

# Full Web Appendix to "Consumption Inequality and Partial Insurance", by Richard Blundell, Luigi Pistaferri, and Ian Preston.

## Appendix A: Data

*The PSID.*—Since the PSID has been widely used for microeconomic research, we shall only sketch the description in this appendix.<sup>1</sup> The PSID started in 1968 collecting information on a sample of roughly 5,000 households. Of these, about 3,000 were representative of the US population as a whole (the core sample), and about 2,000 were low-income families (the Census Bureau’s Survey of Economic Opportunities, or SEO sample). Thereafter, both the original families and their split-offs (children of the original family forming a family of their own) have been followed.

The PSID includes a variety of socio-economic characteristics of the household, including education, food spending, and income of household members. Questions referring to income are retrospective; thus, those asked in 1993, say, refer to the 1992 calendar year. In contrast, the timing of the survey questions on food expenditure is much less clear (see Hall and Mishkin, 1982, and Altonji and Siow, 1987, for two alternative views). Typically, the PSID asks how much is spent on food in an average week. Since interviews are usually conducted around March, it has been argued that people report their food expenditure for an average week around that period, rather than for the previous calendar year as is the case for family income. We assume that food expenditure reported in survey year  $t$  refers to the previous calendar year, but check the effect of alternative assumptions.

Households in the PSID report their taxable family income (which includes transfers and financial income). The measure of income used in the baseline analysis below excludes income from financial assets, subtracts federal taxes on non-financial income and deflates the corresponding value by the CPI. We assume that federal taxes on non-financial income are a proportion of total federal taxes given by the ratio between non-financial income and total income. Before 1991, federal taxes are computed by PSID researchers and added into the data set using information on filing status, adjusted gross income, whether the respondent itemizes or takes the standard deduction, and other household characteristics that make them qualify for extra deductions, exemptions, and tax credits. Federal taxes are not computed in 1992 and 1993. For these two years, we impute taxes using the TAXSIM program at the NBER. Education level is computed using the PSID variable “grades of school finished”. Individuals who changed their education level during the sample period are allocated to the highest grade achieved. We consider two education groups: with and without college education (corresponding to more than high school and high school or less, respectively).

Since CEX data are available on a consistent basis since 1980, we construct an unbalanced PSID panel using data from 1978 to 1992 (the first two years are retained for initial conditions purposes). Due to attrition, changes in family composition, and various other reasons, household heads in the 1978-1992 PSID may be present from a minimum of one year to a maximum of fifteen years. We thus create unbalanced panel data sets of various length. The longest panel includes individuals present from 1978 to 1992; the shortest, individuals present for two consecutive years only (1978-79, 1979-80, up to 1991-92).

The objective of our sample selection is to focus on a sample of continuously married couples headed by a male (with or without children). The step-by-step selection of our PSID sample is illustrated in Table A1. We eliminate households facing some dramatic family composition change over the sample period. In particular, we keep only those with no change, and those experiencing changes in members other than the head or the wife (children leaving parental home, say). We next eliminate households headed by a female and those with missing report on education and region<sup>2</sup>. We keep continuously married couples and drop some income outliers.<sup>3</sup> We then drop those born before 1920 or after 1959.

For most of the analysis we exclude SEO households and their split-offs. However, we do consider the robustness of our results in the low income SEO subsample. Finally, we drop those aged less than 30 or more than 65. This is to avoid problems related to changes in family composition and education, in the

---

<sup>1</sup>See Martha Hill (1992) for more details about the PSID.

<sup>2</sup>When possible, we impute values for education and region of residence using adjacent records on these variables.

<sup>3</sup>An income outlier is defined as a household with an income growth above 500 percent, below  $-80$  percent, or with a level of income below \$100 a year.

first case, and retirement, in the second. We also check sensitivity of results to including people aged 25-30. The final sample used in the minimum distance exercise below is composed of 17,604 observations and 1,765 households. Our income regressions do not use 36 observations with topcoded income, financial income, or federal taxes. We use information on age and the survey year to allocate individuals in our sample to four cohorts defined on the basis of the year of birth of the household head: born in the 1920s, 1930s, 1940s, and 1950s. Years where cell size is less than 50 are discarded.

Table A1—Sample selection in the PSID

Reason for exclusion	# dropped	# remain
(Initial sample from family files, 1967-1992)	NA	172,274
Latino subsample	8,403	163,871
Intermittent headship	17,855	146,016
Interviewed prior to 1978	52,448	93,568
Change in family composition	18,561	75,007
Female head	23,806	51,201
Missing values	1,071	50,130
Change in marital status	5,737	44,393
Income outliers	2,386	42,007
Born before 1920 or after 1959	8,362	33,645
Poverty subsample	12,455	21,190
Aged less than 30 or more than 65	3,586	17,604

*The CEX.*—The Consumer Expenditure Survey provides a continuous and comprehensive flow of data on the buying habits of American consumers. The data are collected by the Bureau of Labor Statistics and used primarily for revising the CPI.<sup>4</sup> The definition of the head of the household in the CEX is the person or one of the persons who owns or rents the unit; this definition is slightly different from the one adopted in the PSID, where the head is always the husband in a couple. We make the two definitions compatible.

The CEX is based on two components, the Diary survey and the Interview survey. The Diary sample interviews households for two consecutive weeks, and it is designed to obtain detailed expenditures data on small and frequently purchased items, such as food, personal care, and household supplies. The Interview sample follows survey households for a maximum of 5 quarters, although only inventory and basic sample data are collected in the first quarter. The data base covers about 95 percent of all expenditure, with the exclusion of expenditures for housekeeping supplies, personal care products, and non-prescription drugs. Following most previous research, our analysis below uses only the Interview sample.<sup>5</sup>

As the PSID, the CEX collects information on a variety of socio-demographic variables, including income and consumer expenditure. Expenditure is reported in each quarter and refers to the previous quarter; income is reported in the second and fifth interview (with some exceptions), and refers to the previous twelve months. For consistency with the timing of consumption, fifth-quarter income data are used.

Our initial 1980-2004 CEX sample includes 1,848,348 monthly observations, corresponding to 192,564 households. We drop those with missing record on food and/or zero total nondurable expenditure, and those who completed less than 12 month interviews. This is to obtain a sample where a measure of annual consumption can be obtained. A problem is that many households report their consumption for overlapping years, i.e. there are people interviewed partly in year  $t$  and partly in year  $t + 1$ . Pragmatically, we assume that if the household is interviewed for at least 6 months at  $t + 1$ , then the reference year is  $t + 1$ , and it is  $t$  otherwise. Prices are adjusted accordingly. We then sum food at home, food away from home and other nondurable expenditure over the 12 interview months. This gives annual expenditures. For consistency with the timing of the PSID data, we drop households interviewed after 1992. We also drop those with zero before-tax income, those with missing region or education records, single households and those with changes

<sup>4</sup>A description of the survey, including more details on sample design, interview procedures, etc., may be found in “Chapter 16: Consumer Expenditures and Income”, from the BLS Handbook of Methods.

<sup>5</sup>There is some evidence that trends in consumption inequality measured in the two CEX surveys have diverged in the 1990s (Attanasio, Battistin and Ichimura, 2004). While research on the reasons for this divergence is clearly warranted, our analysis, which uses data up to 1992, will only be marginally affected.

in family composition. Finally, we eliminate households where the head is born before 1920 or after 1959, those aged less than 30 or more than 65, and those with outlier income (defined as a level of income below the amount spent on food) or incomplete income responses. The final sample used to estimate the food demand equation in Table 1 contains 14,430 households. Table A2 details the sample selection process in the CEX.

Table A2—Sample selection in the CEX

Reason for exclusion	# dropped	# remain
(Initial sample, 1980-2004)	NA	192,564
Missing expenditure data	1,488	191,076
Present for less than 12 months	98,926	92,150
Observed after 1992	47,901	44,249
Zero before-tax income	1,321	42,928
Missing values	4,015	38,913
Marital status	16,141	22,772
Born before 1920 or after 1959	4,696	18,076
Aged less than 30 or more than 65	1,860	16,216
Income outliers and incomplete income response	1,786	14,430

## Appendix B: The Euler Equation Approximation

If preferences are quadratic (and interest rates are not subject to uncertainty), it is possible to obtain a closed form solution for consumption. It is also straightforward to derive an exact mapping between the expectation error of the Euler equation for consumption and income shocks. See Hall and Mishkin (1982), for example. Quadratic preferences have well known undesirable features, such as increasing risk aversion and lack of a precautionary motive for saving. More realistic preferences, such as the CRRA functional form used here, solve these problems but deliver no closed form solution for consumption. The Euler equation can be linearized to describe the behavior of consumption growth. In this appendix we derive an approximation of the mapping between the expectation error of the Euler equation and the income shock.

Consider the consumption problem faced by household  $i$  of age  $t$ . Assuming that preferences are of the CRRA form, the objective is to choose a path for consumption  $C$  so as to:

$$(B1) \quad \max_C E_t \sum_{j=0}^{T-t} \frac{1}{(1+\delta)^j} \frac{C_{i,t+j}^{\beta-1}}{\beta} e^{Z'_{i,t+j} \vartheta_{t+j}}.$$

where  $Z_{i,t+j}$  incorporates taste shifters. Maximization of (B1) is subject to the budget constraint which in the self-insurance model assumes individuals have access to a risk free bond with real return  $r_{t+j}$

$$(B2) \quad \begin{aligned} A_{i,t+j+1} &= (1 + r_{t+j}) (A_{i,t+j} + Y_{i,t+j} - C_{i,t+j}) \\ A_{i,T} &= 0 \end{aligned}$$

with  $A_{i,t}$  given. We set the retirement age after which labor income falls to zero at  $L$ , assumed known and certain, and the end of the life-cycle at age  $T$ . We assume that there is no uncertainty about the date of death. With budget constraint (B2) optimal consumption choices can be described by the Euler equation

$$(B3) \quad C_{i,t-1}^{\beta-1} = \frac{1+r_{t-1}}{1+\delta} e^{\Delta Z'_{i,t} \vartheta_t} E_{t-1} C_{i,t}^{\beta-1}.$$

As it is, equation (B3) is not useful for empirical purposes. We then consider approximating it in the following way. In general, the logarithm of the sum of an arbitrary series  $X_t, X_{t+1}, \dots, X_S$  can be written as:

$$\ln \sum_{k=0}^{S-t} X_{t+k} = \ln X_t + \ln \left[ 1 + \sum_{k=1}^{S-t} \exp(\ln X_{t+k} - \ln X_t) \right]$$

Taking a Taylor expansion around  $\ln X_{t+k} = \ln X_t + \sum_{i=0}^k \delta_{t+i}$ ,  $k = 1, \dots, S-t$  for some path of increments  $\delta_t, \delta_{t+1}, \dots, \delta_S$  with  $\delta_t = 0$ ,

$$(B4) \quad \begin{aligned} \ln \sum_{k=0}^{S-t} X_{t+k} &\simeq \ln X_t + \ln \left[ 1 + \sum_{k=1}^{S-t} \exp \left( \sum_{i=0}^k \delta_{t+i} \right) \right] \\ &+ \sum_{k=1}^{S-t} \frac{\exp \left( \sum_{i=0}^k \delta_{t+i} \right)}{\left[ 1 + \sum_{k=1}^{S-t} \exp \left( \sum_{i=0}^k \delta_{t+i} \right) \right]} \left( \ln X_{t+k} - \ln X_t - \sum_{i=0}^k \delta_{t+i} \right) \\ &\simeq \sum_{k=0}^{S-t} \alpha_{t+k,S}^{\delta} \left[ \ln X_{t+k} - \ln \alpha_{t+k,S}^{\delta} \right] \end{aligned}$$

where  $\alpha_{t+k,S}^\delta = \exp(\sum_{i=0}^k \delta_{t+i}) / \left[1 + \sum_{k=1}^{S-t} \exp(\sum_{i=0}^k \delta_{t+i})\right]$  and the error in the approximation is  $O(\sum_{k=0}^{S-t} (\ln X_{t+k} - \ln X_t - \sum_{i=0}^k \delta_{t+i})^2)$ .

Applying this approximation to the Euler equation (B3) above gives:

$$\Delta \log C_{i,t} \simeq \Delta Z'_{i,t} \vartheta'_t + \eta_{i,t} + \Omega_{i,t}$$

where  $\vartheta'_t = (1 - \beta)^{-1} \vartheta_t$ ,  $\eta_{i,t}$  is a consumption shock with  $E_{t-1} \eta_{i,t} = 0$ ,  $\Omega_{i,t}$  captures any slope in the consumption path due to interest rates, impatience or precautionary savings and the error in the approximation is  $O(E_t \eta_{i,t}^2)$ . Suppose that any idiosyncratic component to this gradient to the consumption path can be adequately picked up by a vector of deterministic characteristics  $\Gamma_{i,t}$  and a stochastic individual element  $\xi_{i,t}$

$$\Delta \log C_{i,t} - \Gamma_{i,t} - \Delta Z'_{i,t} \vartheta'_t = \Delta c_{i,t} \simeq \eta_{i,t} + \xi_{i,t}.$$

From (2) we also have

$$\Delta y_{i,t+k} = \zeta_{i,t+k} + \sum_{j=0}^q \theta_j \varepsilon_{i,t+k-j}.$$

The intertemporal budget constraint is

$$\sum_{k=0}^{T-t} Q_{t+k} C_{i,t+k} = \sum_{k=0}^{L-t} Q_{t+k} Y_{i,t+k} + A_{i,t}$$

where  $T$  is death,  $L$  is retirement and  $Q_{t+k}$  is appropriate discount factor  $\prod_{i=1}^k (1 + r_{t+i})$ ,  $k = 1, \dots, T - t$  (and  $Q_t = 1$ ).

Applying the approximation (B4) appropriately to each side

$$\begin{aligned} & \sum_{k=0}^{T-t} \alpha_{t+k,T}^{\Omega + \Delta Z \vartheta' - r} [\ln C_{i,t+k} - \ln Q_{t+k} - \ln \alpha_{t+k,T}^{\omega - r}] \\ & \simeq \pi_{i,t} \sum_{k=0}^{L-t} \alpha_{t+k,L}^{\Delta Z \vartheta' - r} [\ln Y_{i,t+k} - \ln Q_{t+k} - \ln \alpha_{t+k,L}^{\Delta Z \varphi - r}] \\ & + (1 - \pi_{i,t}) \ln A_{i,t} - [(1 - \pi_{i,t}) \ln(1 - \pi_{i,t}) + \pi_{i,t} \ln \pi_{i,t}] \end{aligned}$$

where  $\pi_{i,t} = \frac{\sum_{k=0}^{L-t} Q_{t+k} Y_{i,t+k}}{\sum_{k=0}^{L-t} Q_{t+k} Y_{i,t+k} + A_{i,t}}$  is the share of future labor income in current human and financial wealth.

Taking differences in expectations gives

$$\eta_{i,t} \simeq \pi_{i,t} [\zeta_{i,t} + \gamma_{t,L} \varepsilon_{i,t}]$$

where  $\gamma_{t,L} = \left(\sum_{j=0}^q \alpha_{t+j,L}^{\Delta Z \vartheta' - r} \theta_j\right)$  and the error on the approximation is  $O([\zeta_{i,t} + \gamma_{t,L} \varepsilon_{i,t}]^2 + E_{t-1} [\zeta_{i,t} + \gamma_{t,L} \varepsilon_{i,t}]^2)$ .

Then<sup>6</sup>

$$(B5) \quad \Delta c_{i,t} \simeq \xi_{i,t} + \pi_{i,t} \zeta_{i,t} + \gamma_{t,L} \pi_{i,t} \varepsilon_{i,t}.$$

with a similar order of approximation error.

If  $\Delta Z'_{i,t+k} \vartheta' = 0$  and  $r_t = r$  is constant then  $\alpha_{t+j,L}^{\Delta Z \vartheta' - r} = \alpha_{t+j,L}^{-r} = \exp(-jr) / \sum_{k=0}^L \exp(-kr) \simeq r / (1+r)^k$  and  $\gamma_{t,L} \simeq \frac{r}{1+r} [1 + \sum_{j=1}^q \theta_j / (1+r)^j]$ .

## Appendix C: Identification

*The simplest model.*—Here we show how the model can be identified with four years of data ( $t+1, t, t-1, t-2$ ), and discuss various extensions. We start with the simplest model with no measurement error, serially uncorrelated transitory component, and stationarity. This is relaxed later. (Unexplained) consumption and income growth in period  $s$  ( $s = t-1, t, t+1$ ) are, respectively:

$$\begin{aligned} \Delta c_s &= \xi_s + \phi \zeta_s + \psi \varepsilon_s \\ \Delta y_s &= \zeta_s + \Delta \varepsilon_s \end{aligned}$$

(where for simplicity we have assumed that the transitory shock to income is i.i.d.).<sup>7</sup> The parameters to identify are:  $\phi, \psi, \sigma_\xi^2, \sigma_\zeta^2$ , and  $\sigma_\varepsilon^2$ .

As in Meghir and Pistaferri (2004), we can prove that:

$$(C1) \quad E(\Delta y_t (\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})) = \sigma_\zeta^2$$

and that:

$$(C2) \quad E(\Delta y_t \Delta y_{t-1}) = E(\Delta y_{t+1} \Delta y_t) = -\sigma_\varepsilon^2$$

Identification of  $\sigma_\varepsilon^2$  through (C2) rests on the idea that income growth rates are autocorrelated due to mean reversion caused by the transitory component (the permanent component is subject to i.i.d. shocks).

<sup>6</sup>Blundell, Low and Preston (2004) contains a lengthier derivation of such an expression, including discussion of the order of magnitude of the approximation error involved.

<sup>7</sup>The proof mechanism can be easily extended to deal with MA(q) transitory shock processes as in Meghir and Pistaferri (2004).

Identification of  $\sigma_\zeta^2$  through (C1) rests on the idea that the variance of income growth ( $E(\Delta y_t \Delta y_t)$ ), subtracted the contribution of the mean reverting component ( $E(\Delta y_t \Delta y_{t-1}) + E(\Delta y_t \Delta y_{t+1})$ ), coincides with the variance of innovations to the permanent component.

In general, if one has  $T$  years of data, only  $T - 3$  variances of the permanent shock can be identified, and only  $T - 2$  variances of the i.i.d. transitory shock can be identified. As said in the text, with panel data on income, the variances of permanent and transitory shock can be identified without recourse to consumption data.

One can also prove that:

$$(C3) \quad \frac{E(\Delta c_t (\Delta y_{t-1} + \Delta y_t + \Delta y_t))}{E(\Delta y_t (\Delta y_{t-1} + \Delta y_t + \Delta y_t))} = \phi$$

$$(C4) \quad \frac{E(\Delta c_t \Delta y_{t+1})}{E(\Delta y_t \Delta y_{t+1})} = \psi$$

$$(C5) \quad E(\Delta c_t (\Delta c_{t-1} + \Delta c_t + \Delta c_{t+1})) - \frac{[E(\Delta c_t (\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1}))]^2}{E(\Delta y_t (\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1}))} + \frac{[E(\Delta c_t \Delta y_{t+1})]^2}{E(\Delta y_t \Delta y_{t+1})} = \sigma_\xi^2.$$

These moment conditions provide complete identification of the parameters of interest. Identification of  $\psi$  using (C4) uses the fact that income and lagged consumption may be correlated through the transitory component ( $E(\Delta c_t \Delta y_{t+1}) = \psi \sigma_\varepsilon^2$ ). Scaling this by  $E(\Delta y_t \Delta y_{t+1}) = \sigma_\varepsilon^2$  identifies the loading factor  $\psi$ . Note that there is a simple IV interpretation here:  $\psi$  is identified by a regression of  $\Delta c_t$  on  $\Delta y_t$  using  $\Delta y_{t+1}$  as an instrument. A similar reasoning applies to (C3): the current covariance between consumption and income growth ( $E(\Delta c_t \Delta y_t)$ ), stripped of the contribution of the transitory component, reflects the arrival of permanent income shocks ( $E(\Delta c_t (\Delta y_{t-1} + \Delta y_t + \Delta y_t)) = \phi \sigma_\zeta^2$ ). Scaling this by the variance of permanent income shock, identified by using income moments alone, identifies the loading factor  $\phi$ . Note that here too there is a simple IV interpretation:  $\phi$  is identified by a regression of  $\Delta c_t$  on  $\Delta y_t$  using  $(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})$  as an instrument. Finally, (C5) identifies the variance of the component  $\sigma_\xi^2$  using a residual variability idea: the variance of consumption growth, stripped of the contribution of permanent and transitory income shocks, reflects heterogeneity in the consumption gradient.

*Measurement error in consumption.*—Consider now the realistic case in which consumption is measured with error, i.e.,

$$c_{i,t}^* = c_{i,t} + u_{i,t}^c$$

where  $c^*$  denote measured consumption,  $c$  is true consumption, and  $u^c$  the measurement error. Measurement error in consumption induces serial correlation in consumption growth. Because consumption is a martingale with drift in the absence of measurement error, the variance of measurement error can be readily recovered using

$$(C6) \quad E(\Delta c_t^* \Delta c_{t-1}^*) = E(\Delta c_t^* \Delta c_{t+1}^*) = -\sigma_{u^c}^2$$

The other parameters of interest are still identified by (C3)-(C5), replacing the unobserved  $c$  with the measured  $c^*$ . One obvious reason for the presence of measurement error in consumption is our imputation procedure. From (1), we can write  $c_t^* = c_t + \beta (D_{it})^{-1} e_{it}$ . Note that the measurement error  $\beta (D_{it})^{-1} e_{it}$  will be non-stationary (which we account for in estimation).

*Measurement error in income.*—Assume now that income is also measured with error, i.e.,

$$y_{i,t}^* = y_{i,t} + u_{i,t}^y$$

Now it can be proved that  $\phi$  and  $\sigma_{u^c}^2$  are still identified by (C3) and (C6) (replacing the unobserved  $y$  and  $c$  with the measured  $y^*$  and  $c^*$ ). However,  $\sigma_\varepsilon^2$  and  $\sigma_{u^y}^2$  cannot be told apart, and  $\psi$  (as well as  $\sigma_\xi^2$ ) thus remains unidentified.<sup>8</sup> It is possible however to put a lower bound on  $\psi$  using the fact that:

$$\psi \geq \frac{E(\Delta c_t^* \Delta y_{t+1}^*)}{E(\Delta y_t^* \Delta y_{t+1}^*)}$$

Thus it is possible to argue that the estimate of  $\psi$  in Tables VI-VIII is downward biased due to measurement error in income. Using estimates contained in Meghir and Pistaferri (2004), a back-of-the-envelope calculation shows that the variance of measurement error in earnings accounts for approximately 30 percent

<sup>8</sup>If one is willing to assume that  $\varepsilon$  is non-stationary while  $u^y$  is stationary, then  $\psi$  is effectively identified using

$$\psi = \frac{E(\Delta c_t^* \Delta y_{t+1}^* - \Delta c_{t-1}^* \Delta y_t^*)}{E(\Delta y_t^* \Delta y_{t+1}^* - \Delta y_{t-1}^* \Delta y_t^*)}.$$

of the variance of the overall transitory component of earnings. Given that our estimate of  $\psi$  is close to zero in most cases, an adjustment using this inflation factor would make little difference empirically. To give an example, the estimate of  $\psi$  in Table 6, column 1, would increase from 0.055 to 0.079.

Using a similar reasoning, one can argue that we have an upper bound for  $\sigma_\xi^2$ , in that

$$\sigma_\xi^2 \leq E(\Delta c_t^* (\Delta c_{t-1}^* + \Delta c_t^* + \Delta c_{t+1}^*)) - \frac{[E(\Delta c_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*))]}{E(\Delta y_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*))} + \frac{[E(\Delta c_t^* \Delta y_{t+1}^*)]^2}{E(\Delta y_t^* \Delta y_{t+1}^*)}$$

The bias, however, is likely negligible. Using the same out-of-the-envelope calculation above, we calculate that the estimate of  $\sigma_\xi^2$  in Table 6, column 1, would decrease from 0.0122 to 0.0121. For this reason, in what follows we assume for simplicity that income is measured without error in deriving the various identification restrictions.

*Non-stationarity.*—Allowing for non-stationarity and with  $T$  years of data, the following moments can be used to identify the variance of the permanent shock:

$$(C7) \quad E(\Delta y_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*)) = \sigma_{\zeta,s}^2$$

for  $s = 3, 4, \dots, T-1$ . The variance of the transitory shock can be identified using:

$$(C8) \quad -E(\Delta y_s^* \Delta y_{s+1}^*) = \sigma_{\varepsilon,s}^2$$

for  $s = 2, 3, \dots, T-1$ . With an MA(1) process for the transitory component, the analog of (C7) and (C8) become:

$$E(\Delta y_s^* (\Delta y_{s-2}^* + \Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^* + \Delta y_{s+2}^*)) = \sigma_{\zeta,s}^2$$

for  $s = 4, 5, \dots, T-2$ , and (assuming  $\theta$  is being already identified)<sup>9</sup>

$$-E(\Delta y_s^* \Delta y_{s+2}^*) = \theta \sigma_{\varepsilon,s}^2$$

for  $s = 2, 3, \dots, T-2$ . In our case,  $s = 1$  corresponds to 1978 and  $s = T$  corresponds to 1992. These are the restrictions that we impose in the empirical analysis.

The other parameters of interest ( $\sigma_{u,c}^2, \phi, \psi, \sigma_\xi^2$ ) can be identified using:

$$-E(\Delta c_s^* \Delta c_{s+1}^*) = \sigma_{u,c}^2$$

$$\frac{E(\Delta c_s^* \Delta y_{s+1}^*)}{E(\Delta y_s^* \Delta y_{s+1}^*)} = \psi$$

pooling data for  $s = 2, 3, \dots, T-1$ , and

$$\frac{E(\Delta c_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*))}{E(\Delta y_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*))} = \phi$$

$$E(\Delta c_s^* (\Delta c_{s-1}^* + \Delta c_s^* + \Delta c_{s+1}^*)) - \frac{[E(\Delta c_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*))]}{E(\Delta y_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*))} + \frac{[E(\Delta c_s^* \Delta y_{s+1}^*)]^2}{E(\Delta y_s^* \Delta y_{s+1}^*)} = \sigma_\xi^2$$

pooling data for  $s = 3, 4, \dots, T-2$ .

Note that we can allow for time-varying insurance parameters:

$$\Delta c_s = \xi_s + \phi_s \zeta_s + \psi_s \varepsilon_s + \Delta u_s^c$$

which would be identified by the moment conditions:

$$\frac{E(\Delta c_s^* \Delta y_{s+1}^*)}{E(\Delta y_s^* \Delta y_{s+1}^*)} = \psi_s$$

$$\frac{E(\Delta c_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*))}{E(\Delta y_s^* (\Delta y_{s-1}^* + \Delta y_s^* + \Delta y_{s+1}^*))} = \phi_s$$

for all  $s = 2, 3, \dots, T-1$  and  $s = 3, 4, \dots, T-2$  respectively. These are the moment conditions that we use when we allow the insurance parameters to vary over time.<sup>10</sup> Equation (C6) still identifies the variance of measurement error in consumption pooling all available data. Again, this is because consumption growth is autocorrelated only because of measurement error (in the absence of it, it would be a martingale).

*More general model.*—Suppose consumption growth is now given by

$$\Delta c_s^* = \xi_s + \phi_0 \zeta_s + \phi_1 \zeta_{s-1} + \psi_0 \varepsilon_s + \psi_1 \varepsilon_{s-1} + \Delta u_s^c$$

while income growth is still:

$$\Delta y_s^* = \zeta_s + \varepsilon_s - \varepsilon_{s-1}$$

<sup>9</sup>The parameter  $\theta$  is identified by non-linear moment conditions, which we omit here.

<sup>10</sup>We experienced convergence problems in the most flexible specification when we allow for yearly variation in  $\phi$  and  $\psi$ . We thus imposed the restriction that they are the same across the sub-periods 1979-84 and 1985-92.

In this case, we assume consumption growth depends on current and lagged income shocks. The parameters to identify (in the stationary case for simplicity) are  $\phi_0, \phi_1, \psi_0, \psi_1, \sigma_\xi^2, \sigma_{uc}^2, \sigma_\zeta^2$ , and  $\sigma_\varepsilon^2$ . The variances of the income shocks are still identified by:

$$E(\Delta y_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*)) = \sigma_\zeta^2$$

and:

$$E(\Delta y_t^* \Delta y_{t-1}^*) = E(\Delta y_{t+1}^* \Delta y_t^*) = -\sigma_\varepsilon^2$$

However, one can prove that, of all the ‘‘insurance’’ coefficients, only  $\psi_0$  can be identified, using

$$\frac{E(\Delta c_t^* \Delta y_{t+1}^*)}{E(\Delta y_t^* \Delta y_{t+1}^*)} = \psi_0$$

while all the others remain not identified in the absence of further restrictions. For example, the expression we used to identify  $\phi$  in the baseline scenario,

$$\frac{E(\Delta c_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*))}{E(\Delta y_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*))}$$

would now identify the sum ( $\phi_0 + \phi_1$ ). Increasing the number of lags of income shocks in the consumption income growth equation has no effects:  $\psi_0$  is still identified, while the other insurance parameters are not.

## Appendix D: Estimation details

The two basic vectors of interest are:

$$\Delta \mathbf{c}_i = \begin{pmatrix} \Delta c_{i,1} \\ \Delta c_{i,2} \\ \dots \\ \Delta c_{i,T} \end{pmatrix} \text{ and } \Delta \mathbf{y}_i = \begin{pmatrix} \Delta y_{i,1} \\ \Delta y_{i,2} \\ \dots \\ \Delta y_{i,T} \end{pmatrix}$$

where, for simplicity, we indicate with 0 the first year in the panel (1978) and with  $T$  the last (1992), and the reference to age has been omitted. Since PSID consumption data were not collected in 1987 and 1988, the vector  $\Delta c_i$  is understood to have  $\dim(\Delta y_i) - 3$ , i.e., the rows with missing consumption data have already been swept out from  $\Delta c_i$ . Moreover, if the individual was not interviewed in year  $t$ , we replace the unobservable  $\Delta c_{i,t}$  and  $\Delta y_{i,t}$  with zeros. Conformably with the vectors above, we define:

$$\mathbf{d}_i^c = \begin{pmatrix} d_{i,1}^c \\ d_{i,2}^c \\ \dots \\ d_{i,T}^c \end{pmatrix} \text{ and } \mathbf{d}_i^y = \begin{pmatrix} d_{i,1}^y \\ d_{i,2}^y \\ \dots \\ d_{i,T}^y \end{pmatrix}$$

where  $d_{i,t}^c = 1 \{ \Delta c_{i,t} \text{ is not missing} \}$  and  $d_{i,t}^y = 1 \{ \Delta y_{i,t} \text{ is not missing} \}$ . Overall, this notation allows us to handle in a simple manner the problems of unbalanced panel data and of missing consumption data in 1987 and 1988.

Stacking observations on  $\Delta y$  and  $\Delta c$  (and on  $d^c$  and  $d^y$ ) for each individual we obtain the vectors:

$$\mathbf{x}_i = \begin{pmatrix} \Delta \mathbf{c}_i \\ \Delta \mathbf{y}_i \end{pmatrix} \text{ and } \mathbf{d}_i = \begin{pmatrix} \mathbf{d}_i^c \\ \mathbf{d}_i^y \end{pmatrix}$$

Now we can derive:

$$\mathbf{m} = \text{vech} \left\{ \left( \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i' \right) \oslash \left( \sum_{i=1}^N \mathbf{d}_i \mathbf{d}_i' \right) \right\}$$

where  $\oslash$  denotes an elementwise division. The vector  $m$  contains the estimates of  $\text{cov}(\Delta y_t, \Delta y_{t+s})$ ,  $\text{cov}(\Delta y_t, \Delta c_{t+s})$ , and  $\text{cov}(\Delta c_t, \Delta c_{t+s})$ , a total of  $T(2T+1)$  unique moments.<sup>11</sup> To obtain the variance-covariance matrix of  $m$ , define conformably with  $m$  the individual vector,  $\mathbf{m}_i = \text{vech} \{ \mathbf{x}_i \mathbf{x}_i' \}$ .

The variance-covariance matrix of  $m$  that can be used for inference is:

$$\mathbf{V} = \left[ \sum_{i=1}^N ((\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})') \otimes (\mathbf{D}_i \mathbf{D}_i') \right] \oslash \left( \sum_{i=1}^N \mathbf{D}_i \mathbf{D}_i' \right)$$

where  $\mathbf{D}_i = \text{vech} \{ \mathbf{d}_i \mathbf{d}_i' \}$  and  $\otimes$  denotes an elementwise product. The square roots of the elements in the main diagonal of  $\mathbf{V}$  provide the standard errors of the corresponding elements in  $m$ .

What we do in the empirical analysis is to estimate models for  $m$ :

$$\mathbf{m} = f(\boldsymbol{\Lambda}) + \boldsymbol{\Upsilon}$$

where  $\boldsymbol{\Upsilon}$  captures sampling variability and  $\boldsymbol{\Lambda}$  is the vector of parameters we are interested in (the variances of the permanent shock and the transitory shock, the partial insurance parameters, etc.). For instance the mapping from  $m$  to  $f(\boldsymbol{\Lambda})$  is:

<sup>11</sup>In practice there are less than  $T(2T+1)$  moments because data on consumption are not available all years.

$$\begin{pmatrix} \text{var}(\Delta c_1) \\ \text{cov}(\Delta c_1, \Delta c_2) \\ \dots \\ \text{cov}(\Delta c_1, \Delta c_T) \\ \dots \end{pmatrix} = \begin{pmatrix} \phi^2 \text{var}(\zeta_1) + \psi^2 \text{var}(\varepsilon_1) + \text{var}(\xi_1) + \text{var}(u_1^c) + \text{var}(u_0^c) \\ -\text{var}(u_1^c) \\ \dots \\ 0 \\ \dots \end{pmatrix} + \boldsymbol{\Upsilon}$$

We solve the problem of estimating  $\Lambda$  by minimizing:

$$\min_{\Lambda} (\mathbf{m} - f(\Lambda))' \mathbf{A} (\mathbf{m} - f(\Lambda))$$

where  $A$  is a weighting matrix. Optimal minimum distance (OMD) imposes  $A = V^{-1}$ , equally weighted minimum distance (EWMD) imposes  $A = I$ , and diagonally-weighted minimum distance (DWMD) requires that  $A$  is a diagonal matrix with the elements in the main diagonal given by  $\text{diag}(\mathbf{V}^{-1})$ .

For inference purposes we require the computation of standard errors. Gary Chamberlain [1984] shows that these can be obtained as:

$$\text{var}(\widehat{\Lambda}) = (\mathbf{G}' \mathbf{A} \mathbf{G})^{-1} \mathbf{G}' \mathbf{A} \mathbf{V} \mathbf{A} \mathbf{G} (\mathbf{G}' \mathbf{A} \mathbf{G})^{-1}$$

where  $G = \frac{\partial f(\Lambda)}{\partial \Lambda} \Big|_{\Lambda = \widehat{\Lambda}}$  is the Jacobian matrix evaluated at the estimated parameters  $\widehat{\Lambda}$ .