Two Views of Inequality Over the Life-Cycle[*†]

Jonathan Heathcote
Georgetown University

Kjetil Storesletten
University of Oslo

Giovanni L. Violante
New York University

September 2004

Abstract

Data on the life-cycle profiles of inequality in wages, earnings, hours worked and consumption contain precious information for answering questions about the ability of households to insure labor market risk and about the sources of this risk. This paper demonstrates that the choice of whether to control for cohort effects or for time effects has a drastic impact on the estimated age profiles for inequality and, thus, on the answers to those questions. It also shows that time effects are required to account for the observed trends in inequality in thirty years of US data, whereas there is no evidence that cohort effects have been important.

JEL Classification Codes: C13, D31, D91, J22, J31

Keywords: Wage inequality, Consumption inequality, Cohort effects, Labor supply.

[*]Acknowledgements: we thank Dirk Krueger and Fabrizio Perri for help with the CEX consumption data. Heathcote and Violante thank the Economics Program of the National Science Foundation for financial support. Violante also thanks the CV Starr Center for support. Storesletten thanks the Institute for International Economic Studies and the Frisch Centre for support.

[†]Email addresses: Heathcote <jhh9@georgetown.edu>; Storesletten <kjstore@econ.uio.no>; Violante <glv2@nyu.edu>.

1 Introduction

This paper measures how inequality in wages, hours worked and consumption varies over the life-cycle in the United States. The steepness of the age profile for wage inequality is informative about whether wage dispersion is primarily driven by fixed-effects or life-cycle shocks. Contrasting the age profile for wage inequality with age profiles for inequality in consumption and labor supply can shed light on the degree to which agents can insure wage shocks, and can also be useful for differentiating between alternative specifications for preferences (Heathcote et al., 2004a).

The task of documenting how inequality changes with age is complicated by the sharp rise in *cross-sectional* inequality in US wages and earnings since 1980. In order to isolate the dynamics of inequality with age, we need to take a stand on what accounts for these changes over time. Wage inequality could potentially have increased in recent decades due to either (1) *time effects* – changes in the economic environment that have increased wage inequality within every age group, or (2) *cohort effects* – younger cohorts were more unequally endowed with labor market skills than older cohorts.[1] To compute life-cycle facts we must disentangle the relative contributions of time and cohort effects. This decomposition is of independent interest, since it sheds light on what factors are behind the observed increases in cross-sectional inequality.[2]

To identify age-effects and disentangle time and cohort effects we must impose restrictions on how age, time and cohort effects can interact, since these are not independent.[3] We follow Juhn et al. (1993) and Ameriks and Zeldes (2001) and assume that the effects of age, time and cohort are additively separable, and that only two of the three effects are

[1] A third possibility is that wage inequality increases with age, and the average age of the population has risen over time. However, this cannot be the source of increases in cross-sectional inequality in our sample, since the age composition of our sample is stable over time.

[2] Examples of time effects include changing returns to ability and trade liberalization. Examples of cohort effects are variations in the size of birth cohorts, and in the dispersion in education quality. Finally, factors such as a changing returns to education, deunionization, and changes in the minimum wage (which binds mostly for the young) are examples of changes in the environment that combine both cohort and time-effects.

[3] To see this, let a denote age, t denote time, and c denote cohort birth year. Let the statistic of interest be $x(a, t, c)$ for agents of age a at date t. By appropriate substitutions of the identity $c = t - a$ it is easy to see that any model of the form $x(a, t, c) = f(a, t, c)$ is equivalent to three other models that are functions of only two of the three arguments.

operative.[4] We then consider two linear regression models: one in which unrestricted age and cohort effects are present, but no time effects, and one with age and time effects, but no cohort effects. We find that age profiles of inequality are quite different depending on whether one controls for cohort or time effects. These differences lead to very different interpretations of the driving wage process, including the fraction of wage inequality that is "pre-determined" at the time of labor market entrance. We will argue that the picture of life-cycle inequality that arises from the time effects model is consistent with a standard life-cycle model with endogenous labor supply. Moreover, the life-cycle profiles suggest that while some fraction of wage inequality is insurable, insurance markets are incomplete.

We then examine the data more closely to gauge whether the observed trends in cross-sectional inequality are best captured by time effects or by cohort effects. According to the cohort effect model, growth in within-cohort inequality should be the same at each date. We find that the cohort model is contradicted by the data; growth in within-cohort inequality varies strongly with time. This holds for wage inequality, earnings inequality, and the covariance between wages and earnings. By contrast, looking at the age profile of inequality in repeated cross-sections, the slope of the age profile is relatively constant, suggesting that cohort effects can safely be abstracted from.

Section 2 describes the data, Section 3 discusses cross-sectional inequality over the period 1979-1996, Section 4 presents our two sets of estimates for inequality over the life-cycle, and Section 5 discusses the relative strengths and weaknesses of the time effect and cohort effect views.

2 Data Description

Our empirical analysis is based on two sources of data: the 1980-1997 waves of the *Panel Study of Income Dynamics* (PSID) for wages, hours worked and earnings, and the 1980-1997 waves of the *Consumer Expenditures Surveys* (CEX) for consumption.

We use the PSID data sample of Heathcote et al. (2004b, HSV henceforth), and the CEX data sample of Krueger and Perri (2002, KP henceforth) and refer to those papers

[4] In addition to simplicity, this approach has the advantage that collinearity would likely remain a problem in any approach incorporating all three effects, even if identification were formally achieved.

for details on sample selection. Most importantly, the sample is restricted to households whose "head" ("head of household" in the PSID, the "reference person" in the CEX) is aged 20-65 and works $h \in [520, 5096]$ annual hours. Hourly wages are computed as annual labor earnings divided by annual hours worked. KP's "total consumption" is non-durables plus services from durables.

Three are three important data issues that we only briefly mention here:

– *Measurement Error in the PSID:* Several papers based on the PSID Validation Studies argue that in PSID data, hours are measured with error. Since wages are defined as the ratio of earnings to hours, one implication is that the absolute magnitude of the covariance between hours and wages can be underestimated ("division-bias"). If the measurement error is "classical", then it will affect the levels of variances and covariances, but not their slopes over time or over the life-cycle, which are the focus here.

– *Measurement Error in the CEX:* While KP and Slesnick and Ulker (2004) focus exclusively on the survey datasets in the CEX, Attanasio et al. (2004) find that CEX diary data suggest a somewhat larger increase in consumption inequality over time.

– *Unit of analysis:* Wages, hours and earnings are recorded at the level of the individual worker, while consumption is measured only at the level of the household. We therefore convert household consumption data into adult equivalents using the scaling factor $\hat{c} = \sqrt{\#A + 0.5\#C}$, where $\#A$ and $\#C$ are the number of adults and children living in the household (see Slesnick and Ulker, 2004).[5]

3 Cross-Sectional Inequality over Time (1967-1996)

Despite the sharp rise in cross-sectional inequality in US wages and earnings since 1980, inequality in consumption and hours worked has been relatively stable. However, the underlying consumption-hours distribution has changed significantly: the covariance between

[5] See HSV for a discussion of the pros and cons of focussing on the individual versus the household as the unit of analysis.

consumption and hours has decreased, while the covariance between wages and hours has increased (see HSV and KP).

In HSV we use the panel dimension of the PSID to estimate a time varying stochastic process for individual wages. We find that a calibrated Bewley-Huggett-Aiyagari incomplete-markets economy with endogenous labor supply and the estimated wage dynamics can account – quantitatively – for all the trends described above. The model predicts a minor increase in the variability of consumption because our wage process attributes some of the increase in wage-inequality to transitory shocks that are easy to self-insure. Moreover, it replicates the observed rise in the wage-hours covariance: as the variance of transitory shocks increases, labor supply tracks wages more closely. Consequently, the model can account for the fact that earnings inequality has increased by more than wage inequality.

In HSV our specification for the individual wage process allows for time effects but not for cohort effects. If cohort effects were important, abstracting from them could lead to misleading predictions for the dynamics of inequality over time in other variables. Fortunately, our analysis of life-cycle data in the next section will lead us to conclude that it is reasonable to focus on time effects when modelling changes in the wage-generating process.

4 Cross-Sectional Inequality over the Life Cycle

We start by constructing variances of wages, hours, total and non-durable consumption, and the covariance between hours and wages for observations grouped by year and age. An individual is defined to be of age a if her actual age lies between $a - 2$ and $a + 2$. The typical "cell" contains several hundred observations, although the size range is large. We then estimate the life-cycle profile for each moment, removing year and cohort effects successively. To remove year effects from the typical moment $x(a, t, c)$ (e.g., the variance of wages of individuals with age index a at time t) we assume away cohort effects and run a linear regression with a full set of year and age dummies. To remove cohort effects from the moment we assume away time effects and run a linear regression with a full set of cohort and age dummies.

Figure 1 reports the results. The "time view" and the "cohort view" yield very different results for the variances of wages and earnings and the covariance between wages and hours.

The rise in wage and earnings inequality over the life cycle is substantially larger when one controls for cohort effects (by a factor of 1.6 for wages and 1.7 for earnings). For the covariance the differences are even more pronounced. When we remove time effects, the covariance falls with age, when we remove cohort effects, it rises slightly. By contrast, the age profiles for the variance of hours and the variance of consumption are largely invariant to the assumptions made about time versus cohort effects.[6]

These age profiles contain information about the driving income process (see also Storesletten et al., 2004). For example, assume that the log of an individual's wages can be decomposed into a fixed effect, a persistent autoregressive shock, and a transitory shock. The fact that both age profiles for wages are linear suggests that the persistent component is close to a unit root. The steepness of the profile identifies the conditional variance of the permanent shock. Thus, the time view suggests substantially less life-cycle risk than the cohort view, implying that a larger fraction of wage and earnings inequality is "pre-determined" prior to labor market entrance.

In Heathcote et al. (2004a), we estimate an incomplete-markets life-cycle model with endogenous labor supply. That model is quantitatively consistent both with the life-cycle profiles according to the time view, and with the evolution of cross-sectional inequality as described in Section 3. As agents age, accumulated permanent shocks account for a larger fraction of the (within-cohort) cross-sectional variance of wages. When preferences are such that agents increase leisure in response to a permanent wage increase, many of the patterns in Figure 1 emerge: the variance of consumption increases less than the variance of earnings, and the covariance between hours and wages declines with age.[7]

In contrast to the incomplete markets model proposed in Heathcote et al. (2004a), Figure 1 seems inconsistent with the existence of *complete markets* against wage risk. Complete insurance implies that workers with good (bad) wage shocks should supply more (fewer)

[6] To save space the picture for total consumption is omitted from Figure 1. Its age-profile is similar to that for non-durable consumption until age 48, but then it flattens out. Moreover, age profiles controlling for time and cohort effects are essentially identical, as for non-durable consumption.

[7] We have not formally estimated the life-cycle model under the assumption that the empirical moments are those arising from the cohort view. However, since (1) the accumulated permanent shocks must account for an even larger fraction of wage-variance than under the time view, and (2) the cohort view implies an *increasing* empirical age profile for the wage-hours correlation, we conjecture that the life-cycle model of labor supply will be rejected.

hours (see Heathcote et al., 2004c).[8] Consequently, the complete markets model has the following sharp implication: as a cohort grows older and shocks account for a larger fraction of within-cohort wage inequality, the correlation between wages and hours should increase. However, Figure 2 reveals that the wage-hours correlation is close to zero at all ages.

5 Time or Cohorts Effects?

Given the sharp differences between the cohort-view and the time-view documented above, we now try to measure the relative importance of time and cohort effects in the determination of the dynamics of inequality in the US over the last 30 years.

Consider the following general model for the moment $x(a, t, c)$ in which we assume separability between the effects of age a, time t and cohort $c = t - a$:

$$x(a, t, t-a) = g_1(a) + g_2(t) + g_3(t-a).$$

In the regressions described in the previous section, the functions g_1, and g_2 and g_3 are implemented as full sets of regression dummies.

To help identify time and cohort effects, we compute three different measures of changes in inequality:

- The change *within age* group a between t and $t+1$:

$$\Delta x^a_{t,t+1} = g_2(t+1) - g_2(t) + g_3(t+1-a) - g_3(t-a) = \Delta g_2(t) + \Delta g_3(t-a).$$

  The average (across age groups) within-age change is

$$\bar{\Delta} x^a_{t,t+1} = \Delta g_2(t) + \bar{\Delta} g_3(t).$$

- The change *within cohort* $c = t - a$ between t and $t+1$:

$$\Delta x^c_{t,t+1} = g_1(a+1) - g_1(a) + g_2(t+1) - g_2(t) = \Delta g_1(a) + \Delta g_2(t).$$

  The average (across cohorts) within-cohort change is

$$\bar{\Delta} x^c_{t,t+1} = \bar{\Delta} g_1 + \Delta g_2(t).$$

[8] This holds even if preferences are non-separable between consumption and leisure. Non-separability is needed in order to account for the rising age profile of consumption inequality (Storesletten et al., 2001).

- The *within period* $t+1$ change between age a and $a+1$:

  $$\Delta x^t_{a,a+1} = g_1(a+1) - g_1(a) + g_3((t+1)-(a+1)) - g_3((t+1)-a) = \Delta g_1(a) - \Delta g_3(t-a).$$

  The average (across age groups) within-period change is

  $$\bar{\Delta} x^t_{a,a+1} = \bar{\Delta} g_1 - \bar{\Delta} g_3(t).$$

Juhn et al. (1993) note that $\bar{\Delta} x^a_{t,t+1}$ and $\bar{\Delta} x^c_{t,t+1}$ have a common component, the time effect. Thus, in the presence of important time effects and negligible cohort effects, these two measures should be strongly correlated. This is the first hypothesis we examine. For each moment Table 1 reports the average within-cohort and within-age-group changes for six periods (four periods for the variance of consumption). First, consider the PSID moments. The correlation between average within-cohort and within-age-group changes (reported in the last column of the table) is always high and positive. In particular, note that in the three time periods from 1973-1987, when most of the rise in wage inequality took place, the magnitude of the within-age and within-cohort changes are remarkably similar for all moments.[9]

A positive correlation between within-cohort and within-age changes might occur even in the absence of time effects if the age effect $\bar{\Delta} g_1$ and the cohort effect $\bar{\Delta} g_3(t)$ were similar. The hypothesis that age effects and cohort effects happen to coincide, and that time effects are small can be examined in two ways. First, in the absence of time effects, the within-cohort change $\bar{\Delta} x^c_{t,t+1}$ should be constant across time periods. This is rejected in our data: the majority of the within cohort changes for any given variable are statistically different between periods. Second, the within-period between-age change $\bar{\Delta} x^t_{a,a+1}$ should be zero, since the age effect and the cohort effect enter with opposite offsetting signs in this expression. However, in our sample the within-period between-age change is statistically different from zero in every period. Thus both tests suggest that we should reject the hypothesis that time effects are small.

To go one step further, note that if cohort effects are small, then $\bar{\Delta} x^t_{a,a+1}$ should be constant across periods. This hypothesis cannot be rejected by the PSID moments. We conclude that the evidence speaks in favor of time effects, rather than cohort effects.

[9] The correlation coefficient restricted to the periods 1973-1977, 1978-1982, and 1983-1987 are above 0.92 for every PSID moment.

We do not analyze changes in inequality in consumption between cohorts and age-groups because changes in the variance of log consumption are essentially not significant and are therefore not useful in distinguishing between time and cohort effects.

6 Concluding Remarks

In this paper we have shown that the dynamics of cross-sectional inequality across time and age groups are consistent with the presence of time effects and absence of cohort effects, but inconsistent with the presence of cohort effects and absence of time effects.

Furthermore, attributing rising inequality over time solely to cohort effects paints a misleading picture of inequality over the life-cycle. Inequality rose dramatically within cohorts during the 1980s, so a model without time effects must attribute this increase in inequality to a very steep age profile for inequality. By contrast, a model with time effects attributes much of the rise in inequality to time effects, implying a flatter "steady state" wage profile for inequality. This might explain why Deaton and Paxson (1994), who abstracted from time effects, found a dramatic increase in consumption inequality by age using data from the 1980s – a period of sharply rising wage inequality – while we and Slesnick and Ulker (2004) find a smaller increase when extending the sample to include the 1990s.

References

Ameriks, John and Stephen P. Zeldes, (2001), How do household portfolio shares vary with age?, Unpublished manuscript, Columbia University.

Attanasio, Orazio; Erich Battistin, and Hide Ichimura, (2004), What really happened to consumption inequality in the U.S?, NBER wp 10338.

Deaton, Angus and Christina Paxson, (1994), Intertemporal choice and inequality, *Journal of Political Economy* 102, 437–467.

Heathcote, Jonathan; Kjetil Storesletten; and Giovanni L. Violante, (2004a), Insurance and opportunities: The welfare implications of rising wage dispersion, Mimeo, Georgetown University.

Heathcote, Jonathan; Kjetil Storesletten; and Giovanni L. Violante, (2004b), The macroeconomic implications of rising wage inequality in the United States, CEPR wp 4296.

Heathcote, Jonathan; Kjetil Storesletten; and Giovanni L. Violante, (2004c), A tractable framework for understanding dispersion in labor supply and consumption, Mimeo, Georgetown University.

Juhn, Chinhui; Kevin M. Murphy; and Brooks Pierce, (1993), Wage inequality and the rise in returns to skill, *Journal of Political Economy* 101, 410–42.

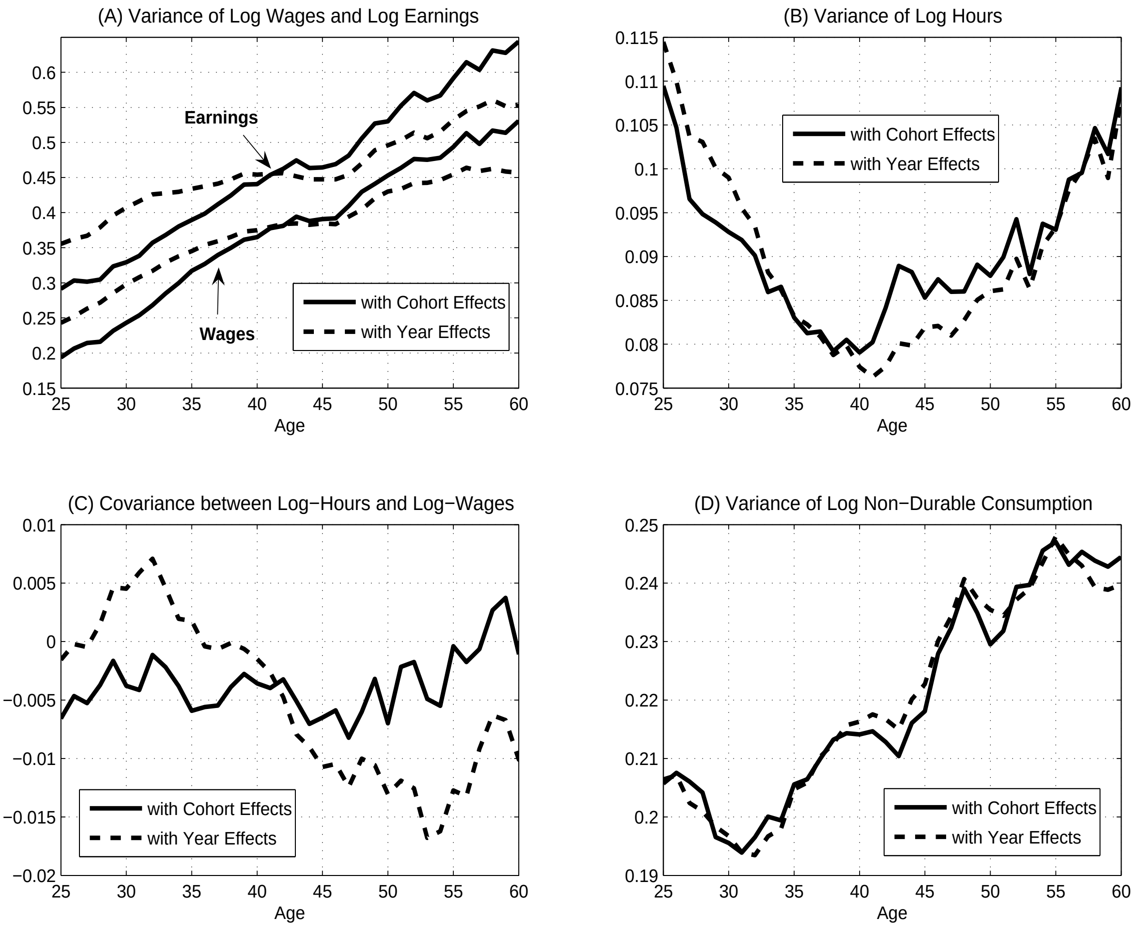
Krueger, Dirk and Fabrizio Perri, (2002), Does income inequality lead to consumption inequality? Evidence and theory, Mimeo, Stanford University.

Slesnick, Daniel T. and Aydogan Ulker, (2004), Inequality and the life cycle: Age, cohort effects, and consumption, Mimeo, University of Texas, Austin.

Storesletten, Kjetil; Christopher I. Telmer; and Amir Yaron, (2001), How important are idiosyncratic shocks? Evidence from labor supply, *American Economic Review* 91 (2), 413–417.

Storesletten, Kjetil; Christopher I. Telmer; and Amir Yaron, (2004), Consumption and risk sharing over the life cycle, *Journal of Monetary Economics* 51, 609–633.

Figure 1
Age-profiles of Inequality: Cohort- versus Time-Effects



The figure portrays the 36 age dummies from running linear regressions abstracting from cohort effects and time effects, respectively. Each graph is normalized to match the unconditional average for the corresponding cross-sectional moment over the period 1980-1997.

Table 1

	1968-1972	1973-1977	1978-1982	1983-1987	1988-1992	1993-1996	Correlation
Variance of log wages							
Within cohort $\Delta x^{c}_{t,t+1}$	-0.2727	-0.4166	0.9909	1.2466	0.8135	0.5877	
S.E.	*0.113*	*0.127*	*0.111*	*0.180*	*0.217*	*0.246*	0.647
Within age group $\Delta x^{a}_{t,t+1}$	0.5426	-0.4642	0.9262	0.9266	0.1813	-0.0226	
S.E.	*0.190*	*0.225*	*0.210*	*0.319*	*0.274*	*0.395*	
Between age group $\Delta x^{t}_{a,a+1}$	0.5485	0.5861	0.3925	0.6137	0.8134	0.8029	
S.E.	*0.113*	*0.127*	*0.111*	*0.180*	*0.217*	*0.246*	
Variance of log earnings							
Within cohort $\Delta x^{c}_{t,t+1}$	-0.3994	-0.2912	1.7410	1.6375	0.6589	0.6158	
S.E.	*0.157*	*0.169*	*0.174*	*0.237*	*0.274*	*0.331*	0.691
Within age group $\Delta x^{a}_{t,t+1}$	0.5948	-0.2434	1.4004	1.0915	-0.1149	-0.0696	
S.E.	*0.326*	*0.294*	*0.289*	*0.414*	*0.421*	*0.518*	
Between age group $\Delta x^{t}_{a,a+1}$	0.4787	0.4946	0.3338	0.4990	0.8192	0.8261	
S.E.	*0.157*	*0.169*	*0.174*	*0.237*	*0.274*	*0.331*	
Variance of log hours							
Within cohort $\Delta x^{c}_{t,t+1}$	-0.1354	-0.2044	0.4152	-0.0181	0.3669	-0.2475	
S.E.	*0.059*	*0.057*	*0.059*	*0.070*	*0.072*	*0.081*	0.804
Within age group $\Delta x^{a}_{t,t+1}$	0.1298	-0.1594	0.2361	-0.0846	0.2358	-0.4560	
S.E.	*0.136*	*0.113*	*0.102*	*0.125*	*0.124*	*0.159*	
Between age group $\Delta x^{t}_{a,a+1}$	-0.0954	-0.0588	-0.0391	0.0050	-0.0049	-0.0226	
S.E.	*0.059*	*0.057*	*0.059*	*0.070*	*0.072*	*0.081*	
Covariance between log wages and log hours							
Within cohort $\Delta x^{c}_{t,t+1}$	0.0030	0.1647	0.1358	0.1707	-0.1602	0.1681	
S.E.	*0.047*	*0.056*	*0.055*	*0.073*	*0.084*	*0.086*	0.969
Within age group $\Delta x^{a}_{t,t+1}$	-0.0385	0.1900	0.0904	0.1029	-0.2163	0.2041	
S.E.	*0.099*	*0.109*	*0.099*	*0.130*	*0.155*	*0.175*	
Between age group $\Delta x^{t}_{a,a+1}$	0.0128	-0.0142	-0.0096	-0.0703	-0.0178	0.0230	
S.E.	*0.047*	*0.056*	*0.055*	*0.073*	*0.084*	*0.086*	
Variance of log consumption							
Within cohort $\Delta x^{c}_{t,t+1}$			-0.9831	-0.7194	-0.4690	0.1987	
S.E.			*0.981*	*0.458*	*0.394*	*0.451*	-0.431
Within age group $\Delta x^{a}_{t,t+1}$			1.1057	-0.3177	-0.2132	0.0700	
S.E.			*0.337*	*0.225*	*0.185*	*0.187*	
Between age group $\Delta x^{t}_{a,a+1}$			-0.1285	-0.0039	0.0383	0.0135	
S.E.			*0.454*	*0.239*	*0.227*	*0.240*	
Variance of log non-durable consumption							
Within cohort $\Delta x^{c}_{t,t+1}$			-1.9826	-0.5245	-0.1804	0.4470	
S.E.			*0.972*	*0.729*	*0.373*	*0.591*	-0.654
Within age group $\Delta x^{a}_{t,t+1}$			0.1678	0.1790	-0.0450	0.0285	
S.E.			*0.252*	*0.212*	*0.176*	*0.279*	
Between age group $\Delta x^{t}_{a,a+1}$			-0.2117	-0.0129	0.0317	0.0348	
S.E.			*0.496*	*0.292*	*0.221*	*0.244*	

Note: Average annual change within cohort, within age group and between successive age groups by period All changes are multiplied by 100. Standard Errors are reported below. The column labelled "Correlation" displays the correlation coefficient between the within-age and the within-cohort changes. The data for consumption in the column labelled "1978-1982" refer to the period 1980-1982.