Consumption and Aggregate Constraints: Evidence from U.S. States and Canadian Provinces

Charlotte Ostergaard
Norwegian School of Management

Bent E. Sørensen
Binghamton University and Centre for Economic Policy Research

Oved Yosha
Tel Aviv University and Centre for Economic Policy Research

State-level consumption exhibits excess sensitivity to lagged income to the same extent as U.S. aggregate data, but state-specific (idiosyncratic) consumption exhibits substantially less sensitivity to lagged state-specific income—a result that also holds for Canadian provinces. We propose the following interpretation: borrowing and lending in response to changes in consumer demand are easier for individual U.S. states than for the United States as a whole, and therefore, the measured deviation from the benchmark permanent income hypothesis model is smaller. However, lagged state-specific variables help predict state-specific consumption, suggesting that the PIH model still requires qualification.

We thank two referees, Lars Hansen, a second editor, Sumru Altug, and Kenneth West for helpful comments; Karen Lewis, Nouriel Roubini, and David Weil for useful discussions; the publishers of Sales and Marketing Management for permission to use their retail sales data; and Marco del Negro for help with the data.
I. Introduction

Personal consumption in the United States amounts to 70 percent of gross domestic product, and the modeling of consumer behavior is still a challenge to the profession in spite of much serious research. Hall's (1978) version of the permanent income hypothesis (PIH) implies that current consumption is independent of lagged disposable income conditional on lagged consumption. Micro evidence for this proposition is mixed, whereas macro evidence overwhelmingly rejects it, resulting in an empirical stylized fact: the excess sensitivity of consumption to lagged income. See Flavin (1981) and Hall and Mishkin (1982), and Hansen, Roberds, and Sargent (1991) for a thorough theoretical analysis.

Hall's model relies on a constant rate of interest; however, when consumers in a fairly closed economy such as the United States wish to increase the share of national income devoted to consumption, there is increased competition for scarce resources since aggregate consumption cannot adjust immediately. For example, it may take time for the United States as a whole to borrow internationally or to increase the quantity of goods imported, which creates upward pressure on the U.S.-wide interest rate, depressing the demand for consumption (see Michtener 1984; Christiano 1987). Hansen and Singleton (1982, 1983) developed and tested the empirical implications of the PIH when asset returns and, in particular, interest rates are time-varying and stochastic, but their model failed to fit U.S. macroeconomic time series, and the literature following Hansen and Singleton has not been successful in improving the fit. If measured interest rates and asset returns do not fully capture the closed-economy constraints on aggregate consumption, we should indeed expect substantial deviations from the benchmark PIH model in aggregate data.

We suggest a way around this problem. Individual U.S. states can more easily borrow and lend in response to changes in consumer demand, so if the PIH model in part fails because of closed-economy effects, we may still expect the model to perform well with idiosyncratic (state-specific) consumption and income, as changes in state-specific consumption demand aggregate to zero by definition in any given year.

To corroborate this conjecture empirically, we examine implications of the PIH using data on personal disposable income and consumption for U.S. states and Canadian provinces. Regional data at the subnational level are much underutilized for the study of consumer behavior. Such data are sufficiently aggregated to be regarded as macroeconomic data.

1 Tests of the PIH model with a time-varying but nonstochastic interest rate and macrolevel time series also fail (see Mankiw 1981; Shapiro 1984). Micro studies allowing for a time-varying (nonstochastic) interest rate have been more successful (see Altonji and Siow 1987; Mariger and Shaw 1995).
yet exhibit considerable cross-sectional variation that can be exploited in empirical analysis. Endogeneity of state-specific income is not likely to be a major problem, and measurement error is less serious than in micro data.

We find that state-level consumption and disposable income exhibit considerable excess sensitivity, similar to that found in aggregate U.S. data. We remove the aggregate U.S.-wide component in the data and find that state-specific consumption exhibits substantially less sensitivity to lagged state-specific disposable income. Similar results are obtained for Canadian provinces. Thus, once aggregate income and consumption fluctuations are controlled for, the deviation from PIH consumption behavior in macroeconomic data is smaller.

However, even after the aggregate component in the data is removed, lagged variables help predict consumption growth, suggesting that the PIH model still requires qualification.

II. Data

U.S. State-Level Data

We use annual data for 1963–95. Disposable personal income data are taken from the Bureau of Economic Analysis. We approximate state-level private nondurable consumption by state-level retail sales of nondurable goods published in the Survey of Buying Power, published in *Sales Management* (after 1976, *Sales and Marketing Management*). Retail sales are a somewhat noisy proxy for state-level private consumption (e.g., travel expenses are not included in retail sales), but to our knowledge, it is the best available. The correlation between annual percentage increments of aggregate U.S. nondurable retail sales and aggregate U.S. nondurable private consumption (in data from the national income and product accounts), both measured in real (consumer price index deflated) terms, is .68. We transform the data series to per capita terms using population data from the Bureau of Economic Analysis. We denote as state $i$'s period $t$ idiosyncratic (state-specific) disposable log income per capita—for brevity, “income”—where $Y_i$ is period $t$ (total) income and $Y$ is period $t$ aggregate (U.S.-wide) income.

It is widely accepted that U.S. aggregate income series are nonstationary. By contrast, the statistical properties of the idiosyncratic components of U.S. state-level data have not been studied. Exploiting the panel structure of our data, we perform the Im, Pesaran, and Shin (1997) (IPS) test for a unit root in $y_i$. The null hypothesis of nonstationarity is not rejected with one, two, and three lags ($p$-values of .13, .45, and .34, respectively). The IPS test is valid for independent observations, and since the idiosyncratic components of income are unlikely
to be fully independent, the critical values of the test statistics must be taken as approximations. State-by-state augmented Dickey-Fuller (ADF) tests reject the unit root null hypothesis for only a few states at conventional significance levels. The ADF tests provide weak evidence since they have low power for samples as short as ours, but along with the IPS test they provide a reasonable guide to specification. Overall, the idiosyncratic component of U.S. state-level disposable income is well described as an integrated process.

We define idiosyncratic state-level per capita log consumption—for brevity, “consumption”—in the same manner: \( c_a = C_a - C_r \). The aggregate per capita nondurable retail sales series, \( C_a \), is clearly nonstationary according to standard ADF tests. Using the IPS test, we find that the null hypothesis of nonstationarity is not rejected for a specification with one lag (p-value of .10) but is rejected with two and three lags (p-values of .01 and .02, respectively). State-by-state ADF tests rarely reject the hypothesis of nonstationarity, so we conclude that the idiosyncratic component of consumption is best regarded as nonstationary.

Most models of consumption imply that consumption tracks income in the long run. An interpretation is that the process \( c_a - y_a \) is stationary, that is, that consumption and income are cointegrated with a coefficient of unity. We test this hypothesis using the IPS test, which consistently rejects the null of a unit root for various lag lengths and whether a drift term is allowed for or not. We therefore feel confident treating \( c_a - y_a \) as a stationary process. Since both \( y_a \) and \( c_a \) are nonstationary, we carry out the empirical analysis using first-differenced series.

We estimated AR(2) models for the income series and found the coefficients of the twice-lagged variables very small and typically insignificant. A test of the hypothesis that the AR(2) coefficients are all zero provided no evidence against the null, so a simple AR(1) model in log differences seems appropriate. All regressions reported in this article are estimated using feasible generalized least squares (GLS) allowing for cross correlations of the disturbances between states.\(^2\) For \( \Delta Y_a \), the estimated average of the AR(1) coefficients is 0.16, with the absolute value of the t-statistics averaging 2.46. For \( \Delta y_a \), the average of the AR(1) coefficients is 0.05, and the average for the absolute value of the t-statistics averaging 2.

\(^2\) We estimate an unrestricted variance-covariance matrix for the 50 states on the basis of the residuals from an initial panel data ordinary least squares (OLS) estimation. Since we have fewer than 50 time-series observations for each state, this estimated variance-covariance matrix is singular; in order to perform the second-stage GLS estimation, we modified the estimated variance-covariance matrix by decreasing the off-diagonal elements by 10 percent (this reduction is not done for the Canadian data). The estimated coefficients are very similar to those obtained from OLS regressions. The estimated standard errors depend on the procedure used, but we verified empirically that the qualitative conclusions of the present paper hold even if covariances across states are set equal to zero, as long as variances are allowed to differ across states.
Fig. 1.—Growth rates of idiosyncratic and aggregate income and consumption. a, Georgia (state 10); b, New Jersey (state 30); c, Wyoming (state 50); d, U.S. aggregate. For panels a–c, income is the log difference of per capita disposable income of the state minus the log difference of per capita U.S.-wide disposable income (similarly for the consumption series). Consumption is proxied by nondurable retail sales. For panel d, income is the log difference of aggregate (U.S.-wide) per capita disposable income. Consumption is the log difference of aggregate (U.S.-wide) per capita nondurable retail sales. Sample period: 1964–95.

statistics is 2.65. In both regressions, the null hypothesis that the AR(1) coefficients are equal across U.S. states is strongly rejected with p-values of .00.\(^3\)

To get a sense of the variation in the idiosyncratic series, we display them in figure 1 for three states.\(^4\) We also display the aggregate series for the United States. The idiosyncratic nondurable retail sales series show more variation than the idiosyncratic disposable income series,

\(^3\)We ran similar regressions including the lagged consumption/income ratio as an additional regressor. The estimated coefficient on this variable is positive and significant, and the estimated AR(1) and AR(2) coefficients remain roughly the same. This type of regression was suggested by Cochrane (1994) in the framework of a bivariate system to predict U.S. aggregate income (gross national product) and to identify the transitory component in income. Our results are broadly similar to those obtained by Cochrane for the prediction equation for income.

\(^4\)We chose three states at random (states 10, 30, and 50 in alphabetical order).
with average standard deviations (across the 50 states) $\sigma_{c_{it}} = 4.19$ and $\sigma_{\Delta c_{it}} = 2.31$. For the series $\Delta C_{it}$ and $\Delta Y_{it}$, $\sigma_{\Delta C_{it}} = 3.12$ and $\sigma_{\Delta Y_{it}} = 5.01$. These standard errors do not indicate that measurement error in the idiosyncratic state-level series is larger than in the state-level series.

*Canadian Province-Level Data*

Data are available from the CANSIM database maintained by Statistics Canada. We use annual data for 1961–96 for personal disposable income, nondurable consumption (defined as the sum of nondurables, semidurables, and services), population, and aggregate consumer prices. The statistical properties of the province-level income and consumption series are similar to those of their U.S. state-level counterparts. The IPS test for a unit root in $Y_{it}$ rejects the null hypothesis of nonstationarity for an autoregressive model with one lag, but not for models with two or three lags ($p$-values of .00, .07, and .07, respectively). Province-by-province ADF tests provide no evidence against unit roots in these series. For $c_{it}$, the IPS test easily accepts nonstationarity. The AR(2) models of $\Delta Y_{it}$ and $\Delta c_{it}$ are similar to their U.S. states counterparts: the AR(2) coefficients are typically very small, typically insignificant, and jointly not different from zero. The average AR(1) coefficient for $\Delta Y_{it}$ is 0.14, whereas for $\Delta c_{it}$ it is $-0.08$. Since province-level consumption is part of province-level “national accounts,” measurement error is likely to be less severe than in the U.S. state-level retail sales data.

III. Empirical Results

We turn to our central empirical question: Is excess sensitivity lower when aggregate fluctuations are controlled for? Table 1 displays regressions of consumption growth on lagged income growth with and without controlling for aggregate fluctuations. We display the results from two specifications: one with two lags of income and one that further includes an error correction term as in Cochrane (1994) (see n. 3). In all the regressions we include state fixed effects. (The results are not affected substantially when they are omitted.)

The main empirical finding is that when aggregate fluctuations are controlled for, the coefficient of one-year lagged income is smaller. In the specification with an error correction term, the coefficient even drops to zero, and the coefficient of two-year lagged income is also not significantly different from zero.

---

5 Including the lagged consumption/income ratio in these regressions results in somewhat smaller coefficients for lagged income and a positive significant coefficient to the lagged consumption/income ratio. In conjunction with the results found for U.S. states, this indicates that Cochrane’s (1994) results are very robust.
### Table 1

**Sensitivity of U.S. State-Level Consumption to Lagged Income**

<table>
<thead>
<tr>
<th>Estimate</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta C_i = \alpha_i + b_1 \Delta Y_{i,t-1} + b_2 \Delta Y_{i,t-2} + \varepsilon_i$</td>
<td></td>
</tr>
<tr>
<td>$b_1$</td>
<td>0.21</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Model:

$\Delta C_i = \alpha_i + b_1 \Delta Y_{i,t-1} + b_2 \Delta Y_{i,t-2} + b_3 (C_{i,t-1} - Y_{i,t-1}) + \varepsilon_i$

<table>
<thead>
<tr>
<th>Estimate</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>0.19</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.07</td>
</tr>
<tr>
<td>$b_3$</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

Model:

$\Delta C_i = \alpha_i + b_1 \Delta Y_{i,t-1} + b_2 \Delta Y_{i,t-2} + \varepsilon_i$

<table>
<thead>
<tr>
<th>Estimate</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>0.12</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Note.—The term $\Delta Y_t$ is the period $t$ log difference of state $i$’s (total) per capita disposable income. The equation $\Delta Y_i = \Delta Y_t - \Delta Y_t$ is the period $t$ log difference of state $i$’s idiosyncratic (state-specific) per capita disposable income, where $\Delta Y_t$ is the period $t$ log difference of aggregate (U.S-wide) per capita disposable income (similarly for the consumption series). State-level consumption is proxied by nondurable retail sales. Two-stage GLS estimation is used, where the first stage is OLS and the second stage is GLS with covariance matrix $I_T \otimes \Omega$, where $I_T$ is an identity matrix with dimension equal to the time dimension of the sample and $\Omega_i$ has typical element $(1/\hat{\text{ee}}_i)$, are reduced by 10 percent to avoid singularity of the covariance matrix. The sample period is 1965–95.

The coefficient of the error correction term is negative and significant, suggesting that this specification captures important dynamics in the data. The negative coefficient is considerably larger and more strongly significant when aggregate fluctuations are controlled for. The high $t$-statistic associated with this coefficient (especially in the regression that controls for aggregate fluctuations) implies that the lagged income/consumption ratio helps forecast consumption growth, contrary to the predictions of the PIH.

Table 2 displays analogous regressions for Canadian provinces. Again, when aggregate fluctuations are controlled for, the coefficient of one-year lagged income is smaller. A similar pattern also holds for the coefficient of two-year lagged income. The coefficient of the error correction term is positive, in contrast to that for U.S. states, but is smaller in magnitude (and not significant at the 5 percent level) when aggregate fluctuations are controlled for.

When the U.S. and Canadian results are considered together, the lower sensitivity of consumption to lagged variables when aggregate variables are controlled for appears to be a robust stylized fact. The impact of the lagged consumption/income ratio on current consumption is not robustly estimated since its sign varies between the U.S. and
Canadian data. We suspect that this may be due to measurement error in the U.S. retail sales data, making these data unsuited as right-hand-side variables. The lower sensitivity of consumption to the idiosyncratic component of lagged income is not likely to be driven by higher measurement error because, as was reported in Section II for the U.S. data, the idiosyncratic income series exhibit low variance compared to idiosyncratic consumption and overall income.

Further Robustness Checks

The finding that the coefficient of one-year lagged income is smaller when aggregate fluctuations are controlled for is extremely robust. We tried out specifications (not reported in the tables) in which we control for aggregate fluctuations by including time dummy variables (time fixed effects), rather than subtracting the aggregate variables, in the regressions of $\Delta C_i$ on $\Delta Y_{i,t-1}$ or by including aggregate consumption as a regressor. We ran the various specifications with only one lag, with two lags, with and without an error correction term, and with and without state fixed effects. In all cases, we obtained a much lower excess sensitivity coefficient compared to the coefficient in the regression of $\Delta C_i$ on $\Delta Y_{i,t-1}$ without controlling for aggregate fluctuations. Including
lagged consumption in the prediction equations for consumption resulted in an insignificant coefficient to this variable for the United States, with little impact on the other regressors. A positive significant coefficient to lagged consumption was found for Canada whether aggregate variables were controlled for or not. However, in every specification we estimated using Canadian data, the coefficients to all lagged variables showed a clear decline when aggregate variables were controlled for.

IV. Discussion

Interpretation

Our preferred interpretation relies on the “closedness” of the U.S. economy. There are (at least) two ways of thinking about closedness. The first stresses frictions and imperfections in international capital and credit markets rendering international borrowing and lending difficult and preventing rapid adjustment of aggregate consumption to U.S.-wide changes in consumption demand. By contrast, individual states are relatively open in the sense that they can more easily borrow and lend among themselves. Thus the adjustment of state-specific consumption to changes in state-specific consumption demand should be faster.

An alternative manner of thinking about closedness is centered on the slow adjustment of U.S. net imports in response to fluctuations in U.S. consumption demand. In a fully integrated and frictionless world, aggregate net imports would immediately increase in response to higher consumption demand. In reality, it may take time to adjust aggregate imports (not to speak of exports). For example, an increased demand for Toyota cars in the United States will typically be reflected in higher prices and less attractive financing opportunities since adjustment of Japanese exports cannot be done instantaneously. By contrast, net imports of a state within the United States can adjust much more rapidly. If, in some year, Massachusetts residents have a large idiosyncratic demand for consumption, this demand may be satisfied relatively quickly by moving goods from other states in which idiosyncratic demand is low.

These economic mechanisms may be independent or complementary (e.g., imports adjust slowly because international credit markets are imperfect), and we do not have a model or adequate data to disentangle them. Our empirical results strongly suggest that such mechanisms are part of the explanation for the seeming deviation from optimal consumer behavior in macroeconomic data.

Our empirical strategy allows us to circumvent the problem of how to measure the prevailing equilibrium interest rate. In practice, measured interest rates are affected by many factors such as monetary or
fiscal policy that are typically not incorporated in theoretical models and in empirical analyses of consumption. Moreover, consumers are often unable to obtain credit at posted interest rates. Our work avoids directly addressing these (important) issues and focuses on one key point: controlling for aggregate constraints improves the empirical performance of the PIH. The results in tables 1 and 2 indeed indicate that doing so substantially reduces excess sensitivity but does not eliminate it. The remaining excess sensitivity may be due to other frictions that have been extensively researched.

Relation to Tests of Full Risk Sharing

Suppose that consumers can insure their consumption ex ante, before shocks occur. (This assumption is clearly stronger than the assumption underlying the PIH, namely that consumers need only to have full access to a credit market in which they can borrow and lend ex post, after shocks occur.) Then, under commonly used assumptions—symmetric information, no transaction costs, constant relative risk aversion utility, and identical rate of time preference for all agents—full (Pareto-efficient) risk sharing within a group implies that \( \Delta C_a = \Delta C_o \), that is, the growth rate of each agent’s consumption and the growth rate of aggregate consumption are the same. The central empirical implication is that an agent’s consumption growth should not depend on any idiosyncratic characteristic of the agent—in particular, income growth. Thus, when aggregate fluctuations are controlled for, a regression of consumption growth on contemporaneous or lagged income should yield a coefficient of zero. This is the essence of the tests suggested by Cochrane (1991), Mace (1991), and Townsend (1994) (see also Obstfeld 1994). Asdrubali, Sørensen, and Yoshia (1996) and Sørensen and Yoshia (1998) nested this test within a cross-sectional variance decomposition, applied to U.S. states and OECD countries, respectively. They measured how much risk sharing is achieved via different mechanisms (e.g., portfolio diversification or federal taxes and transfers) by estimating a system of equations one of which is similar to those estimated by Cochrane, Mace, and Townsend. The regressions displayed in tables 1 and 2 constitute evidence that full risk sharing does not hold across U.S. states or Canadian provinces; if it did, a regression of \( \Delta C_a \) on \( \Delta y_{i,t-1} \), or on \( C_{i,t-1} - y_{i,t-1} \), should yield a coefficient of zero.  

6 There are many explanations for a lack of full risk sharing. Kocherlakota (1996), e.g., stresses limited enforceability and commitment, whereas others (e.g., Constantinides and Duffie 1996; Heaton and Lucas 1996) study the conditions that ensure that the full risk-sharing allocation is approximated (or even achieved) among heterogeneous agents in the absence of insurance opportunities when only intertemporal smoothing is present.
In our analysis, we ignored potential heterogeneity in state-level patterns of income and consumption. But the results reported in Section II indicate that the AR(1) processes for U.S. state-level income are not identical. Exploratory analysis points to a potentially interesting regularity: excess sensitivity is larger the higher the persistence of income. We shall not pursue this issue further in this paper, but we believe that utilizing state-level differences in order to sort through the many models of consumer behavior is a fruitful area for research.7

References


———. “Intranational (regional-level) heterogeneity is a promising source of information for other important questions in macroeconomics and international finance; see the collection of articles in Hess and van Wincoop (2000).
CONSUMPTION


