Risk Sharing*

Martin Gervais
University of Western Ontario
gervais@uwo.ca

Paul Klein
University of Western Ontario
paul.klein@uwo.ca

First draft: November 18, 2004
This draft: December 7, 2004

Abstract

The purpose of this paper is to measure the degree to which households smooth consumption in the presence of variable earnings using data from the Survey of Consumer Expenditure (CEX). The main problem of the CEX is that consumption and income are not observed for coincident periods. We show that an ordinary least squares (OLS) estimator in which the unobserved regressor is replaced by its conditional expectation is consistent. We construct this conditional expectation by imposing an AR(1) structure on monthly income. We contrast our estimates with the output of two estimators used in the literature: one is an OLS estimator in which the unobserved regressor is replaced by a proxy, the other is an instrumental variable (IV) estimator which uses another proxy as an instrument. We show that while the first (OLS) estimator tends to overstate the degree of risk-sharing, the second (IV) estimator grossly understates it.

Journal of Economic Literature Classification Numbers: C13; C8; D12; E21

Keywords: Risk Sharing; Income Risk; Consumption Risk; Projection

*We would like to thank seminar participants at the 2004 CEA meeting in Toronto, Dalhousie University, and The University of Western Ontario, as well as Lance Lochner, Lars Osberg, Fabrizio Perri, Chris Robinson and Todd Stinebrickner for useful comments. Michio Suzuki provided excellent research assistance. Both authors gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

In this paper we propose a method for measuring the degree to which households smooth variable earnings using data from the Survey of Consumer Expenditures (CEX). Following Dynarski and Gruber (1997), we operationalize the notion of smoothing in terms of the population regression coefficient of the idiosyncratic consumption change on the idiosyncratic income change. A value of zero for this coefficient is interpreted as perfect consumption risk sharing, and a value of one is interpreted as autarky.

Our analysis deals with the main problem of the CEX, which is that consumption and income are not observed for coincident periods. In particular, a household reports information regarding 7 quarters (21 months). Consumption expenditure data is available for the last four of these quarters, and income data is available for two 12-month periods, one covering the first 12 months and the other covering the last 12 months.

The paper makes a critical contribution and a constructive one. The critical contribution is to show that the approach of Dynarski and Gruber (1997) yields misleading results; specifically, it tends to understate the degree of consumption risk sharing. The constructive contribution is to present an alternative estimator, which is shown to be consistent under fairly weak assumptions.

The fundamental problem of the approach of Dynarski and Gruber (1997) is that it involves regressing the consumption change from quarter 4 to quarter 7 on the income change from quarters 1–4 to quarters 4–7. Thus the actual regressor, the income change from quarter 4 to quarter 7, is replaced by a proxy variable, giving rise to measurement error. Dynarski and Gruber (1997) do not consider this source of measurement error, but do take seriously the possibility of reporting error in income. In order to deal with this possible reporting error, they use an instrumental variables (IV) estimator. As an instrument they use the gross amount of the last pay cheque, reported in the CEX for month 12 and month 21.

The structure of the CEX renders this IV approach invalid, since the measurement error is by construction strongly negatively correlated with the instrument. Intuitively, if true income in quarter 7 (months 19–21) is high relative to its proxy (income in months 10 to 21), then measured income change is low relative to true income change, and a negative measurement error arise. But if income in months 19–21 is high relative to income in months 10–18, then the income change between month 21 and month 12, the instrument, is likely to be high. As a result, the IV estimator has a strong upward bias, thus understating the degree of risk sharing in the data.[1]

Our alternative estimator is based on the following simple result. If the proxy for true income is such that the resulting measurement error is orthogonal to the proxy itself, then the OLS estimator using the proxy as the regressor is consistent.[2] The required orthogonality property is achieved by replacing the unobserved regressor by its conditional expectation. In order to compute the conditional expectation of income contemporaneous with our consumption measure, we impose an AR(1) structure on monthly income whose parameters are estimated using the generalized method of moments (GMM).

The main result is that we can reject the hypothesis of perfect consumption risk sharing, but that the degree of risk sharing is quite high. In particular, it is much higher than it would appear if measured using the method of Dynarski and Gruber (1997).

The paper is organized as follows. In Section 2 we describe the CEX data. In Section 3 we describe the instrumental variables approach of Dynarski and Gruber (1997) and show why it is invalid. Section 4 describes our solution to the asynchronicity problem. Section 5 presents some simulation results, contrasting the properties of existing estimators with our own. In Section 6 we present some estimation results from the CEX data. Section 7 concludes.

[1] The IV estimate in Dynarski and Gruber (1997) is more than three times greater than the OLS estimate, which they recognize is inconsistent with a simple measurement error explanation.

[2] This is in contrast to "classical" measurement error, where the measurement error is orthogonal to true income.

2 Description of the Data

The problem with measuring the degree to which households smooth consumption in the presence of income variability has always been the scarcity of reliable consumption data. It is well known that both the CPS and the PSID contain high quality income data. However, the CPS provides no consumption data and the PSID only has information on food and housing consumption, which is clearly insufficient. The Survey of consumer expenditures (CEX) is the only survey in the U.S. which collects detailed consumption data. In addition, the CEX collects a limited amount of income data, which essentially makes the CEX the only dataset suitable for our purpose.[3] Another desirable feature of the CEX is that is provides a (short) panel for both consumption and income. The structure of the Survey, however, is not ideal.

Each household or consumer unit (CU) in the Survey is interviewed 5 times.[4] Interviews occur every 3 months, but the interview month varies across CU's: if the first interview occurs in January of a given year, the last interview takes place in January of the following year. The first interview is only used to collect data on characteristics of the CU and its members as well as information on some durable goods. Whereas consumption data pertaining to each of the last three months is collected at each subsequent interview, income data is only collected during the second and fifth interviews and pertains to the last 12 months.

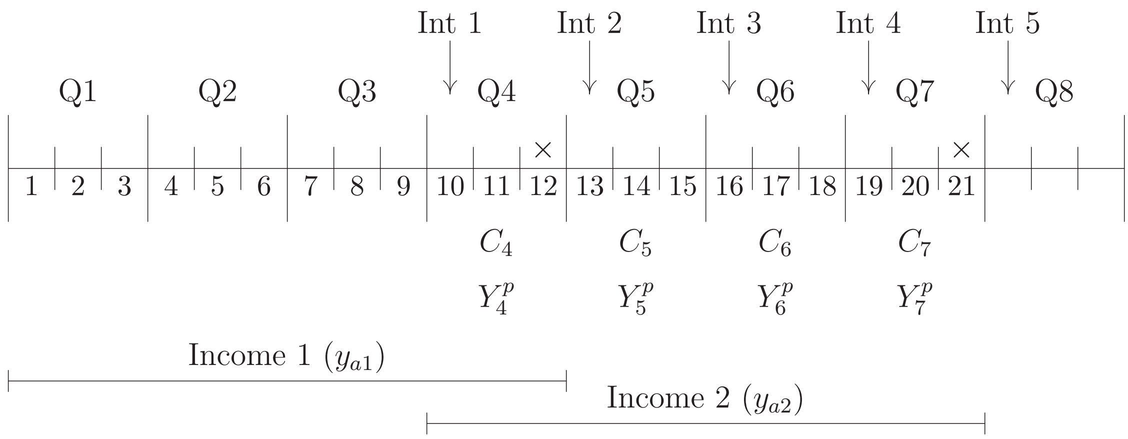
Figure 1 illustrates the life of a CU in the survey for one of three possible cases. In this Figure quarters, labelled Q1 to Q8, refer to calendar quarters.[5] Accordingly, this figure applies to households whose interview month corresponds to the first month of a calendar quarter. Notice that while the 5 interviews only span 12 months, the data

[3]Attanasio and Davis (1996) combine high quality consumption data from the CEX to high quality income data from the CPS. While this strategy is suitable to study risk sharing across groups of individuals, it masks the degree of risk sharing at the household level as idiosyncratic risk washes out in the aggregation procedure.

[4]The primary sampling unit in the CEX is called a consumer unit. A consumer unit consists of individuals who are either related or share their income to make joint expenditures. The CEX makes a subtle distinction between households and consumer units, but we use household and CU interchangeably.

[5]Note, however, that Q1 need not refer to the first calendar quarter of a year.

Figure 1: Life of an Interviewee: Case 1

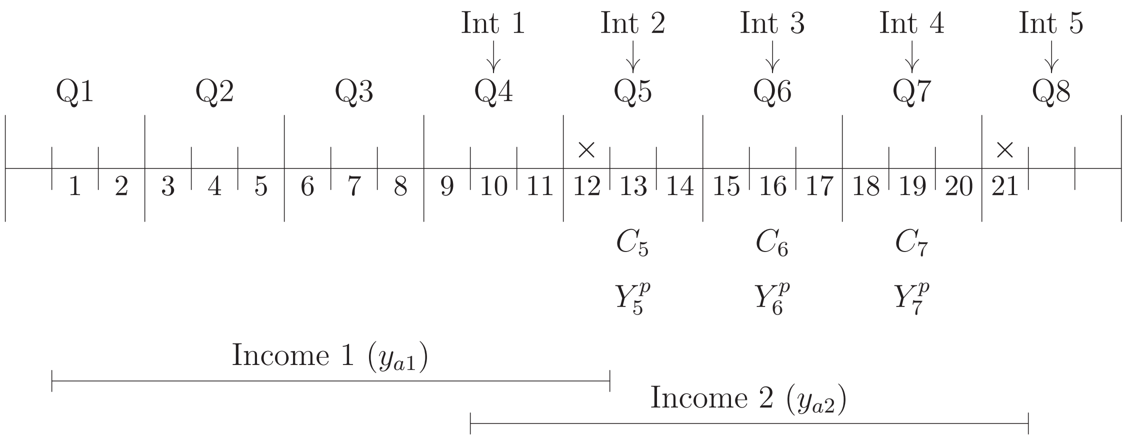


collected span 7 quarters which, in this case, correspond to calendar quarters. Since annual income data (y_{a1} and y_{a2}) are collected at the second and fifth interviews, they overlap for 3 months and thus only span 21 months. The $\times$'s in the figure (at months 12 and 21 in this case) indicate that information regarding the last pay check received by members of the household is collected for these months.[6] Notice that since these households are interviewed in the first month of a calendar quarter, their 4 consumption observations actually correspond to 4 calendar quarters, labelled C_4 though C_7. Finally, the Y^p's refer to projected income, which we will explain in detail in section 4.

Figures 2 and 3 illustrate the life of a household whose interview month corresponds to the second and third month of a calendar quarter, respectively. Although the structure of the data is the same as the previous case, the consumption observations no longer coincide with calendar quarters. Nevertheless, we can construct three calendar quarter observations of consumption, labelled C_5 to C_7, as consumption data, although only collected every three months, pertains to *each* of the past three months. Accordingly, our first observation of consumption (C_5) corresponds to consumption

[6]This last pay check need not refer only to that month, not does it need to cover the entire month. Fortunately, information on the period of time the pay check covers is also collected.

Figure 2: Life of an Interviewee: Case 2



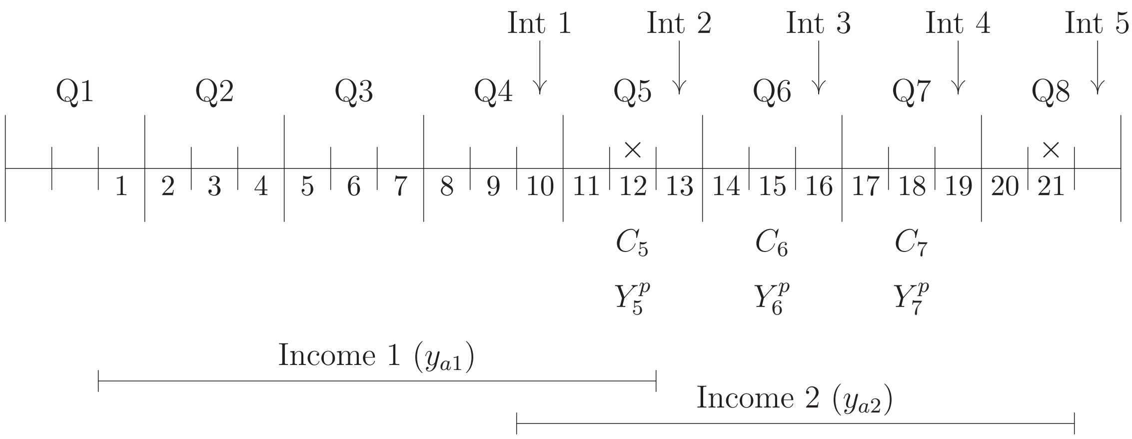
data collected in the second interview for month 12 in addition to data collected in the third interview for months 13 and 14. We do, however, lose the equivalent of one quarter of consumption information as the data only fully covers three calendar quarters.[7]

2.1 Income Data

Our definition of income consists of after-tax labor earnings plus transfers minus mandatory deductions. More precisely, for every CU, yearly income is defined as wages and salary (including all compensations from the employer) plus a fraction (0.864) of self-employment (farm and non-farm) income. Since the government-run social security system is mandatory, we treat social security as a system that taxes working individuals and makes transfers to retired (or old) individuals. Accordingly, we add social security transfers in our definition of income. We also include the following government transfers: unemployment compensation, public assistance and

[7]Dynarski and Gruber (1997) and Krueger and Perri (2003) do not face this problem as they only use the information collected in the second and fifth interviews, irrespective of whether they correspond to calendar quarters or not.

Figure 3: Life of an Interviewee: Case 3



welfare payments, as well as other transfers.[8] From that amount, we then deduct total taxes paid (federal, state and local, including property taxes, all net of refunds), social security contributions, and (government and railroad) retirement contributions. Total income is then deflated by the CPI for the relevant 12 months.[9]

To extract the idiosyncratic component of income, we regress log income on a constant, a cubic in age, aggregate log GDP per head (not seasonally adjusted), the number of earners in the household and dummies representing marital status, education and race; we then keep the residuals.[10] Rather than use an explicit individual fixed effect, we regress consumption changes on income changes, thus eliminating any individual fixed effect.

[8] Since food stamps data is only available from 1985, we do not include it in our definition of income.

[9] A future version of the paper will deal with top-coding and outlier issues.

[10] Dynarski and Gruber (1997) do not take logs, and so their numerical results are not directly comparable to ours.

2.2 Consumption Data

We define quarterly consumption as the expenditure value of all items purchased by the household during the quarter, except for cars and houses purchased. Our measure of consumption includes all the expenses born by homeowners, but does not impute a service flow. By contrast, the rent expenses of renters are included. For all other durables, we do not attempt to impute any income from the stocks, and simply add expenditures during the quarter. All observations are deflated by the CPI.

Idiosyncratic consumption is exctracted in the same way as income except that we also introduce a seasonal dummy in order to take care of the possibility that consumption varies systematically with the time of year.

2.3 Our Sample

Our sample runs from the first quarter of 1980 to the last quarter of 2001. We exclude CU's whose income is considered incomplete.[11] We also exclude households whose characteristics are inconsistent over time, either because individuals grow younger or who age by more than one year from one quarter to the next, get less or more educated too fast, or individuals who undergo a sex or race change. Because of a coding mistake in the CEX data, all households whose interviews span the years 1981 and 1982 are dropped. Finally, we also exclude households with at least one consumption or income observation equal to zero.

3 Measurement Error and the IV Solution

The fundamental problem with estimating the degree to which households can smooth consumption in the presence of income risk from CEX data is the fact that we do not observe consumption and income for the same periods of time. Having no observations

[11] The CEX produces a variable whose value is 1 if the respondent's income data is complete and 0 otherwise.

of income corresponding to consumption, a natural though problematic strategy is to use the change in annual income as a proxy for the income change between quarters Q7 and Q4. Doing so, however, introduces measurement error in the income change even in the absence of misreporting.

As is well known, the OLS estimator of the consumption response to an income change is biased toward zero in the presence of classical measurement error in the income change; this is noted both by Altonji and Siow (1987) and Dynarski and Gruber (1997). Both of these papers use an instrumental variable approach in order to deal with the problem, though they differ in their choice of instrument. Altonji and Siow (1987) use data from the Panel Study of Income Dynamics (PSID) where the only consumption data, as we stress above, is on food and housing; therefore their work is of limited relevance here. Hence we confine our attention to Dynarski and Gruber (1997). We will argue that because of the particular way in which measurement error arises in their approach, their particular choice of instrument is invalid.

The problem is to estimate β in the regression equation

$$C = \beta Y + \varepsilon, \tag{1}$$

where C measures the change in the idiosyncratic component of consumption $(C_7 - C_4)$, Y measures the change in the idiosyncratic component of income $(Y_7 - Y_4)$, and $\mathsf{E}[\varepsilon] = \mathsf{E}[Y\varepsilon] = 0$. However, since Y is unobserved, it is replaced in Dynarski and Gruber (1997) by a proxy $\widehat{Y}$, given by the difference between the two yearly observations $(y_{a2} - y_{a1})$. Even in the absence of reporting error, this gives rise to measurement error. We will denote it by η, defined via $\widehat{Y} = Y + \eta$. The OLS estimator in this case is not likely to be consistent. We have

$$\operatorname*{plim}_{N\to\infty} \beta_N^{\text{OLS}} = \frac{\mathsf{E}[\widehat{Y}C]}{\mathsf{E}[\widehat{Y}^2]} = \frac{\beta\mathsf{E}[\widehat{Y}Y] + \mathsf{E}[\widehat{Y}\varepsilon]}{\mathsf{E}[Y^2] + \mathsf{E}[\eta^2] + 2\mathsf{E}[Y\eta]}.$$

If we assume that $\mathsf{E}[\eta\varepsilon] = 0$, then

$$\operatorname*{plim}_{N\to\infty} \beta_N^{\text{OLS}} = \beta\frac{\mathsf{E}[Y\widehat{Y}]}{\mathsf{E}[\widehat{Y}^2]}.$$

If the measurement error is "classical," i.e. $\mathsf{E}[Y\eta] = 0$, we have

$$\frac{\mathsf{E}[Y\widehat{Y}]}{\mathsf{E}[\widehat{Y}^2]} = \frac{\mathsf{E}[Y^2]}{\mathsf{E}[Y^2] + \mathsf{E}[\eta^2]} < 1,$$

so that the OLS estimator is asymptotically biased towards zero, thereby overstating the degree of risk sharing.[12]

To deal with the problem of measurement error, Dynarski and Gruber (1997) use an instrumental variable (IV) approach with a second measure of the income change as an instrument. As shown in Figure 1, the CEX provides information about the amount of the CU's last paycheck, as well as its frequency, which we refer to as (log) monthly incomes y_{12}^m and y_{21}^m. This is then used to define the instrument via $Z = y_{21}^m - y_{12}^m$. This instrument is invalid because the measurement error arises from asynchronicity, as opposed to misreporting of income. To see this, notice that

$$\operatorname*{plim}_{N\to\infty} \beta_N^{\mathrm{IV}} = \frac{\mathsf{E}[CZ]}{\mathsf{E}[\widehat{Y}Z]}.$$

Assuming that $\mathsf{E}[\eta\varepsilon] = 0$, we have

$$\operatorname*{plim}_{N\to\infty} \beta_N^{\mathrm{IV}} = \beta\frac{\mathsf{E}[YZ]}{\mathsf{E}[YZ] + \mathsf{E}[Z\eta]}. \tag{2}$$

This instrumental variable strategy will thus only be valid if the measurement error in change in annual earnings is uncorrelated with the instrument, $\mathsf{E}[Z\eta] = 0$.[13]

There are very strong reasons to believe that this condition is violated in this context, simply because of the structure of the CEX as illustrated in Figure 1. As an example, suppose that income in the second year (y_{a2}) is high relative to what it is trying to measure, i.e. relative to income in the seventh quarter (Y_7). This would be interpreted as measurement error within the IV strategy. That is, because Y_7 is unobservable and proxied by y_{a2}, a positive measurement error results when one computes Y as $y_{a2} - y_{a1}$. Similarly, if y_{a2} is high relative to Y_7, it must be that income during quarters Q4 to

[12] Notice that if $\mathsf{E}[Y\eta] > 0$, then the estimator is biased even further towards zero. Our analysis in section 5 suggests that this is indeed the case.

[13] Although Dynarski and Gruber (1997) assume that the measurement error η is uncorrelated with the measurement error in Z, they do not explicitly discuss whether $\mathsf{E}[Z\eta] = 0$.

Q6 was also high relative to Y_7. In particular, income in the first quarter is likely high, and income in the third month of Q4, which corresponds to the instrument y_{12}^m, is also likely to be high. But if y_{12}^m is high relative to Y_7, it is also likely to be high relative to income in month 21, i.e. relative to y_{21}^m. That means that the instrument, $y_{21}^m - y_{12}^m$, is likely to be negative. For such reasons, one would expect measurement error to be negatively correlated with the instrument. As Equation (2) shows, the IV estimator in that case is biased upward, that is, this estimator tends to understate the degree of risk sharing.

4 Projection-based Estimation

4.1 The Projection Estimator

We have seen that, when the regressor is measured with error, the OLS estimator is asymptotically biased by the factor

$$\frac{\mathsf{E}[Y\widehat{Y}]}{\mathsf{E}[\widehat{Y}^2]}$$

where $\widehat{Y} = Y + \eta$. We have also seen that if measurement error is classical, i.e. $\mathsf{E}[Y\eta] = 0$, then this ratio is strictly less than one. However, if $\mathsf{E}[\widehat{Y}\eta] = 0$, then $\mathsf{E}[\widehat{Y}^2] = \mathsf{E}[\widehat{Y}Y]$ and hence

$$\frac{\mathsf{E}[Y\widehat{Y}]}{\mathsf{E}[\widehat{Y}^2]} = 1.$$

Thus consistency is achieved if the measurement error is orthogonal not to the regressor but to its proxy.

This result can be generalized to higher dimensions.[14]

Proposition 1 *Let X, Y and ε be mean-zero random vectors. Consider the linear regression model*

$$Y = \beta X + \varepsilon$$

[14] A related result can be found in the literature on the missing data problem in multiple regression analysis, e.g. Buck (1960) and Dagenais (1973).

where $\mathsf{E}[X\varepsilon'] = 0$. *Suppose* X *is measured with error and denote the measure by* $\widehat{X}$. *If* $\mathsf{E}[\widehat{X}\varepsilon'] = 0$ *(or equivalently, that* $\mathsf{E}[(\widehat{X} - X)\varepsilon'] = 0$*) and*

$$\mathsf{E}[(X - \widehat{X})\widehat{X}'] = 0.$$

Then

$$\beta = \mathsf{E}[Y\widehat{X}']\mathsf{E}[\widehat{X}\widehat{X}']^{-1}$$

and hence the OLS estimator of β *using* $\widehat{X}$ *as a regressor is consistent.*

Proof. Using $\mathsf{E}[\widehat{X}\varepsilon'] = 0$ we get $\mathsf{E}[Y\widehat{X}'] = \beta\mathsf{E}[X\widehat{X}']$. Using $\mathsf{E}[(X - \widehat{X})\widehat{X}'] = 0$ we get $\mathsf{E}[\widehat{X}\widehat{X}'] = \mathsf{E}[X\widehat{X}']$. ■

One proxy for the regressor that will certainly have the desired orthogonality property is the linear projection of the unobserved regressor on something we can observe, such as our two observations of annual income. By definition of the projection, the projection error is orthogonal to the projection itself.[15] The only remaining problem is to construct that projection. To do that, we define a vector that we can observe, namely

$$W = \begin{bmatrix} y^i_{a1} \\ y^i_{a2} \end{bmatrix}.$$

Denote the actual (unobserved) income change by Y and its linear projection on W by Y^p. Then

$$Y^p = \alpha W$$

where

$$\alpha = \mathsf{E}[YW']\mathsf{E}[WW']^{-1}.$$

Thus in order to compute the projection we need to estimate the variance of annual income changes and the covariance between quarterly income changes and annual income changes. For the latter we need to impose some structure.

[15] See, for example, Chung (2001).

4.2 Constructing the Projection: Structure and GMM

Our strategy is as follows. First impose a parameterized structure on the data generating process of income. Then use GMM to estimate the parameters of that structure. Then we use the estimated parameter values to compute the desired covariance matrix.

4.2.1 Income Process

Let y_t^i denote (log) monthly income for individual i. Assume that the stochastic process governing y_t^i is given by

$$y_t^i = \rho y_{t-1}^i + \varepsilon_t^i, \tag{3}$$

where ε_t^i is the idiosyncratic shock received by individual i in period t and ρ measures the persistence of income. Recall that the data we have consist of two annual observations of income, denoted y_{a1}^i and y_{a2}^i, which overlap for exactly 3 months. What we want is a measure of quarterly income, which will be constructed from estimates of monthly income y_t^i, $t = 1 \ldots, 21$. First note that given y_1,

$$y_t^i = \rho^{t-1} y_1^i + \sum_{k=2}^{t} \rho^{t-k} \varepsilon_k^i.$$

We can then express our first annual income observation in terms of monthly income;

$$y_{a1}^i = \ln \left(\frac{1}{12} \sum_{t=1}^{12} \exp \left\{ \left(\rho^{t-1} y_1^i + \sum_{k=2}^{t} \rho^{t-k} \varepsilon_k^i \right) \right\} \right).$$

Similarly, we can express our second annual income observation in terms of monthly income;

$$y_{a2}^i = \ln \left(\frac{1}{12} \sum_{t=10}^{21} \exp \left\{ \left(\rho^{t-1} y_1^i + \sum_{k=2}^{t} \rho^{t-k} \varepsilon_k^i \right) \right\} \right).$$

As moments for GMM we use $\mathsf{E}[y_{a1}^i \, y_{a1}^i]$, $\mathsf{E}[y_{a2}^i \, y_{a2}^i]$ and $\mathsf{E}[y_{a1}^i \, y_{a2}^i]$. We have the following **approximate** results:

$$\mathsf{E}[y_{a1}^i \, y_{a1}^i] \approx \left(\frac{1-\rho^{12}}{1-\rho} \right)^2 \frac{\sigma_{y_1}^2}{12^2} + \frac{1}{(1-\rho)^2} \left(11 - 2\rho \frac{1-\rho^{11}}{1-\rho} + \rho^2 \frac{1-\rho^{22}}{1-\rho^2} \right) \frac{\sigma_\varepsilon^2}{12^2} \; ;$$

$$\mathsf{E}[y^i_{a2}\, y^i_{a2}] \approx \rho^{18} \left(\frac{1-\rho^{12}}{1-\rho}\right)^2 \frac{\sigma^2_{y_1}}{12^2}$$
$$+ \left[\left(\frac{1-\rho^{18}}{1-\rho^2}\right) \left(\frac{1-\rho^{12}}{1-\rho}\right)^2 + \frac{1}{(1-\rho)^2} \left(11 - 2\rho \frac{1-\rho^{11}}{1-\rho} + \rho^2 \frac{1-\rho^{22}}{1-\rho^2}\right) \right] \frac{\sigma^2_{\varepsilon}}{12^2} \;;$$

$$\mathsf{E}[y^i_{a1}\, y^i_{a2}] \approx \frac{\rho^9 (1-\rho^{12})^2}{(1-\rho)^2} \frac{\sigma^2_{y_1}}{12^2}$$
$$+ \left[\frac{1-\rho^{12}}{(1-\rho)^2} \left(\frac{1-\rho^9}{1-\rho} + \frac{\rho^{21}-\rho^3}{1-\rho^2} \right) + \frac{(1+\rho)(1-\rho^{11})}{1-\rho} + \frac{(1-\rho^{10})}{1-\rho} \right] \frac{\sigma^2_{\varepsilon}}{12^2} \,.$$

5 Simulations

To illustrate the properties of our estimator relative to the OLS and IV estimator, we simulate data according to the processes for income implied by the estimates from the previous section. We further assume that the true value of β is 0.2. Our results, shown in Table 1 indicate that the IV estimate is more than 5 times higher than the OLS estimates, as one would expect given the results in Dynarski and Gruber (1997). While the estimate from our proposed projection method, labelled β^{PRO}, is equal to the true value of β, the IV estimate understates the degree of risk sharing and the OLS estimate understates it.

Table 1: Simulation Results

	β^{PRO}	β^{OLS}	β^{IV}
Estimate	0.2000	0.1492	0.8084

6 Estimation results

Table 2 shows our estimate of the degree of risk sharing together with the OLS and IV estimates. The standard errors in that table were computed using a bootstrap strategy, with 1000 samples of the same number of observation as in the original sample (56,600). As expected, the OLS estimates is low and the IV estimate is high relative to our proposed projection method estimate.

Table 2: Estimates

	β^{PRO}	β^{OLS}	β^{IV}
Estimate	0.1798	0.0621	0.3616
Standard Error	0.0198	0.0047	0.0380

7 Conclusion

This paper proposes an estimation method that can be used in contexts where the regressor is unobservable but its co-variance with observables is either known or can be estimated. Our main proposition is that the OLS estimator is consistent if the proxy variable used in the regression is orthogonal to measurement error. A proxy variable with that property can easily be constructed by projecting the unobserved regressor on observables. This method can thus be used in a variety of cases where the desired regressor is unobservable but other observable variables can be used to construct a suitable proxy. Of course, some structure is needed to construct such a proxy variable.

We use this estimation technique to estimate the degree of risk sharing from CEX data. Our simulations and estimates confirm that whereas the OLS estimator overstates the degree of risk sharing, the IV estimator greatly understates it.

References

Altonji, J. G. and A. Siow (1987). Testing the response of consumption to income changes with (noisy) panel data. *Quarterly Journal of Economics 102*(2), 293–328.

Attanasio, O. and S. J. Davis (1996). Relative wage movements and the distribution of consumption. *Journal of Political Economy 104*(6), 1227–1262.

Buck, S. F. (1960). A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *Journal of the Royal Statistical Society, Series B (Methodological) 22*(2), 302–306.

Chung, K. L. (2001). *A Course in Probability Theory, Third Edition*. Academic Press.

Dagenais, M. G. (1973). The use of incomplete observations in multiple regression analysis. *Journal of Econometrics 1*, 317–328.

Dynarski, S. and J. Gruber (1997). Can families smooth variable earnings? *Brookings Papers on Economic Activity 8*(1), 229–303.

Krueger, D. and F. Perri (2003). Does income inequality lead to consumption inequality? Evidence and theory. Unpublished Manuscript.