we bound the left-hand side for each realization, that is, without the expectations operator. However, because $S(+,t_1,\Delta)$ and $S(-,t_1,\Delta)$ are disjoint events, we can only do this using the "shock processes," which have the same distribution as the original processes conditioned on the shock events, but are defined on a common probability space.

Suppose first that the shock is positive. Suppressing in notation that we work with the shock process, we have $\log \overline{C}^h_{t_3} \geq \log \overline{A}_{t_3} = \log \widetilde{A}_{t_3} + \Delta$. Moreover, by Lemma 10, for t_2 close to t_0, we have

$$\log \overline{C}^h_{t_2} = \log\bigl[e^{\Delta}\overline{A}_{t_0} + \overline{X}_{t_0}\bigr] + K(\bar{t}, \Delta) = \log \overline{A}_{t_0} + \Delta + K(\bar{t}, \Delta),$$

where the second equality follows because, given that we condition on the history up to t_0, $\overline{X}^h_{t_0}/\overline{A}_{t_0}$ is a constant. We can now write, for a positive shock, that

$$\begin{aligned}\log \overline{C}^h_{t_3} - \log \overline{C}^h_{t_2} &\geq (\log \widetilde{A}_{t_3} + \Delta) - \bigl(\log \overline{A}_{t_0} + \Delta + K(\bar{t}, \Delta)\bigr)\\ &= \log \widetilde{A}_{t_3} - \log \overline{A}_{t_0} + K(\bar{t}, \Delta).\end{aligned}$$

Now suppose that the shock is negative. Then, using Lemma 10,

$$\begin{aligned}\log \overline{C}^h_{t_2} &= \log\bigl[e^{-\Delta}\overline{A}_{t_0} + \overline{X}_{t_0}\bigr] + K(\bar{t}, \Delta)\\ &= \log \overline{A}_{t_0} + \log\bigl[e^{-\Delta} + \overline{X}_{t_0}/\overline{A}_{t_0}\bigr] + K(\bar{t}, \Delta) \geq \log \overline{A}_{t_0} + K(\bar{t}, \Delta)\end{aligned}$$

because $\overline{X}_{t_0}/\overline{A}_{t_0}$ is a constant. Moreover, using the fact that $\log(1+z) \leq z$,

$$\begin{aligned}\log \overline{C}^h_{t_3} &= \log\bigl[\overline{A}_{t_3} + \overline{X}^h_{t_3}\bigr] = \log \overline{A}_{t_3} + \log\bigl[1 + \overline{X}^h_{t_3}/\overline{A}_{t_3}\bigr]\\ &\leq \log \widetilde{A}_{t_3} - \Delta + \overline{X}^h_{t_3}/\overline{A}_{t_3} = \log \widetilde{A}_{t_3} - \Delta + \widetilde{X}^h_{t_3}/\widetilde{A}_{t_3} + K(\bar{t}, \Delta),\end{aligned}$$

where at the last step we used Lemma 11. It follows that for a negative shock,

$$\log \overline{C}^h_{t_3} - \log \overline{C}^h_{t_2} \leq \log \widetilde{A}_{t_3} - \Delta + \widetilde{X}^h_{t_3}/\widetilde{A}_{t_3} - \log \overline{A}_{t_0} + K(\bar{t}, \Delta).$$

Combining the inequalities for the positive and the negative shocks yields, for the shock processes, the bound

$$\begin{aligned}&\bigl[\log \overline{C}^h_{t_3}(+, t_1, \Delta) - \log \overline{C}^h_{t_2}(+, t_1, \Delta)\bigr]\\ &\qquad - \bigl[\log \overline{C}^h_{t_3}(-, t_1, \Delta) - \log \overline{C}^h_{t_2}(-, t_1, \Delta)\bigr]\\ &\quad \geq \log \widetilde{A}_{t_3} - \log \overline{A}_{t_0} - \bigl(\log \widetilde{A}_{t_3} - \Delta + \widetilde{X}^h_{t_3}/\widetilde{A}_{t_3} - \log \overline{A}_{t_0}\bigr) + K(\bar{t}, \Delta)\\ &\quad = \Delta - \widetilde{X}^h_{t_3}/\widetilde{A}_{t_3} + K(\bar{t}, \Delta).\end{aligned}$$

Finally,

$$\frac{\widetilde{X}_{t_3}}{\widetilde{A}_{t_3}} = \frac{1}{\widetilde{A}_{t_3}} \int_0^{t_3} \xi^*(t_3 - s)\, d\widetilde{A}_s + \bar{x}\frac{\widetilde{A}_0}{\widetilde{A}_{t_3}}$$

$$= \bar{x} + \frac{1}{\widetilde{A}_{t_3}} \int_0^{t_3} \left[\xi^*(t_3 - s) - \bar{x}\right] d\widetilde{A}_s,$$

and by Lemma 7, the last term is bounded in L_p for all t_3. Thus the above difference is Δ plus a term bounded in L_p, and the claim of the lemma follows. *Q.E.D.*

S.3.6. *Proofs of Propositions 6 and 7*

PROOF OF PROPOSITION 6: (i) Taking expectations in the regression equation (12) conditional on the shock being positive, respectively negative, and differencing, we obtain

$$\begin{aligned}(20)\qquad & E\left[\log \overline{C}_{t_1} - \log \overline{C}_{t_0} | S(+, t_1, \Delta)\right] - E\left[\log \overline{C}_{t_1} - \log \overline{C}_{t_0} | S(-, t_1, \Delta)\right] \\ & \quad = 2\beta_1(t_1, \Delta) \cdot \Delta,\end{aligned}$$

which gives an expression for $\beta_1(t_1, \Delta)$. An analogous formula expresses $\beta_1^h(t_1, \Delta)$. Because X_t/A_t is bounded from below by L and from above by U, we have $|\log(\overline{C}_{t_1}/\overline{C}_{t_0}) - \log(\overline{A}_{t_1}/\overline{A}_{t_0})| \leq \log(1 + U) - \log(1 + L) = K_1$, and therefore,

$$\begin{aligned}& E\left[\log \overline{C}_{t_1} - \log \overline{C}_{t_0} | S(+, t_1, \Delta)\right] - E\left[\log \overline{C}_{t_1} - \log \overline{C}_{t_0} | S(-, t_1, \Delta)\right] \\ & \quad \geq E\left[\log(\overline{A}_{t_1}/\overline{A}_{t_0}) | S(+, t_1, \Delta)\right] \\ & \qquad - E\left[\log(\overline{A}_{t_1}/\overline{A}_{t_0}) | S(-, t_1, \Delta)\right] - 2K_1 \\ & \quad = 2(\Delta - K_1).\end{aligned}$$

Hence $\beta_1(t_1, \Delta) \geq 1 - K_1/\Delta$.

(ii) Lemma 10 implies that for any positive K_2, we can choose t_1 close enough to t_0 such that

$$\begin{aligned}& E\left[\log \overline{C}_{t_1} | S(+, t_1, \Delta)\right] - E\left[\log \overline{C}_{t_1} | S(-, t_1, \Delta)\right] \\ & \quad \leq \log\left[e^{\Delta}\overline{A}_{t_0} + \overline{X}^h_{t_0}\right] - \log\left[e^{-\Delta}\overline{A}_{t_0} + \overline{X}^h_{t_0}\right] + K_2.\end{aligned}$$

The right-hand side can be bounded as

$$\begin{aligned}\log\left[\frac{e^{\Delta}\overline{A}_{t_0} + \overline{X}^h_{t_0}}{e^{-\Delta}\overline{A}_{t_0} + \overline{X}^h_{t_0}}\right] &= \log\left[\frac{e^{\Delta} + \overline{X}^h_{t_0}/\overline{A}_{t_0}}{e^{-\Delta} + \overline{X}^h_{t_0}/\overline{A}_{t_0}}\right] \leq \log\left[\frac{e^{\Delta} + \overline{X}^h_{t_0}/\overline{A}_{t_0}}{\overline{X}^h_{t_0}/\overline{A}_{t_0}}\right] \\ & \leq \Delta + \log\left[\frac{1 + \overline{X}^h_{t_0}/\overline{A}_{t_0}}{\overline{X}^h_{t_0}/\overline{A}_{t_0}}\right] = \Delta + K_3,\end{aligned}$$

where—given that we condition on the history up to t_0—K_3 is a constant. It then follows from (20) that, for a given Δ, we can choose t_1 close enough to t_0 such that $\beta_1^h(t_1, \Delta) < 1/2 + (K_2 + K_3)/\Delta$. *Q.E.D.*

PROOF OF PROPOSITION 7: (i) Taking expectations in (13) and differencing, we obtain

$$E\big[\log \overline{C}_{t_3} - \log \overline{C}_{t_2} | S(+, t_1, \Delta)\big] - E\big[\log \overline{C}_{t_3} - \log \overline{C}_{t_2} | S(-, t_1, \Delta)\big] = 2\beta_2(\bar{t}, \Delta) \cdot \Delta, \tag{21}$$

which gives an expression for $\beta_2(\bar{t}, \Delta)$. An analogous formula expresses $\beta_2^h(\bar{t}, \Delta)$. Because X_t/A_t is bounded from below by L and from above by U, we have $|\log(\overline{C}_{t_3}/\overline{C}_{t_2}) - \log(\overline{A}_{t_3}/\overline{A}_{t_2})| \le \log(1 + U) - \log(1 + L) = K_1$, and therefore,

$$\begin{aligned} &E\big[\log \overline{C}_{t_3} - \log \overline{C}_{t_2} | S(+, t_1, \Delta)\big] - E\big[\log \overline{C}_{t_3} - \log \overline{C}_{t_2} | S(-, t_1, \Delta)\big] \\ &\quad \le E\big[\log(\overline{A}_{t_3}/\overline{A}_{t_2}) | S(+, t_1, \Delta)\big] \\ &\qquad - E\big[\log(\overline{A}_{t_3}/\overline{A}_{t_2}) | S(-, t_1, \Delta)\big] + 2K_1 \\ &\quad = 2K_1. \end{aligned}$$

Using (21), we obtain $\beta_2(t_1, t_2, t_3, \Delta) \le K_1/\Delta$.

(ii) Using Lemma 12, we can find t_2 and t_3, and t_1 close enough to t_0, such that $\beta_2^h(t_1, t_2, t_3, \Delta) \ge 1 - K_2/\Delta$. This gives the proof along a sequence Θ_n in which $\sigma_A \to 0$. Finally, we discuss the case when, as $n \to \infty$, we have $\sigma_I \to \infty$. The only step we need to verify is that Lemma 12 also holds for n large enough. To show this, just like in the proof of our main result, we change the clock. Using the transformation introduced in Lemma 8, we let $\tau = 1/\sigma_I^2$ and slow down the model by rescaling deep parameters with τ. In the habit representation of that "rescaled" model, for n large enough, Lemma 12 holds, because all the assumptions, in particular, the requirement that σ_A is small enough, are satisfied. And because the habit representation of the model after the clock change is the same as changing the clock in the habit representation of the original model, it follows that—with appropriately unscaled values for t_2 and t_3—Lemma 12 also holds in the original model. *Q.E.D.*

S.4. PROOFS FOR SECTIONS 4.2 AND 4.3

PROOF OF PROPOSITION 8: In $\overline{\Theta}^*$, agents in the interior of the band never adjust, hence $T_*(\widetilde{p}|x_0) = \infty$. For n finite, agents do adjust eventually, but since the drift and variance of y go to zero, the expected time to adjustment approaches infinity. In the habit model, x never changes, hence $T^{h,n}(\widetilde{p}|x_0) = \infty$. *Q.E.D.*

PROOF OF PROPOSITION 9: (i) Our first goal is to compute the value function of the habit agent. Let ψ be defined so that the value function of the Merton consumption problem in the environment of the representative habit consumer, but without habit, is $\psi W^{1-\gamma}/(1-\gamma)$. By the envelope theorem, this Merton agent has consumption policy $c = \psi^{-1/\gamma} W$. The surplus consumption of our habit agent is identical to the consumption of a Merton agent, because they solve the same maximization problem. Hence, if the habit consumer sets his initial surplus consumption to be A_0, the dollar cost of his lifetime surplus consumption expenditure is $A_0 \psi^{1/\gamma}$.

To proceed, we now evaluate the lifetime budget constraint of the habit consumer. Each dollar of consumption spending in a period also creates future expenditure in the form of increased habit. Suppose $1 + B$ dollars is the present value of these future expenditures for a dollar of consumption spending today, where $B = 0$ with no habits. Then B must satisfy

$$B = \int_{u=0}^{\infty} \theta(u) e^{-ru}\, du \cdot (1 + B)$$

because each dollar of consumption creates $\theta(u)$ habit spending u periods ahead, which has a total cost of $\theta(u)(1 + B)$ in period u dollars, which we must then discount back at the risk-free rate because these payments are certain. Solving yields

$$B = \frac{1}{1 - \int_{u=0}^{\infty} \theta(u) e^{-ru}\, du}.$$

At any time t, our habit consumer also has pre-existing habit created by his past consumption. The dollar value of the expenditures generated is

$$Z_t = (1 + B) \cdot \left[\int_{s=0}^{t} C_{t-s} \int_{s}^{\infty} \theta(u) e^{-r(u-s)}\, du\, ds + \int_{s=t}^{\infty} \theta_0(u) X_0 e^{-ru}\, du \right],$$

where the term in parentheses measures future consumption expenditures created by habits established before t, discounted back at the risk-free rate because these are certain; and the factor $1 + B$ is included because each dollar of consumption spending has this total expenditure cost.

The consumer's lifetime budget constraint must then satisfy

$$W_t = A_t \cdot \psi^{1/\gamma}(1+B) + Z_t$$

and his lifetime utility from surplus consumption, by the Merton value function, is simply $\psi^{1/\gamma} A_t^{1-\gamma}/(1-\gamma)$. Combining these equations yields

$$V_t^{\text{habit}}(W_t, X_t) = \frac{\psi}{1-\gamma}\left(\frac{W_t - Z_t}{1+B}\right)^{1-\gamma}.$$

The welfare of an individual commitment agent for a move-inducing negative wealth shock is proportional to $(w - \lambda_1 x)^{1-\gamma}/(1-\gamma)$.

Now compare the welfare cost of shocks in the commitment and the habit economies. As wealth falls to zero, if $Z_t > 0$, then the marginal utility of the habit agent will be driven to infinity even with a finite shock. In contrast, when $\lambda_1 = 0$, the marginal utility of the commitment agent only blows up when all his wealth is taken. It follows that for large finite shocks, $\Pi(q, b)$ is higher for the habit agent than in the commitment economy.

(ii) Begin with the commitment model. The agent in the limit economy never moves, and hence his value function is proportional to $(W - x/r)^{1-\gamma}/(1-\gamma)$. It follows that the coefficient of relative risk aversion $\text{CRRA}^*(W_0, x_0) = \gamma W_0/(W_0 - x_0/r)$. Now consider an agent in economy n. Let p_0 denote the total dollar value at date zero of his total commitment expenditures on his current home. Given positive risk and growth, this agent does move eventually, implying $p_0 < x_0/r$. One policy available to this consumer at any wealth W is to maintain his spending and moving patterns on current commitments, and adjust spending proportionally on all other goods relative to the optimal policy with initial wealth W_0. Given that $\lambda_1 = 0$, this policy yields lifetime utility $V_n(W_0, x_0)(W - p_0)^{1-\gamma}/(W_0 - p_0)^{1-\gamma}$. This is a lower bound for the agent's true value function, and both equal $V_n(W_0, x_0)$ at W_0. It follows that the lower bound has higher curvature at W_0. As a result, $\text{CRRA}^n(W_0, x_0) \leq \gamma W_0/(W_0 - p_0)$. Since $p_0 < x_0/r$, we have $\text{CRRA}^n(W_0, x_0) < \text{CRRA}^*(W_0, x_0)$. Hence, for b small, the Arrow–Pratt approximation implies $\Pi^n(q, b) < \Pi^*(q, b)$ uniformly in n.

In the habit model, the value function in every economy is proportional to $(W - x/r)^{1-\gamma}/(1-\gamma)$, and hence $\Pi^{h,n}(q, b) = \Pi^{h*}(q, b)$. *Q.E.D.*

S.5. SIMULATIONS

Solving the Commitments Model

In the simulations, we use an ODE characterization of the optimal policy that builds on a similar characterization for the one-good model by Grossman and Laroque. To develop this ODE, we must study the Bellman equation of the commitment agent. By the mutual fund theorem, the agent will combine the risky assets available to him in fixed proportions, effectively sharing his wealth between the mutual fund and the risk-free asset. Let π_r and σ_r denote the mean and standard deviation of the mutual fund's excess return.[2] Denote the value function by $V(W, x)$; then the Bellman equation between adjustment dates is

$$\rho V(W,x) = \max_{\alpha,\alpha}\left[\kappa\frac{a^{1-\gamma}}{1-\gamma} + \frac{x^{1-\gamma}}{1-\gamma} + V_1(W,x)E\,dW + \frac{1}{2}V_{11}(W,x)\,\mathrm{Var}(dW)\right].$$

Following Grossman and Laroque, let $y = W/X - \lambda_1$ and define $h(y) = x^{-1+\gamma}V(W,x) = V(W/x, 1)$. Dividing through by $x^{1-\gamma}$ in the Bellman equation, we obtain

$$\rho h(y) = \max_{a,\alpha}\left[\kappa\frac{(a/x)^{1-\gamma}}{1-\gamma} + \frac{1}{1-\gamma} + h'(y)E\,dy + \frac{1}{2}h''(y)\,\mathrm{Var}(dy)\right],$$

and the budget constraint yields

$$dy = \big((y+\lambda_1)(r+\alpha\pi_r) - 1 - a/x\big)\,dt + (y+\lambda_1)\alpha\sigma_r\,dz.$$

Maximizing in α, the optimal portfolio satisfies

$$\alpha(y+\lambda_1) = \frac{-h'(y)}{h''(y)}\frac{\pi_r}{\sigma_r^2}$$

and adjustable consumption is

$$\frac{a}{x} = \left[\frac{h'(y)}{\kappa}\right]^{-1/\gamma}.$$

Substituting back into the Bellman equation, we obtain

$$\rho h(y) = h'(y)^{1-1/\gamma}\kappa^{1/\gamma}\frac{\gamma}{1-\gamma} + \frac{1}{1-\gamma} + h'(y)\big[(y+\lambda_1)r - 1\big] - \frac{1}{2}\frac{h'(y)^2}{h''(y)}\frac{\pi_r^2}{\sigma_r^2}.$$

[2]In our setting we can use $\pi_r = [(\pi_M/\sigma_M)^2 + (\pi_M/\sigma_M)^2]/[\pi_M/\sigma_M^2 + \pi_M/\sigma_M^2]$ and $\sigma_r^2 = \pi_r/[\pi_M/\sigma_M^2 + \pi_M/\sigma_M^2]$.

This is an ordinary differential equation for $h(y)$. To obtain boundary conditions, note that, on an adjustment date, the value function equals

$$\begin{aligned}&\frac{V(W,x)}{x^{1-\gamma}}\\&=\frac{1}{x^{1-\gamma}}\max_{x'}V\big(W-\lambda_1 x-\lambda_2 x',x'\big)\\&=\left(\frac{W-\lambda_1 x}{x}\right)^{1-\gamma}\cdot\max_{x'}\left(\frac{x'}{W-\lambda_1 x}\right)^{1-\gamma}\cdot V\left(\frac{W-\lambda_1 x}{x'}-\lambda_2,1\right)\\&=\left(\frac{W-\lambda_1 x}{x}\right)^{1-\gamma}\cdot\max_{y}(y+\lambda_1+\lambda_2)^{-1+\gamma}h(y).\end{aligned}$$

Define

$$M=\max_{y}(y+\lambda_1+\lambda_2)^{-1+\gamma}h(y);$$

then by the above reasoning, at the edges of the inaction band, denoted y_1 and y_2, we have

$$h(y_i)=My_i^{1-\gamma}.$$

Moreover, smooth pasting implies

$$h'(y_i)=M(1-\gamma)y_i^{-\gamma}.$$

Finally, the target value of y satisfies

$$y^*=\arg\max(y+\lambda_1+\lambda_2)^{-1+\gamma}h(y).$$

To numerically solve the ODE subject to these conditions, we follow the approach outlined by Grossman and Laroque. We first pick some M, pick y_1, solve the ODE with initial conditions as given above. If there is no y_2 for which the boundary conditions are satisfied, then we start with a different y_1. If the boundary conditions do hold for some y_2, then we check if M satisfies the equation above; if not, we start with a different M.

Simulating Dynamics

We simulate the dynamics of an economy populated by a continuum of commitment agents using the partial differential equation of Proposition 3. We discretize the differential equation following the approach presented in Caballero (1993). We use this methodology to compute the steady-state density f^*, to compute the impulse response (Definition 1), and to simulate dynamics along a sequence of aggregate shocks. We compute the matching consumption habit weights using Lemma 5 of Appendix A, and simulate the dynamics of the habit model using equation (10) of the main text.

ADDITIONAL REFERENCES

BEN-ARI, I., AND R. PINSKY (2007): "Spectral Analysis of a Family of Second-Order Elliptic Operators With Nonlocal Boundary Condition Indexed by a Probability Measure," *Journal of Functional Analysis*, 251, 122–140. [6]

——— (2009): "Ergodic Behavior of Diffusions With Random Jumps From the Boundary," *Stochastic Processes and Their Applications*, 119, 864–881. [6]

DARLING, D. A., AND A. J. F. SIEGERT (1953): "The First Passage Problem for a Continuous Markov Process," *The Annals of Mathematical Statistics*, 24, 624–639. [5]

GRIGORESCU, I., AND M. KANG (2002): "Brownian Motion on the Figure Eight," *Journal of Theoretical Probability*, 15 (3), 817–844. [5]

KARATZAS, I., AND S. SHREVE (1988): *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, Vol. 113. New York: Springer-Verlag. [10]

LEW, J. (1972): "On Linear Volterra Integral Equations of Convolution Type," *Proceeedings of the American Mathematical Society*, 35 (2), 450–456. [8]

REVUZ, D., AND M. YOR (1994): *Continuous Martingales and Brownian Motion* (Second Ed.). Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 293. Berlin: Springer-Verlag. [5]

Dept. of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, U.S.A. and NBER; chetty@stanford.edu

and

Dept. of Economics, Central European University, Nádor utca 9, Budapest 1051, Hungary and CEPR; szeidla@ceu.edu.

Co-editor Harald Uhlig handled this manuscript.

Manuscript received June, 2010; final revision received September, 2015.