Jérôme Adda

James Banks

Hans-Martin von Gaudecker

The Impact of Income Shocks on Health: Evidence from Cohort Data

ABSTRACT

The Impact of Income Shocks on Health: Evidence from Cohort Data[*]

We study the effect of permanent income innovations on health for a prime-aged population. Using information on more than half a million individuals sampled over a twenty-five year period in three different cross-sectional surveys we aggregate data by date-of-birth cohort to construct a 'synthetic cohort' dataset with details of income, expenditure, socio-demographic factors, health outcomes and selected risk factors. We then exploit structural and arguably exogenous changes in cohort incomes over the eighties and nineties to uncover causal effects of permanent income shocks on health. We find that such income innovations have little effects on a wide range of health measures, but do lead to increases in mortality and risky health behaviour.

Corresponding author:

Jérôme Adda
University College London
Department of Economics
Gower Street
London WC1E 6BT
United Kingdom
E-mail: j.adda@ucl.ac.uk

1 Introduction

One of the most striking and frequently studied correlations observed in individual level data, is that between measures of income or socio-economic status and measures of health. Data from all over the world, collected in many different years, consistently show that those with greater levels of economic resources have better health. Consequently, an extensive and continually evolving literature has concerned itself with the precise nature of the relationship between income and health and in particular the direction and magnitude of causal links between the two.

Interpretations of the origin of the health-income gradient require researchers to deal with a classic simultaneity issue. Three broad channels of causality can be imagined. Epidemiologists stress the importance of causality from income to health (Marmot 1999). Economists have often tried to quantify the reverse direction going from poor health to impaired current and future earnings capacity (Smith 2004). A third explanation is that there are one or more underlying factors determining both income and health. Among many potential candidates are genetics, ability/intelligence and other background or early life factors such as parental income (Case, Lubotsky, and Paxson (2002), Dehejia and Lleras Muney (2004) or van den Berg, Lindeboom, and Portrait (2006)) or education (Lleras Muney 2005).

In this paper, we focus on the effect of income shocks on health over the life-cycle. First, we specify an individual dynamic model of both income and health which allows us to decompose health and income shocks into transitory and permanent ones. We also allow for a non-linear relationship between income and health at the individual level. Second, we aggregate and estimate this model, using synthetic cohorts followed for up to twenty-five years, which covers a substantial part of the life-cycle. Third, we estimate the effect of *permanent* shocks to income on a number of health outcomes including subjective and objective measures, as well as health behaviour and mortality.

We use a synthetic cohort methodology for two reasons. First, by using it we aggregate out individual level variation, allowing us to focus on exogenous changes in income at the cohort level. In many countries, including the UK, there have been important changes in the income structure over time (Buchinsky (1994), Gottschalk and Smeeding (1997) or Gosling, Machin, and Meghir (2000)). The causes appear to range from changes in the return to education and experience due to skill-biased technology changes, declines in unionization and increased competition. Importantly, however, these changes are not thought to be caused by changes in health and we will use this as an exogeneity assumption to identify a causal effect of income on health. In an individual level model it would be much harder to disentangle the effects in question. [1] In the spirit of robustness analysis we also investigate the results we would obtain if the exogeneity assumption was violated. The calculations show that varying the strength of the effect of cohort health shocks on cohort incomes does not seriously affect our conclusions.

A second advantage of cohort data is that we can exploit a wealth of data with detailed information on both income and many health outcomes because it allows the combination of various datasets. This is particularly important in our analysis because data on health and socio-economic variables typically have not been recorded

---

[1] Some attempts have been made using exogenous changes in income such as lottery gains (Lindahl 2005), or "unexpected" inheritances (Meer, Miller, and Rosen 2003) as instruments although these instruments apply only to subgroups of the population which may not necessarily be representative.

jointly. Longitudinal data sets have either good information on income dynamics (e.g. the PSID in the US, or the BHPS in the UK) but no or very limited information on health, or detailed health information but very limited information on income or other socioeconomic variables. This has changed recently, but only over the last decade, and mostly for data covering the elderly and near-elderly population (e.g. the Health and Retirement Survey). Hence there are no long-T panels currently available that facilitate the study of the joint dynamics of income and health for the prime-aged population.

In our analysis we use techniques recently developed in the microeconometric literature for modelling income processes and income uncertainty and apply them to the question of identifying the effect of income shocks on subsequent health outcomes. More precisely, we use the methods of Meghir and Pistaferri (2004), who in turn build on the methods of Abowd and Card (1989), Gottschalk and Moffitt (1994), Banks, Blundell, and Brugiavini (2001), and Blundell, Pistaferri, and Preston (2005) to identify the covariance of permanent income shocks with changes in a series of health indicators. In this way we are able to extend the Granger-causality type analysis, previously carried out for US individuals aged over 70 (Adams, Hurd, McFadden, Merrill, and Ribeiro 2003) and aged over 50 (Smith 2004), and apply it to a working age population in England. These previous studies found no effect of income shocks on subsequent health changes, once one conditions on initial health and income. Differences in health care institutions, between the UK and the US, and particularly the relative lack of private provision in the UK, would hardly be expected to reverse such a result. The use of a different age group, however, could lead to different findings than those of the above studies, regardless of the country of study.

We show that different cohorts were affected by sizable permanent shocks to income over that period, especially those with low education. We find, as in Ruhm (2000), that these shocks have an impact on mortality. However, our results show that these shocks are not transmitted to any of our other health measures, whether we consider subjective ones (self-assessed health, longstanding illness) or objective ones (high blood pressure, cardiovascular diseases, or respiratory diseases). In contrast, we show that individuals change some of their behaviour such as total expenditure as well as expenditures on tobacco and alcohol. Our interpretation of these three findings is that risk behaviours do not seem to transmit directly into mortality or morbidity for the prime-aged population. Procyclical mortality is probably rather driven by work-related accidents and similar mechanisms, but our results are not more than suggestive on this point.

Our analysis is structured as follows. Section 2 describes our empirical strategy by setting out the structural framework for the dynamics of income and health processes, outlining our key identification assumptions and documenting our estimation methodology. Section 3 describes the various datasets that we draw upon in order to carry out our analysis, and provides some simple summary statistics. Section 4 presents our empirical results, first with reference to the effect of income shocks on mortality and then with reference to a host of health indicators and behaviour. The last part of that section (4.6) contains a detailed comparison of our approach to related ones found in the literature. Finally Section 5 concludes.

2 Empirical Strategy

In this section we describe our approach to modelling income and health as stochastic processes that evolve over the life cycle. We then turn to our strategy to estimate the relevant parameters. In either of the two areas, there exists substantial previous work that we can build on. Starting with MaCurdy (1982), a large literature has emerged that models life-cycle income or earnings as stochastic processes. In terms of methodology, Abowd and Card (1989) and especially Meghir and Pistaferri (2004) are close to our approach. All these articles deal with individual panel data, applications to consumption with synthetic cohorts include Blundell and Preston (1998) and Banks, Blundell, and Brugiavini (2001).

Regarding the literature on the evolution of health, Smith (2004) gives an extensive overview of different specifications. Few authors have used pseudo panel data in connection with health outcomes, a notable exception being a series of papers by Deaton and Paxson (1998b, 2001, 2004). We provide a comparison of our approach and results to theirs (and those of Ruhm (2000)) in Section 4.6.

We adopt the following notation throughout the analysis. ε and u denote transitory shocks, ζ and v permanent shocks. Parameters corresponding to them carry a superscripts T and P, respectively. x^Y income-related variables and x^H health-related variables. i/N denote an individual and the number of individuals respectively, I indexes parameters corresponding to individual-level variables. Cohorts carry an index c and parameters corresponding to cohort-level variables a superscript C. A time period (one year) is subscripted by t and age by a. L denotes the lag-operator.

2.1 Stochastic Processes for Individual Income and Health

We assume the following nonstationary process for the evolution of log household income over the life-cycle:

$$(1) \qquad Y_{it} = Y_{i,t-1} + v_{it}^Y + (1 - L)u_{it}^Y$$

The permanent and transitory shocks are decomposed into health and income shocks either at individual or cohort level:

$$(2) \qquad v_{it}^Y = \zeta_{ct}^Y + \zeta_{it}^Y + \sum_{j=0}^{p+1} \xi_j^{C,P} \zeta_{c,t-j}^H + \sum_{j=0}^{p+1} \xi_j^{I,P} \zeta_{i,t-j}^H$$

$$(3) \qquad u_{it}^Y = \sum_{j=0}^{p} \theta_j^C \varepsilon_{c,t-j}^Y + \sum_{j=0}^{p} \theta_j^I \varepsilon_{i,t-j}^Y + \sum_{j=0}^{p} \xi_j^{C,T} \varepsilon_{c,t-j}^H + \sum_{j=0}^{p} \xi_j^{I,T} \varepsilon_{i,t-j}^H$$

$$(4)$$

with $\theta_0 = 1$ and $|\theta_j| < 1$ for $j > 0$. p is the order of the moving average process governing the persistence of transitory shocks and the number of periods in which permanent health shocks may impact income.[2] We deliberately distinguish between cohort-wide shocks to income denoted by x_{ct}^Y and individual-level shocks represented by x_{it}^Y. Examples of the former include changes in the returns to education, the latter may be promotions, bonuses or

---

[2]In principle, p could be a different number in each of the terms where it appears. However, we shall see that this point will not be relevant for our analysis and we choose to not further complicate notation.

layoffs. Cohort and individual health shocks are denoted by x_{ct}^H and x_{it}^H respectively, with the impact on income being determined by the coefficients ξ. All our estimation equations further include demographic terms which remain suppressed in this section for ease of notation.

This set up is effectively a version of the stochastic process in Blundell and Preston (1998). They also show how equation (1) can be derived from a decomposition of income into a permanent and a transitory component. There are two main differences in our approach. First, we augment the specification by a moving average component in the transitory errors (see Meghir and Pistaferri (2004) or Banks, Blundell, and Brugiavini (2001)). With respect to the identification problem concerning cohort, age, and time effects, we assume the latter to be zero although a time trend is captured by our empirical methodology. Results reported in Banks, Blundell, and Brugiavini (2001) show that shocks to cohort income dominate any aggregate effects so that we do not view this assumption to be problematic. The second difference to the above-mentioned papers is that we take individual health shocks explicitly into account. Similar to the income shocks, permanent shocks to health are represented by the ζ^H's and transitory shocks are denoted by the ε^H's. After describing the stochastic process for health, we will go into more detail on the nature and role of these shocks.

For the purposes of this section, we treat health as a unidimensional stock variable. To model its evolution over the life course, we use the process employed in Deaton and Paxson (1998a):

$$(5) \qquad H_{it} \quad = \quad H_{i,t-1} + v_{it}^H + (1-L)u_{it}^H$$

Again, we assume that the error terms can be broken down additively into several components:

$$(6) \qquad v_{it}^H \quad = \quad \zeta_{ct}^H \quad + \quad \zeta_{it}^H + \sum_{j=0}^{q+1} \gamma_j^{C,P} \zeta_{c,t-j}^Y + \sum_{j=0}^{q+1} \gamma_j^{I,P} \zeta_{i,t-j}^Y$$

$$(7) \qquad u_{it}^H \quad = \quad \sum_{j=0}^{q} \phi_j^C \varepsilon_{c,t-j}^H + \sum_{j=0}^{q} \phi_j^I \varepsilon_{i,t-j}^H + \sum_{j=0}^{q} \gamma_j^{C,T} \varepsilon_{c,t-j}^Y + \sum_{j=0}^{q} \gamma_j^{I,T} \varepsilon_{i,t-j}^Y$$

for some order q of the moving-average process. Examples of individual transitory health shocks ε_{it}^H are broken bones or influenza, which in most cases do not have important long-term effects. On the other hand ζ_{it}^H can be thought of as a permanent (incurable) shock like many cancers or AIDS. Transitory cohort level shocks ε_{ct}^H could be highly infectious diseases whereas the most important example for permanent innovations that affect entire cohorts ζ_{ct}^H would be medical progress.

The income shocks described previously are allowed to have a direct effect on health that is governed by the parameters γ. We impose the restriction that permanent shocks to income influence health permanently and transitory shocks to income affect health only in a transitory fashion. This is analogous to what we imposed on the income process in (2) and (3). [3] In this specification, transitory income shocks may influence health in time periods t until $t+q$. Period-t permanent income shocks are assumed to impact health not beyond period $t+q+1$. In Section 4.3 we report results for $q = 0, 1, 2$. [4]

---

[3]Following the suggestion of a referee, we derived the implications of a more general model where transitory shocks to income could have permanent effects on health as well. We could not reject the more parsimonious model in our data and thus do not present the more involved specification.

[4]The same comments apply to q that we mentioned before with respect to p. There could be different values of q for each of the

Note that we do not give (5) to (7) an interpretation that is structural in the sense that all relevant decision parameters are incorporated in a model of health production. We rather specify a model that is rich enough to incorporate the most important conceivable causation channels and arrive at a causal interpretation through appropriate exclusion restrictions. This is different from the approach taken in the consumption literature. For example, Blundell, Pistaferri, and Preston (2005) derive a rather similar condition on the evolution of consumption from intertemporal utility maximisation. A natural point of departure in our case would have been a version of the classic model of health capital by Grossman (1972). We do not employ such a model because research in epidemiology is richly suggestive of mechanisms linking income changes to health outcomes over which individuals have only limited control (for example through byproducts of physiological stress markers). Put differently, we think that too little is known about the causation channels running from income to health outcomes as to isolate and parameterise the relevant ones a priori. As a consequence, we stick to the general statistical modelling approach.

In our analysis we also focus on health behaviour. For these models we reinterpret H_{it} in equation (5) as a measure of expenditures or quantity of consumption goods, such as fruits and vegetables, cigarettes and alcohol. In these cases, the economic model of informed decision-making has received a widespread application, as an example see the review of Chaloupka and Warner (2000) with respect to smoking behaviour. Nevertheless we stick to our approach for coherency reasons. The comments made in the last paragraph still apply: It is quite general and we do not see any interesting causation channels that we miss. There is a qualifier to this. It is well known that in our type of data it is not possible to control for age, cohort, and time effects simultaneously. In our particular case, the exclusion of time effects over and above a trend seems more justified in the case of health outcomes than consumption goods. For example, a jump in the relative price of cigarettes due to a tax increase may manifest itself as a shock that goes into the same direction for all cohorts. The assumption of iid-shocks would be violated. While we cannot make our inference strategy robust to this type of shocks, our point estimates remain valid. Having said this, we note that relative prices of cigarettes in the UK experienced a rather smooth upward trend over the period under scrutiny (with the exception of two slight declines in the late seventies and late eighties). All our estimates contain a full set of cohort dummies and a flexible age trend which adequately capture such trends and since we are not interested in interpreting either the age or cohort coefficients this seems sufficient for our purposes.

2.2 Aggregation

We need some assumptions on the idiosyncratic shocks before we proceed to the cohort level. In particular, for the purposes of exposition we impose the restriction that their mean depends on age only. This is equivalent to specifying an age trend and zero-mean errors. It captures the age-dependence of income and health. Other time-constant covariates such as gender or education are straightforward to incorporate in the same way and we include them in our empirical modelling thus allowing a considerable heterogeneity in the key income and health processes. However we do not discuss them further in this section for ease of notation.

In addition to this, we assume all idiosyncratic shocks to be independently and identically distributed in any

---

terms where it appears, in particular there could be much more persistence (i.e. $q > 2$) in the individual-level shocks. These will not affect our subsequent analysis.

given cross-section of individuals. This is an identifying assumption because correlated shocks will be captured by the cohort-level stochastic terms. Note that shocks may be correlated within individuals over time, which captures slowly progressing health conditions such as diabetes or obesity. For the first two moments, we have:

$$E(\varepsilon_{it}^Y) = m_a^{Y,T}, \qquad V(\varepsilon_{it}^Y) = \sigma_{\varepsilon^{Y,I}}^2, \qquad cov(\varepsilon_{it}^Y, \varepsilon_{jt}^Y) = 0 \quad \forall \quad i \neq j$$

$$E(\zeta_{it}^Y) = m_a^{Y,P}, \qquad V(\zeta_{it}^Y) = \sigma_{\zeta^{Y,I}}^2, \qquad cov(\zeta_{it}^Y, \zeta_{jt}^Y) = 0 \quad \forall \quad i \neq j$$

$$E(\varepsilon_{it}^H) = m_a^{H,T}, \qquad V(\varepsilon_{it}^H) = \sigma_{\varepsilon^{H,I}}^2, \qquad cov(\varepsilon_{it}^H, \varepsilon_{jt}^H) = 0 \quad \forall \quad i \neq j$$

$$E(\zeta_{it}^H) = m_a^{H,P}, \qquad V(\zeta_{it}^H) = \sigma_{\zeta^{H,I}}^2, \qquad cov(\zeta_{it}^H, \zeta_{jt}^H) = 0 \quad \forall \quad i \neq j$$

where m^T stands for constants associated with transitory shocks, and m^P for those associated with permanent ones. The I-superscripts on the variance terms indicate that they refer to the respective individual-level parameters.

Aggregating (1) and (5) up for individuals within each cohort, c, leads to:

$$
\begin{aligned}
(8) \qquad Y_{ct} =\ & Y_{c,t-1} + m_a^{Y,P} + \sum_{j=0}^{p+1} \xi_j^{I,P} m_{a-j}^{H,P} + (1-L)\left(\sum_{j=0}^{p} \theta_j^I m_{a-j}^{Y,T} + \sum_{j=0}^{p} \xi_j^{I,T} m_{a-j}^{H,T} \right) \\
& + \zeta_{ct}^Y + \sum_{j=0}^{p+1} \xi_j^{C,P} \zeta_{c,t-j}^H + (1-L)\left(\sum_{j=0}^{p} \theta_j^C \varepsilon_{c,t-j}^Y + \sum_{j=0}^{p} \xi_j^{C,T} \varepsilon_{c,t-j}^H + \nu_{ct}^Y \right)
\end{aligned}
$$

$$
\begin{aligned}
(9) \qquad H_{ct} =\ & H_{c,t-1} + m_a^{H,P} + \sum_{j=0}^{q+1} \gamma_j^{I,P} m_{a-j}^{Y,P} + (1-L)\left(\sum_{j=0}^{q} \phi_j^I m_{a-j}^{H,T} + \sum_{j=0}^{q} \gamma_j^{I,T} m_{a-j}^{Y,T} \right) \\
& + \zeta_{ct}^H + \sum_{j=0}^{q+1} \gamma_j^{C,P} \zeta_{c,t-j}^Y + (1-L)\left(\sum_{j=0}^{q} \phi_j^C \varepsilon_{c,t-j}^H + \sum_{j=0}^{q} \gamma_j^{C,T} \varepsilon_{c,t-j}^Y + \nu_{ct}^H \right)
\end{aligned}
$$

where ν_{ct} represents cell variation as induced by summing up the stochastic terms. It can also serve to incorporate classical measurement error (which was omitted from equations (1) and (5) purely for ease of notation). Hence, both terms will be iid across time (but not necessarily across cohorts) with zero mean:

$$E(\nu_{ct}^Y) = 0, \qquad V(\nu_{ct}^Y) = \sigma_{\nu^Y}^2, \qquad cov(\nu_{ct}^Y, \nu_{ct'}^Y) = 0 \quad \forall \quad t \neq t'$$

$$E(\nu_{ct}^H) = 0, \qquad V(\nu_{ct}^H) = \sigma_{\nu^H}^2, \qquad cov(\nu_{ct}^H, \nu_{ct'}^H) = 0 \quad \forall \quad t \neq t'$$

The transitory and permanent shocks to income and health at the cohort level may be correlated across cohorts within periods. We only impose the standard assumption of no correlation over time and an innocuous zero-mean condition:

$$E(\varepsilon_{ct}^Y) = 0, \qquad V(\varepsilon_{ct}^Y) = \sigma_{\varepsilon^{Y,C}}^2, \qquad cov(\varepsilon_{ct}^Y, \varepsilon_{ct'}^Y) = 0 \quad \forall \quad t \neq t'$$

$$E(\varepsilon_{ct}^H) = 0, \qquad V(\varepsilon_{ct}^H) = \sigma_{\varepsilon^{H,C}}^2, \qquad cov(\varepsilon_{ct}^H, \varepsilon_{ct'}^H) = 0 \quad \forall \quad t \neq t'$$

$$E(\zeta_{ct}^Y) = 0, \qquad V(\zeta_{ct}^Y) = \sigma_{\zeta^{Y,C}}^2, \qquad cov(\zeta_{ct}^Y, \zeta_{ct'}^Y) = 0 \quad \forall \quad t \neq t'$$

$$E(\zeta_{ct}^H) = 0, \qquad V(\zeta_{ct}^H) = \sigma_{\zeta^{H,C}}^2, \qquad cov(\zeta_{ct}^H, \zeta_{ct'}^H) = 0 \quad \forall \quad t \neq t'$$

ζ_{ct}^Y and ζ_{ct}^H need to be contemporaneously uncorrelated which does not seem to be very restrictive with respect to health. Remember that the most important example for cohort-level health shocks is probably medical progress

which usually does not affect all age groups in the same way. With respect to behaviour we refer to the discussion in the concluding paragraph of the previous subsection.

In order to simplify (8) and (9), note that all m-terms including their coefficients are not separately identified. Thus they can be included in a single intercept term that depends on age only:

$$m_a^Y \;=\; m_a^{Y,P} + \sum_{j=0}^{p+1} \xi_j^{I,P} m_{a-j}^{H,P} + (1-L)\left(\sum_{j=0}^{p} \theta_j^I m_{a-j}^{Y,T} + \sum_{j=0}^{p} \xi_j^{I,T} m_{a-j}^{H,T}\right)$$

$$m_a^H \;=\; m_a^{H,P} + \sum_{j=0}^{q+1} \gamma_j^{I,P} m_{a-j}^{Y,P} + (1-L)\left(\sum_{j=0}^{q} \phi_j^I m_{a-j}^{H,T} + \sum_{j=0}^{q} \gamma_j^{I,T} m_{a-j}^{Y,T}\right)$$

Hence we can rewrite (8) and (9) as:

$$(10) \qquad Y_{ct} \;=\; Y_{c,t-1} + m_a^Y + \zeta_{ct}^Y + \sum_{j=0}^{p+1} \xi_j^{C,P} \zeta_{c,t-j}^H + (1-L)\left(\sum_{j=0}^{p} \theta_j^C \varepsilon_{c,t-j}^Y + \sum_{j=0}^{p} \xi_j^{C,T} \varepsilon_{c,t-j}^H + \nu_{ct}^Y\right)$$

$$(11) \qquad H_{ct} \;=\; H_{c,t-1} + m_a^H + \zeta_{ct}^H + \sum_{j=0}^{q+1} \gamma_j^{C,P} \zeta_{c,t-j}^Y + (1-L)\left(\sum_{j=0}^{q} \phi_j^C \varepsilon_{c,t-j}^H + \sum_{j=0}^{q} \gamma_j^{C,T} \varepsilon_{c,t-j}^Y + \nu_{ct}^H\right)$$

This is our basic system of equations which we use to identify permanent shocks to income and to investigate the relation between them and various health variables.

2.3   Identification

Identification of the parameters in (10) and (11) closely traces the analysis in Meghir and Pistaferri (2004) and Blundell, Pistaferri, and Preston (2005) once we make the system triangular in terms of permanent shocks. Our key identifying assumption in this respect is that $\xi^{C,P} = 0$, i.e. that permanent cohort-level shocks to health do not affect income (or are not present at all). Before presenting the steps that lead to identification of the parameters of interest, we provide a brief discussion of the identification issues. In addition to the discussion below, a more empirical approach will also be taken when we estimate the model and we will attempt to examine potential departures from exogeneity. Whilst we cannot identify $\xi^{C,P}$ we can consider how our estimates would be effected were the true value to be other than zero so section 4.5 provides estimates for the key parameters of our model under alternative values for $\xi^{C,P}$.

It is worth beginning this discussion, however, by noting that the model allows for any type of health to income causality at the individual level. Here, both permanent and transitory shocks to health may affect income in an arbitrary fashion. This covers most typically mentioned channels ranging from work disability to transitorily impaired earnings capacities such as broken bones. The aggregation to the cohort level, combined with the time-series structure we impose, permits us to abstract from these issues. We also allow for transitory health innovations at the cohort level to affect current income which takes care of infectious diseases such as influenzas that might impact on different age-education cohorts differentially.

With respect to permanent changes in cohort health, the most salient example is medical progress that affects each cohort in a specific way. The introduction of by-pass surgery could be seen as a permanent health shock for

older cohorts. Our identifying assumption claims that these changes are small enough as to not affect the incomes of entire cohorts. Rather, permanent income fluctuations are driven by the factors described above such as differential returns to education. This exclusion restriction appears to be rather plausible to us, especially at younger ages. Differences in early retirement seem to be more problematic in this respect. This is the reason why we restrict ourselves to people until sixty years of age. We did robustness checks with even younger cohorts only and results were not sensitive to the cutoff age.

Having triangularised the system, we use very similar moment conditions as Meghir and Pistaferri (2004) to identify the permanent shocks. In contrast to them, however, we do not consider bounding the transitory ones.[5] We start by defining g_{ct}^Y (g_{ct}^H) as the part in the growth rate of income (health) that is not explained by age effects.

$$
g_{ct}^Y = \zeta_{ct}^Y + \sum_{j=0}^{p+1} \xi_j^{C,P} \zeta_{c,t-j}^H + (1-L) \left(\sum_{j=0}^{p} \theta_j^C \varepsilon_{c,t-j}^Y + \sum_{j=0}^{p} \xi_j^{C,T} \varepsilon_{c,t-j}^H + \nu_{ct}^Y \right)
$$

$$
g_{ct}^H = \zeta_{ct}^H + \sum_{j=0}^{q+1} \gamma_j^{C,P} \zeta_{c,t-j}^Y + (1-L) \left(\sum_{j=0}^{q} \phi_j^C \varepsilon_{c,t-j}^H + \sum_{j=0}^{q} \gamma_j^{C,T} \varepsilon_{c,t-j}^Y + \nu_{ct}^H \right)
$$

These equations relate the annual changes in income and health for each cohort to the parameters of our model. We rely on second order moments to identify the parameters of interest. Our data allow us to characterize the variance of the changes in income or health, as well as their covariances at different lags. We can then relate these moments from the data to the theoretical moments. For instance, the variance of the permanent shock to income at the cohort level can be expressed as a function of the covariances of the changes in income at various lags and the variance of the changes in income. This strategy can be formalised in the following way:

$$
(12) \qquad E\left[g_{ct}^Y \left(\sum_{j=-(p+1)}^{p+1} g_{c,t+j}^Y \right) \right] - E\left[g_{ct}^Y \right] E\left[\sum_{j=-(p+1)}^{p+1} g_{c,t+j}^Y \right] = \sigma_{\zeta^{Y,C}}^2
$$

$$
(13) \qquad E\left[g_{ct}^H \left(\sum_{j=-(q+1)}^{q+1} g_{c,t+j}^H \right) \right] - E\left[g_{ct}^H \right] E\left[\sum_{j=-(q+1)}^{q+1} g_{c,t+j}^H \right] = \left(\sum_{j=0}^{q+1} \gamma_j^{C,P} \right)^2 \sigma_{\zeta^{Y,C}}^2 + \sigma_{\zeta^{H,C}}^2
$$

$$
(14) \qquad E\left[g_{ct}^H \left(\sum_{j=-(q+1)}^{p+1} g_{c,t+j}^Y \right) \right] - E\left[g_{ct}^H \right] E\left[\sum_{j=-(q+1)}^{p+1} g_{c,t+j}^Y \right] = \sum_{j=0}^{q+1} \gamma_j^{C,P} \sigma_{\zeta^{Y,C}}^2
$$

We note that the terms in the sum $\sum_{j=0}^{q+1} \gamma_j^{C,P}$ are not separately identified. Hence we can say something about the aggregate impact over time of permanent income shocks on health, but we cannot break it down further by year. Treating this sum as one parameter $\gamma^{C,P}$, we have three parameters in three equations that can be estimated via GMM.

---

[5]Meghir and Pistaferri (2004) find a MA(1)-process for the transitory shock to income which gives them a system of three parameters (the variance of the transitory shock to income, the MA(1)-parameter, and the variance of measurement error) in two equations based on the first and second order autocovariances of income. With a MA(0)-process, the two variances are completely confounded. In trying to exploit the cross-covariance restrictions further (5 parameters in 3 equations for the MA(0) case in the health process), we passed the limits of our synthetic cohort data.

2.4 Estimation

Our estimation strategy consists of three steps. We first sketch them very briefly and then describe each one in more detail.

1. Regress the health and income variables on a suitable set of regressors capturing cohort and age effects.

2. Use the autcovariances of the first-differenced residuals to determine p and compute the terms inside the expectations of (12) to (14).

3. Calculate $\sigma^2_{\zeta^{Y,C}}$ and $\sum_{j=0}^{q+1} \gamma_j^{C,P}$ for $q = 0, 1, 2$ via GMM, obtain confidence intervals by bootstrapping.

With respect to the first step of the regression, we choose a full set of cohort dummies and a quadratic in age interacted with sex and education. Since our cohorts will be based upon the latter two variables in addition to date-of-birth intervals (see Section 3), this allows for very flexible profiles across the sex and education groups as well as an additive effect for year of birth. For all variables with the exception of mortality, results were invariant to higher order polynomials in age and the respective interaction terms. This modelling strategy allows us to control for cohort specific effects determined by macro-economic conditions at birth which may impact individuals as shown by van den Berg, Lindeboom, and Portrait (2006) or Dehejia and Lleras Muney (2004). They also capture the downward trends in overall mortality and morbidity over time. We checked for robustness towards the additivity restriction on cohort effects by permitting the income and health profiles to be cohort-specific. Point estimates did not change much but standard errors rose a lot as expected for an overfitted model.

First-differencing the residuals from these regressions immediately gives us the growth rates g_{ct}^Y and g_{ct}^H. In principle we could directly run a regression in first differences. This is inhibited by the fact that our time series dimension is not long enough to make the attenuation bias of OLS in this type of regressions negligible.

We then use GMM techniques to estimate the parameters of interest. All variance terms are allowed to vary across sex and education groups. After some experimentation, we restrict the $\gamma^{C,P}$-parameters to be the same. Due to the small-sample bias that is often observed in optimal minimum distance estimation of covariance structures, we use equally weighted minimum distance (Altonji and Segal 1996). Since it is very hard to keep track of the pre-estimation error analytically, we obtain confidence intervals by using the bootstrap (500 replications). As a consequence of the nonstationarity of the model, we sample entire time series of each cohort. Inference hence rests upon the cross-sectional dimension of the estimation problem.

3 Data

In order to study the dynamic effects of income shocks on health, we construct a synthetic cohort dataset, based on successive years of microdata from several cross-sectional surveys. This type of data has proven to be very useful in the consumption literature because it facilitates the use of large datasets in long time series. While restricting the range of admissible models to additive ones, it permits much more generality than aggregate data. Moreover the

approach allows the combination of various datasets. This is particularly important in our analysis because data on health and socio-economic variables typically have not been recorded jointly. This has changed over the last decade, but mostly for data covering the elderly and near-elderly population. Hence there are no panels currently available that allow to study dynamics for the prime-aged population.

We use English data from three repeated cross-section surveys: The Family Expenditure Survey (FES), the General Household Survey (GHS), and the Health Survey for England (HSE). We will describe these datasets in turn after some general remarks. All of them contain basic demographic information such as individual education level as measured by the age left full time education, marital status, number of persons in the household, number of children and number of children under five years of age in the household, and labour market status.

We construct cohorts based on 3-year date-of-birth intervals, sex, and whether individuals attended school beyond the compulsory schooling age or not. We consider the working-age population of 30 to 60 years of age. The upper bound is set because of cohort differences in early retirement. In our data, early retirement (to the extent it is associated with a fall in income) would show up as a negative income shock, but there may be an independent influence of retirement on health that we cannot control for (see for example Snyder and Evans (2006)). We chose the lower bound because we model household income and many people live with their parents until their mid-twenties. One may well expect that risk sharing in these households is not perfect and that there is a large deal of financial autonomy between generations. If this is the case, then cohort differences in the changes of household income at these ages would not necessarily reflect the type of income shocks that we are interested in.

Table 1 lists the data availability by study and year and Table 2 provides summary statistics of the cell sizes within each date-of-birth, education and gender cohort in each of the surveys. As documented in Table 2, the cell sizes are reasonably large and throughout the analysis we do not explicitly consider aggregation error in calculating the standard errors. This is common practice in the consumption literature (e.g. Browning, Deaton, and Irish (1985)). The HSE displays very small sample sizes in its first two years of existence, as can be inferred from the last two rows of Table 2. We check robustness to exclusion of these two years in all specifications. All data are used on an annual basis from 1978 to 2003, although the HSE and some variables in the GHS do not cover the full period. The rest of this section describes the relevant variables used from each of the datasets in turn. The main variables are on household income and consumption as well as individual-level risk behaviours and health outcomes. Descriptive Statistics for all variables can be found in Table 3 which also indicates which variables are used at an individual or at the household level.

In addition to the three repeated cross sections, we also use all-population mortality from the Human Mortality Database (http://www.mortality.org) which is available for the period 1978-1998. For this dataset, no covariates beyond sex and age are available. Hence we cannot base cohorts on education groups when it comes to mortality. The Longitudinal Study would have allowed for such stratification, but its education measure is based on the highest degree obtained and proved to be not comparable to the years of schooling measure.

3.1 The Family Expenditure Survey (FES)

The FES contains detailed information on household income and consumption. Together with the Consumer Expenditure Survey in the US, it has emerged as one of the workhorses in the consumption literature (see for example Blundell, Pashardes, and Weber (1993)). Our data cover the period from 1978 to 2003 and our selected sample comprises 148,517 individuals. Unless indicated otherwise, we use expenditure and income in logarithms and weighted by the OECD-equivalence scale

$$\text{EQUIV} = 1 + 0.6 \cdot (\#\text{ADULTS} - 1) + 0.4 \cdot \#\text{CHILDREN}.$$

We conduct sensitivity analyses with unequivalised values and demographic variables on the right-hand side, results were qualitatively similar and are not reported. All variables are deflated with a general price index to obtain January 2000 prices. The evolution of mean income and mean expenditure on non- and semi-durable items is shown in Figures 1 and 2 for every other cohort.

The familiar hump-shape over the life-cycle becomes apparent for both variables and it seems to be more pronounced for expenditure as well as for the group with low education. Values for men are slightly higher than those for women, reflecting the higher income of male single households with respect to female single ones. As expected, we find pronounced education differences. Compared to these, cohort effects appear to be moderate with younger cohorts generating more income and consumption than older ones. These effects seem to be stronger for the higher educated, which is consistent with the findings of Gosling, Machin, and Meghir (2000) for male wages. In addition to total expenditure, we also use food expenditure in order to validate our methodology.

By virtue of the detailed consumption measures in the FES we are also able to obtain expenditure on items related to risk behaviours. Specifically, we consider expenditures on fruits and vegetables, alcohol, and cigarettes. While the first two of these items suffer from the problem that expenditure changes might reflect quantity or quality changes, this is less the case for cigarettes. Compared to other products, the price differences among brands are very small at any given point in time. By using a special price index for cigarettes and assuming between-product price variation to be negligible, we are able to back out actual cigarette quantity.

3.2 The General Household Survey (GHS)

The GHS is a general purpose survey that has been conducted since 1971. Questions on health measures and risk behaviours have been included since 1978, our data cover the period until 2003. It was not fielded in 1997 and 1999, we explain in Section 4.3 how we deal with these years. Please refer to Table 1 in the Appendix for the availability of specific variables. We select the following health variables: Self-reported health status on a three-dimensional scale (good-fair-poor), whether respondents have a chronic disease, and whether they are limited in their activity by this disease. As an example, consider the proportion reporting good health that is shown in Figure 3. Consistent with many findings from the literature (Case and Paxson 2005), women report worse health than men at younger ages. However, the decline is slower, leading to little if any gender differences at age 60. Education effects seem to be even stronger. At age 30, only small differences between education groups can be seen. Proportions remain

nearly constant for people in their thirties, the subsequent decline begins earlier and is stronger for persons with low education. There are no apparent cohort differences in either direction.

Furthermore, the GHS includes some self-reported data on risk behaviours. We make use of questions on the quantity of cigarettes smoked and a six-category question regarding the amount of alcohol consumed, running from zero consumption to heavy drinking. With respect to the former, we use whether or not the respondent is currently smoking and an unconditional cohort average of the number of cigarettes smoked per day. Regarding alcohol consumption, we simply average the variable across cohort members.

3.3 The Health Survey for England (HSE)

Further health measures are available from the HSE, albeit only in a relatively small number of years, namely from 1991 to 2003. What makes matters worse is that cell sizes are very small in 1991 and 1992 which is made clear by the last two rows in Table 2. Accordingly, we checked robustness to exclusion of these two years in all specifications.

The HSE was set up to monitor population health in England and it includes a physical examination by a nurse. However, each year a different subset of health outcomes is emphasised (examples include cardiovascular disease and the health of older people). Consequently, few variables are contained every year in the nurse interview. In fact, the only one we are able to use is bloodpressure. We construct a binary variable that is one if the measured bloodpressure is above 140/90 mmHg (i.e. if the systolic pressure is above 140 or the diastolic pressure is above 90 or both) or if the respondent is currently taking bloodpressure medication. The fraction of persons with such a condition is listed in Figure 4. The age gradient is very steep compared to other variables. There is a large gender effect with women being much less likely to suffer from such a condition. There seems to be a slight tendency for lower prevalence among the better educated. We also use an indicator variable that takes the value one only if measured bloodpressure is above 140/90 mmHg and doesn't take account of medication. The two bloodpressure variables treat medical progress in a different fashion. In the course of the analysis, we term the two variables "Bloodpressure Condition" and "High Bloodpressure", respectively. Going in the same direction, we also use a self-reported measure of whether the respondent suffers from cardiovascular disease. A major reason to put an emphasis on diseases of the cardiovascular system stems from research in epidemiology which suggests causal mechanisms running from stress (that may be caused by sudden income drops) to cardiovascular markers which operate rather quickly as compared to many other diseases such as cancer or liver disease. See Brunner (1997) for a review and Steptoe, Willemsen, Owen, Flower, and Mohamed-Ali (2001) for a specific experiment.

Besides bloodpressure, we obtain some self-reported measures from the HSE. In particular, similar questions to those in the GHS are asked and we can use them to complement the years 1997 and 1999 although there is an issue with respect to comparability (see section 4.3). The longstanding illness question is further broken down by type of disease, of which we use the prevalence of of respiratory diseases and the just-described measure of cardiovascular disease. Again, epidemiological research suggests that smoking, a risk factor we consider below, has more rapid effects on respiratory diseases than lung cancer, for example see Peto, Darby, Deo, Silcocks, Whitley, and Doll (2000). The last health measure we include in the analysis is the mean score on the 12-item General Health

Questionnaire, a widespread instrument to measure mental health status. The questions call for yes or no answers and the individual score is simply the number of affirmations. A higher score indicates poorer mental health.

The first thing to establish is that the ubiquitous health-SES gradient is indeed present in our data. From 1997 on, the HSE contains an income measure that we can use to estimate the conditional correlation at the micro level. Table 4 shows the results of regressions of the various health variables on log household income and a set of controls at the level of the individual. The general pattern that emerges is that persons with higher income have better health, whatever the health measure we consider, consistent with many findings from the literature.

4   Results

We present our results in four stages. First, we describe the income process and the variance of permanent income shocks. We then show the effect on mortality, health outcomes and risk behaviours. The concluding part of a section describes the relation of our approach to other works in the literature.

4.1   The Variance of Income Shocks

Table 5 lists the autocovariances of the unexplained growth rate of income up to order 5. Estimation is only based on equation (10) and moment condition (12). All numbers in this table and Table 6 are multiplied by 1000 for easier reading. An MA(0)-process implies that $E[g_t^Y g_{t-1}^Y] < 0$ and that all autocovariances of higher order are zero. The data seems to point to such a model as all covariances two years apart or more are both small in magnitude and not significantly different from zero. The exception is for men as the covariance for higher order lags rises somewhat again in absolute value. However, they are rather imprecisely estimated. Hence, Table 5 suggests that an order 0 moving average process for the transitory shock to income is sufficient to capture its dynamics at the cohort level. This is consistent with the analyses of Banks, Blundell, and Brugiavini (2001) on the same data, albeit over a slightly earlier time period. They find that an MA(1)-process fits the quarterly income data from the FES very well. Hence one would expect an MA(0)-process for the yearly data.

Table 6 presents the estimated variance of the permanent shock to income, applying the methodology described above. All variances are significantly larger than zero. The results correspond to annual shocks with a standard deviation between 3 to 5% of income. The magnitude of the shocks is larger for individuals with low education, although the difference is not statistically significant. Gender differences in the point estimates are only visible for the highly educated with men displaying a somewhat higher variance of income shocks. However, remember that our measure is household income and differences are caused only by single households which constitute around 25% of the total sample. To get a very rough idea about the magnitude of these shocks, it is useful to compare them to the findings reported in Meghir and Pistaferri (2004). They estimate the variance of permanent shocks to individual income in the US using the PSID. The magnitude of the shocks we find is somewhere around 5% of their estimates. Obviously one should not overstate cross-country comparisons, but this seems to be a reasonable

magnitude of cohort-wide income shocks as compared to individual ones. [6]

4.2   Permanent Income Shocks and Mortality

Table 7 presents our estimates of $\gamma^{C,P}$ (the effect of cohort-level permanent income shocks) on mortality outcomes in nine different specifications – allowing for three different time lags for the income effects and three different age trend specifications. The first thing to note is that a sufficiently flexible specification of the age trend is extremely important when it comes to mortality: Moving from the quadratic to the cubic specification all coefficients change their sign and confidence intervals do not overlap. On the other hand, the cubic trend seems to be sufficiently general as a comparison with the results on the quartic trend shows.

In general, we find that an increase in income leads to an increase in mortality. When we allow no lagged effect of income, a one percent increase in income leads to about 0.7 to 1 more deaths per 100,000 persons among the prime aged population in any given year. By and large, we obtain similar results when we extend the lag with which income shocks can affect health although the null hypothesis of no causation is just not rejected at the 95% confidence level (P-value 0.058).

The demography literature has generally considered the log odds of mortality rather than mortality in levels (Elo and Preston 1996) because it is closer to a linear function of key variables such as age. This model cannot be derived from an individual-level specification because it depends on a nonlinear transformation of an aggregated variable. Nevertheless, a legitimate concern with our specification is that the dependence of mortality on the order of the age trend is because the levels of mortality are highly nonlinear in age. Estimating a log mortality version of model reveals similar results — we find zero or negative coefficients for a quadratic trend in the first step; and significantly positive coefficients for higher order trends. None of the results for the other variables were sensitive to higher order age trends than a quadratic.[7]

4.3   Permanent Income Shocks and Health

Table 8 contains our basic results for the relation between permanent income shocks and health outcomes.[8] Except for mental health, all dependent variables are measured prevalence rates in the population. These can be thought of as stemming from a linear probability model with our unidimensional health stock on the right hand side. Since we use household income in logarithms, the coefficients measure absolute changes in prevalence rates with respect to percentage changes in income. Taking the first entry of Table 8 as an example, a one percent shock to permanent income would decrease the fraction of people reporting good health by 0.065 percentage points. In other words a one standard deviation increase in income would reduce the fraction of individuals reporting good health from about 60%

---

[6]We also estimated the income process on different subset period and found no statistically significant differences.

[7]Detailed results are not presented here but were provided to the editor and referees and are available from the authors on request.

[8]In Section 3 and Table 1 we noted that the GHS was not fielded in 1997 and 1999. The HSE also contains a question on self-assessed health, but the answer is recorded on a five-point scale (very good, good, fair, poor, very poor). We collapse the "good" and "poor" categories into one each and use a dummy variable in the first estimation step for these years. With respect to the longstanding and limiting illness variables, we also use the values from the HSE and a dummy. In all four cases, results were largely insensitive to using either the values from HSE for all years beginning in 1993; or to cutting off the analysis in 1996.

to 59.7%. Moreover, the effect is not significantly different from zero. This effect is obtained assuming that income has only an instantaneous effect on health. When we additionally allow a one year delay, we still find an insignificant negative effect on the probability of reporting good health. With a two year lag, the effect of an increase in income becomes slightly beneficial to health, but with a small magnitude: a one percent increase in income increases the probability of reporting good health by 0.02 percentage points. This effect is not significant either.

We obtain a symmetric effect for the probability of reporting poor health. At lags zero and one year, an increase in income increases the proportion of individuals reporting poor health while after two years, the proportion is reduced. Once again, all the effects are not significantly different from zero. Income does not appear to affect other measures of health such as the proportion of individuals with longstanding illness or limiting illness. If anything, a one percent increase in income is associated to a 0.02 percentage point increase in these proportions.

Rows 5 to 7 in Table 8 present the effect on objective health outcomes relating to cardiovascular health: The proportion of individuals with blood pressure condition, with high blood pressure or with cardiovascular diseases.[9] We do not find any significant effects of income on these health outcomes either. This is important as one may think that this set of diseases could be linked to changes in socio-economic position due to stress, see the references listed in Section 3.3 among others. Neither can we reject the null hypothesis that income shocks have no effect on respiratory diseases: All point estimates in row 8 of Table 8 are negative but confidence intervals are too large to make statistically meaningful statements.

Mental health is measured as the simple sum of the scores on the 12-item general health questionnaire. A higher score indicates poorer mental health. We find that a positive shock to permanent income leads to poorer mental health. The confidence interval is rather large, but we can exclude the possibility of no effects at the 5% level. Looking at more lags, the coefficient appears ill identified and we cannot exclude the possibility that income has no effect on mental health.

Our strategy here has been to extract permanent shocks to income directly from income data. The life-cycle model predicts consumption to react stronger to permanent income innovations in income than to transitory ones. An alternative strategy to use as sensitivity analysis is to use data on expenditure to filter out permanent shocks as described in Blundell and Preston (1998) and Blundell, Pistaferri, and Preston (2005). We operationalise this idea by using total expenditure instead of income in the estimation procedure. The three moment conditions (12) to (14) remain unchanged even though we do not need to take care of transitory consumption movements if the life-cycle model is taken to be literal. This is because the presence of measurement and aggregation error introduces a transitory component in the consumption process anyhow. Results based on this strategy are largely comparable to our estimates from the income process. In particular, none of the coefficients changed significantly in magnitude. [10]

---

[9]One year each is missing for blood pressure and mental health (see Table 1), we interpolate the adjacent values linearly for these years.

[10]The tables corresponding to Tables 8 and 9 calculated using this methodology are available from the authors upon request.

4.4 Permanent Income Shocks and Behaviour

We now turn to the effect of income on behaviour and particularly on risk factors for health. We first look at the effect of permanent shocks on total (non durable) consumption. While total consumption is not thought to be directly linked to health, it is a useful benchmark to compare to more health related consumption decision.

The first row of Table 9 presents the elasticity of total expenditure with respect to permanent income shocks. A one percent increase in income is associated with a 0.4 to 0.6 percent increase in total expenditure, depending on the order of the moving average. All effects are significantly different from zero. Our estimates are in the same range as Blundell, Pistaferri, and Preston (2005) on individual-level US data. In contrast with the findings in the previous section, permanent shocks lead to sizable changes in behaviour. We conclude that the income data contains pertinent variation (as Table 6 indicated as well) and that the lack of any significant effect of income on health is not due to lack in variation in the explanatory variable. Therefore, these results establish the ability of our empirical methodology to identify changes at the cohort level.

We now focus on a subset of total expenditure, the expenditure on food. A one percent increase in income leads to a 0.20 to 0.25 percent increase in food expenditure, although estimates are just shy of being significant. The elasticity of food is about half of the elasticity of total expenditure, suggesting that households are able to smooth food consumption. The third row presents the effect of permanent shocks to income on expenditure on fruits and vegetables. The effect varies with the lag order for the income effect, with an elasticity between 0.45 and zero but we cannot exclude the possibility that income has no effect. More importantly, point estimates have roughly the same magnitude as the ones on food consumption and we can never tell them apart. Hence small income shocks do not seem to cause individuals to switch to a healthier diet.

We now turn to risky behaviours such as smoking and alcohol consumption. We find a significant impact of permanent income shocks on cigarettes smoked. [11] We get similar results when we use self-reported quantities instead of quantities inferred from expenditures. This is reassuring as these two measures come from different data sets and reporting schemes. Overall, a one percent increase in income increases the unconditional number of cigarettes smoked per day by about 0.06 to 0.1. We also get significant results for the probability of being a smoker, although the magnitude of the point estimates seems a bit large as to have much confidence in this result.

Rows 7 and 8 of Table 9 present the results for alcohol consumption. Looking at the self-reported quantity in row 7, we find a significant effect if looking at contemporaneous effects only. Point estimates decrease in size and fail to be significant for the more general models. Turning to the elasticity of alcohol expenditure, we find it to be around 0.5. As we consider a delayed response to income shocks, the coefficients appear to be less stable with large standard errors.

To summarise, we find some evidence that permanent income shocks affect behaviour and particularly health behaviour. Broadly speaking, permanent increases in income appear to lead to poorer health behaviour. We also estimated separately the effects shown in Tables 8 and 9 by gender, but we could not find significant differences

---

[11]An exception is when we select a lag of one year which gives an unstable parameter.

across males and females.[12]

4.5 Robustness of Results

The results presented above rely on the identification assumption that variations in permanent income at the cohort level are not caused by permanent shocks to cohort health, i.e. $\xi^{C,P} = 0$. While we believe there is a plausible justification for this assumption, we can explore the robustness of our results to this central assumption, i.e. for values of $\xi^{C,P} > 0$.

Assuming $p = 0$, the effect of income shocks on health can then be expressed as:

$$(15) \qquad \sum_{j=0}^{q+1} \gamma_j^{C,P} = \frac{E\left[g_{ct}^H \left(\sum_{j=-(q+1)}^{1} g_{c,t+j}^Y\right)\right] - E\left[g_{ct}^H\right] E\left[\sum_{j=-(q+1)}^{1} g_{c,t+j}^Y\right] - \xi^{C,P} \sigma_{\zeta^H}^2}{E\left[g_{ct}^Y \left(\sum_{j=-1}^{1} g_{c,t+j}^Y\right)\right] - E\left[g_{ct}^Y\right] E\left[\sum_{j=-1}^{1} g_{c,t+j}^Y\right] - (\xi^{C,P})^2 \sigma_{\zeta^H}^2}$$

Note that when $\xi^{C,P} = 0$, we get back to equation (14). The extent of the bias of the effect of income on health depends both on the magnitude of the reverse effect but also on the covariance between income, lagged income and health. The sign of the bias is a priori ambiguous.

Table 10 displays the effect of income on health for $\xi^{C,P} = 0$ (baseline), for $\xi^{C,P} = 0.1$ and for $\xi^{C,P} = 1$, which means that a 10 percentage point rise in the proportion reporting good health would lead to an 1 (respectively 10) percent increase in cohort income. The rest of the formula is calculated using moments from the data and setting $\sigma_{\zeta^H}^2 = 0.1$ (estimated from the data). The baseline case $\xi^{C,P} = 0$ simply reproduces the results from the first column of Table 8.

Our results are surprisingly robust to a relaxation of our main identifying assumption. For $\xi^{C,P} = 0.1$, the coefficients are remarkably similar to the baseline estimates. For larger effects of cohort health on cohort income ($\xi^{C,P} = 1$), the change in the coefficients are larger, especially for blood pressure conditions. However, if anything, these estimates reinforce our previous results in the sense that, from the point estimates at least, an increase in income appears to worsen health.

4.6 Relation to the Literature

In this section we briefly compare our methodology and results to those previously found by authors who looked at mortality, health, and health behaviour with some sort of aggregate data, in particular the work of Deaton and Paxson (1998b, 2001, 2004), Ruhm (2000, 2003, 2005, 2006) , and Ruhm and Black (2002). The main differences relate to the question of how the relationship between income (or macroeconomic conditions) and health is restricted by the various modelling approaches. Common to all studies is that causality running from health (behaviours) to income is not modelled. It is useful to reformulate our model of how health evolves over the life cycle (11) as:

$$H_{ct} = H_{c,0} + \sum_{s=0}^{t} \left(m_{s-c}^H + \gamma^{C,P} \zeta_{c,s}^Y + \zeta_{c,s}^H\right) + (1-L)\left(\gamma^{C,T} \varepsilon_{c,t}^Y + \varepsilon_{c,t}^H\right)$$

---

[12]Detailed results are not presented here but are available from the authors on request.

We suppress all lagged effects of the income shock on health and moving average parameters in the transitory terms as well as the measurement error terms for ease of notation. The formulation of the age trend makes use of the fact that $a = t - c$. This shows that current health is a function of some initial condition, the difference between the current and initial transitory effects, and the accumulation of the age trend, income and health shocks. From this equation (and the equivalent one on the income process) we move on to identify the parameters of interest from deviations in the cohort growth rates of income and health from their projections.

Deaton and Paxson construct datasets similar to ours for the United States (and also England and Wales in the 2004 paper). Their basic model [13] looks as follows in our notation:

$$H_{ct} = m_a + \beta_1 t + \beta_2 Y_{ct} + \beta_3 \cdot X_{ct} + u_{ct}$$

where X_{ct} represents a vector of covariates. H_{ct} is operationalised by using the log odds of mortality. Identification is then based on cohort differences in income and mortality conditional upon a nonparametric age profile, the linear time trend, and the other covariates. This strategy does not allow for cohort fixed effects that are correlated with the residuals. A cause for this could be long-term effects of macroeconomic conditions at birth on mortality (van den Berg, Lindeboom, and Portrait 2006). More general than such a fixed effect is the question of whether or not mortality may experience persistent shocks. If this is the case, then

$$\mathbb{E}[u_{c,t+s}|\mathbb{I}_{c,t}] \neq 0, \quad \text{some } s > 0$$

where $\mathbb{I}_{c,t}$ represents the information available in period t. The relevance of this problem depends on the persistence of the process. In our nonstationary formulation it holds for all $s > 0$ and is potentially serious. If shocks exhibit only limited persistence, it is probably not too important in practice.

In order to compare the two approaches on our data, we ran a mortality regression based on the following model:

$$H_{ct} = m_a + \beta_c + \beta_2 Y_{ct} + \beta_3 \cdot X_{ct} + u_{ct}$$

where β_c represents the coefficient on a cohort dummy, the age trend is interacted with gender and allowed to be quadratic (cubic, quartic), X_{ct} contains the education measure. Y_{ct} is again specified as log equivalised household income. Overall, results for β_2 go into the same direction as the ones for $\gamma^{C,P}$, but there are important differences. If the age trend is specified to be quadratic, results show a significantly negative correlation of income and health. However, this again seems to be an artefact of specifying the age trend in too restrictive a fashion. Turning to higher order polynomials, we get point estimates of $\hat{\beta}_2 = 0.1$ which is about a tenth of the magnitude observed for the coefficient on the permanent shock. It is never significant. This is consistent with the results reported in Deaton and Paxson (2004) for a specification that is just slightly different. It is difficult to pin down the exact reason for this discrepancy. Either it could be the importance of the focus on permanent shocks and the filtering out of transitory movements and measurement error; or else the coefficients based on (??) are biased due to reasons laid out in the preceding paragraph. In any case, we conclude that the more involved strategy employed here is not a futile complication of matters.

---

[13]The authors also investigate a direct effect of income inequality on mortality. We do not consider the part of their analysis that focusses on this issue here.

Another strand of the literature has investigated the impact of macroeconomic fluctuations on aggregate mortality (Ruhm 2000), on mortality from heart disease (Ruhm 2006), on various health measures (Ruhm 2003), on drinking behaviour (Ruhm and Black 2002) and on several other risk behaviours and markers of unhealthy behaviour(Ruhm 2005). The identification strategy in this work has been based on deviations in state unemployment rates from the (US) national average, controlling for a time trend and population (in case of the mortality studies) or UK individual (health and health behaviour studies) characteristics. This research has shown that for the population above 20 years of age, most of these variables move in accordance with employment rates. The interpretation was that unhealthy behaviour is one of the factors underlying the procyclical variation in mortality rates.

While these results are broadly consistent with our findings in this paper, we want to highlight some important differences. The results in the mortality studies (Ruhm 2000, 2006), have been based on aggregate data. It is well known that this type of data does not permit models that are nonlinear in the *variables*. While this is probably not an issue if employment rates are the dependent variable, it does pose a problem if the research question relates to the effect of income. The function linking income and health is typically thought to be concave. In our particular case, it turned out to be very important indeed to specify household income in logarithms as opposed to levels and we can aggregate such variables exactly from the underlying microdata. Another point concerns the use of log mortality as the outcome of interest which underlies most specifications in the two just-cited papers. While this may be important to capture the aggregate relationship well, it inhibits the interpretation of the results at the individual level. Finally, macroeconomic conditions may affect cohorts in a very different way. For example, unfavourable conditions at the beginning of the working life may impact negatively upon all future labour market outcomes (Raaum and Røed 2006), thereby constituting a permanent shock; whereas they may affect older cohorts' incomes only transitorily.

The only contrasting results and interpretations are found with respect to health outcomes as contained in Ruhm (2003). The basic identification strategy is the same as the one for the mortality studies, but individual-level data from repeated cross sections are used. Unfortunately, no results are reported for the overall prime-aged population which renders specific comparisons difficult. In the cases where analysis is restricted to a similar age range as the one considered here the sample is also restricted to the working population introducing a problem of sample selection. Furthermore, it is not quite clear what kind of effect is picked up by state-level unemployment rates if the sample is restricted to employed persons. Finally, the data analysis is cross-sectional in the sense that no individual-level fixed effects are controlled for, which relates relates immediately to the sample selection issue. Results based on these data in the US show a procyclical variation of health.

We cannot detect such an effect. On the other hand, we are rather confident that we would be able to pin it down if it were there. Our conclusion is that there may not be much instantaneous causation running from risk behaviours to health outcomes and on to mortality. Maybe one would expect the most direct effects with respect to smoking. However, if we look at respiratory diseases (which are expected to react much quicker than cancers), if anything, we find a countervailing effect. In our view, procyclical mortality is more likely to be driven by work-related accidents than by health and health behaviours. We can only provide some indirect evidence using data from the UK Longitudinal Study. Accidents make up 30% of all deaths at age 30 and this proportion declines roughly

linearly to 5% at age 60. This decline is mainly driven by the rise of other deaths; there is only a very slight decrease in the absolute number of accidents. Hence the baseline is large enough to drive our results (0.7 to 1 more deaths per 100,000 people for a 1% rise in income). Using a similar approach as Deaton and Paxson (2004), we do not find statistically significant results, however. This is not surprising in the light of the above-mentioned results on overall mortality and the low number of cohorts that we have available in the Longitudinal Study. Finally, Ruhm (2000) finds that accidents are the source of death that is most responsive to changes in macroeconomic conditions. Among the age groups, the largest effect of the unemployment rate on log mortality is found for those below 45 years. This combination is certainly suggestive for our interpretation, but we do not pursue it further. Given the indirect nature of this evidence, more research is needed to disentangle the precise causation channels.

Finally, reconciliation of the lack of evidence for time series causation found here with the strong correlations observed in many studies, even those which control for many possible covariates, is not really necessary given the often limited extent to which the latter studies can isolate exogenous changes in income or economic resources. The results of Lindahl (2005) and Meer, Miller, and Rosen (2003), however, using lottery gains and unanticipated bequests respectively to exploit exogenous changes in incomes at the individual level, warrant more of a comparison. One obvious point is that different relationships may of course be obtained in different institutional settings, or for different subgroups of the population. An alternative factor may be that the relationship of health to income shocks could be non-linear or even discontinuous, with the effect that large changes in resources, of the nature used in these two studies, yield more substantial (and more identifiable) health effects than the smaller changes in cohort averages that we use here, perhaps through channels such as changes in neighbourhoods and peer groups or major changes in housing or lifestyle more generally. More importantly, perhaps, such studies cannot, by nature of the empirical strategy on which identification is based, control for possible general equilibrium effects that would occur were income changes to affect larger groups of individuals. To the extent that we are using genuine fluctuations in aggregate incomes as drivers of potential health changes general equilibrium effects would be captured by our estimates, at least in the short run. One possibility is that such equilibrium responses, whether through prices or wages, act to dampen down any initial effects of the income change itself.

5   Conclusions

In this paper we estimate the effect of permanent shocks to income on health and health behaviour. We use a long time series of cohort data from multiple surveys which covers the life-cycle and report detailed information on income and health (both subjective and objective measures) as well as health behaviour. We exploit the fact that at the cohort level, over the eighties and nineties, there have been sizable changes in income mainly related to changes in the macro-economic environment (skill biased technological change, changes in the return to education and experience, increased competition). Our identification scheme relies on the fact that these changes are not due to changes in cohort health. While we allow for reverse causation at the individual level, as well as a rich stochastic structure, we aggregate it out at the cohort level to focus only on the effect of these observed income shocks on cohort health.

As in Ruhm (2000), we find some evidence that income and mortality are related, although in our case we decompose income and filter out the transitory component. While we find some evidence that permanent income shocks lead to poorer health behaviour, we find no evidence that it affects directly any of our health measures. Despite sizable shocks to income at the cohort level during our period of analysis, which can be empirically linked to changes in consumption expenditures, we fail to find any changes in (cohort) health, even if we allow those shocks to affect health with a lag. Moreover, the point estimates often hint that an increase in income leads to poorer health, except for blood pressure. This is in contrast with micro, cross-sectional evidence, where higher income is related to better health. Taken literally, our results suggest that this observed "gradient" does not stem from income innovations of the magnitude we observe at the cohort level.

Of course, we cannot rule out potential effects of income shocks that take longer to feed through into health outcomes than we have allowed for in the time-series processes estimated here (up to three years). Nevertheless, from a policy perspective our results seem particularly relevant, suggesting that the kind of changes to incomes that might be expected from redistribution towards particular age-education cohorts of working age adults are unlikely to lead to improvements in health at least in the short to medium run.

References

ABOWD, J. M., AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

ADAMS, H. P., M. D. HURD, D. MCFADDEN, A. MERRILL, AND T. RIBEIRO (2003): "Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status," *Journal of Econometrics*, 112(1), 3–56.

ALTONJI, J. G., AND L. M. SEGAL (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14(3), 353–366.

BANKS, J., R. BLUNDELL, AND A. BRUGIAVINI (2001): "Risk Pooling, Precautionary Saving and Consumption Growth," *Review of Economic Studies*, 68, 757–779.

BLUNDELL, R., P. PASHARDES, AND G. WEBER (1993): "What do we Learn about Consumer Demand Patterns from Micro Data?," *American Economic Review*, 83(3), 570–597.

BLUNDELL, R., L. PISTAFERRI, AND I. PRESTON (2005): "Consumption Inequality and Partial Insurance," IFS Working Paper 04/28.

BLUNDELL, R., AND I. PRESTON (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, 113(2), 603–640.

BROWNING, M., A. DEATON, AND M. IRISH (1985): "A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle," *Econometrica*, 53(3), 503–544.

BRUNNER, E. J. (1997): "Socioeconomic Determinants of Health: Stress and the Biology of Inequality," *British Medical Journal*, 314(7092), 1472–1476.

BUCHINSKY, M. (1994): "Changes in the US Wage Structure 1963-1987: Application of Quantile Regression," *Econometrica*, 62(2), 405–458.

CASE, A., D. LUBOTSKY, AND C. PAXSON (2002): "Economic Status and Health in Childhood: The Origins of the Gradient," *American Economic Review*, 92(5), 1308–1334.

CASE, A., AND C. PAXSON (2005): "Sex Differences in Morbidity and Mortality," *Demography*, 42(2), 189–214.

CHALOUPKA, F. J., AND K. E. WARNER (2000): "The Economics of Smoking," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse, pp. 1539–1627. Elsevier.

DEATON, A., AND C. PAXSON (1998a): "Aging and Inequality in Income and Health," *American Economic Association - Papers and Proceedings*, 88(2), 248–253.

——— (1998b): "Health, Income, and Inequality over the Life Cycle," in *Frontiers in the Economics of Aging*, ed. by D. A. Wise, pp. 431–460. University of Chicago Press, Chicago, IL.

——— (2001): "Mortality, Education, Income, and Inequality Among American Cohorts," in *Themes in the Economics of Aging*, ed. by D. A. Wise, pp. 165–170. University of Chicago Press, Chicago, IL.

——— (2004): "Mortality, Income, and Income Inequality over Time in Britain and the United States," in *Perspectives on the Economics of Aging*, ed. by D. A. Wise, pp. 247–280. University of Chicago Press, Chicago, IL.

DEHEJIA, R., AND A. LLERAS MUNEY (2004): "Booms, Busts, and Babies' Health," *Quarterly Journal of Economics*, 119(3), 1091–1130.

ELO, I. T., AND S. H. PRESTON (1996): "Educational Differentials in Mortality: United States, 1979–1985," *Social Science & Medicine*, 42(1), 47–57.

GOSLING, A., S. MACHIN, AND C. MEGHIR (2000): "The Changing Distribution of Male Wages in the UK," *Review of Economic Studies*, 64(4), 635–666.

GOTTSCHALK, P., AND R. MOFFITT (1994): "The Growth of Earnings Instability in the U.S. Labor Market," *Brookings Paper on Economic Activity*, 2, 217–254.

GOTTSCHALK, P., AND T. SMEEDING (1997): "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature*, 35(2), 633–687.

GROSSMAN, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2), 223–255.

LINDAHL, M. (2005): "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as Exogenous Source of Variation in Income," *Journal of Human Resources*, 40(1), 144–168.

LLERAS MUNEY, A. (2005): "The Relationship Between Education and Adult Mortality in the United States," *Review of Economic Studies*, 72(1), 189–221.

MACURDY, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 83–114.

MARMOT, M. (1999): "Multi-Level Approaches to Understanding Social Determinants," in *Social Epidemiology*, ed. by L. Berkmann, and I. Kawachi. Oxford University Press, Oxford.

MEER, J., D. L. MILLER, AND H. S. ROSEN (2003): "Exploring the Health-Wealth Nexus," *Journal of Health Economics*, 22(5), 713–730.

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

PETO, R., S. DARBY, H. DEO, P. SILCOCKS, E. WHITLEY, AND R. DOLL (2000): "Smoking, Smoking Cessation, and Lung Cancer in the UK Since 1950: Combination of National Statistics with Two Case-Control Studies," 321, 323–329.

RAAUM, O., AND K. RØED (2006): "Do Business Cycle Conditions at the Time of Labor Market Entry Affect Future Employment Prospects?," *The Review of Economics and Statistics*, 88(2), 193–210.

RUHM, C. J. (2000): "Are Recessions Good For Your Health?," *Quarterly Journal of Economics*, 115(2), 617–650.

——— (2003): "Good Times Make You Sick," *Journal of Health Economics*, 22(4), 637–658.

——— (2005): "Healthy Living in Hard Times," *Journal of Health Economics*, 24(2), 341–363.

——— (2006): "A Healthy Economy Can Break Your Heart," NBER Working Papers 12102, National Bureau of Economic Research, Inc.

RUHM, C. J., AND W. E. BLACK (2002): "Does Drinking Really Decrease in Bad Times?," *Journal of Health Economics*, 21(4), 659–678.

SMITH, J. P. (2004): "Unraveling the SES-Health Connection," IFS Working Paper 04/02.

SNYDER, S. E., AND W. N. EVANS (2006): "The Effect of Income on Mortality: Evidence from the Social Security Notch," *Review of Economics and Statistics*, 88(3), 482–495.

STEPTOE, A., G. WILLEMSEN, N. OWEN, L. FLOWER, AND V. MOHAMED-ALI (2001): "Acute Mental Stress Elicits Delayed Increases in Circulating Inflammatory Cytokine Levels," *Clinical Science*, 101, 185–192.

VAN DEN BERG, G. J., M. LINDEBOOM, AND F. PORTRAIT (2006): "Economic Conditions Early in Life and Individual Mortality," *American Economic Review*, 96(1), 290–302.

A   Tables

Table 1: Data Availability by Year
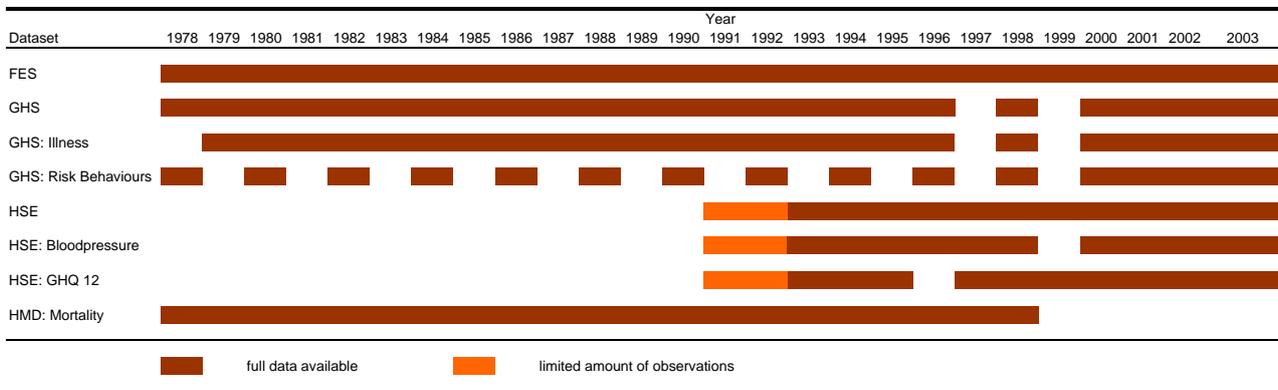
Table 2: Cell sizes after aggregation

Dataset	Mean	Median	First Percentile	Minimum	Total Number of Individuals
FES	137.3	134.5	65	55	148,517
GHS	257.3	221	84	58	277,084
HSE	149.6	140.5	25	18	80,541
HSE, 1993+	164.0	152	62	51	76,943

Table 3: Descriptive Statistics Averaged Across Cohorts

Men	Low Education				High Education			
	Age 30	Age 40	Age 50	Age 60	Age 30	Age 40	Age 50	Age 60
Education (FES) *)	0.555	0.489	0.476	0.589	0.445	0.511	0.524	0.411
Education (GHS) *)	0.507	0.469	0.478	0.563	0.493	0.531	0.522	0.437
Education (HSE) *)	0.521	0.440	0.430	0.538	0.479	0.560	0.570	0.462
Log Income (HH)	5.008	5.040	5.166	5.045	5.355	5.347	5.490	5.374
Log Total Expend. (HH)	4.612	4.672	4.789	4.625	4.833	4.894	5.023	4.980
Log Food Expend. (HH)	3.001	3.116	3.186	3.231	3.047	3.188	3.260	3.309
Log FruitVeg Expend. (HH)	0.741	0.904	1.031	1.085	1.088	1.175	1.357	1.382
Cigarettes (FES, HH)	2.167	2.113	2.286	2.323	2.042	1.954	1.956	2.134
Log Alcohol Expend. (HH)	1.808	1.742	2.076	1.847	1.964	1.788	2.001	1.928
Good Health	0.715	0.669	0.600	0.456	0.796	0.770	0.722	0.630
Poor Health	0.047	0.080	0.141	0.213	0.031	0.045	0.090	0.122
Longstanding Illness	0.265	0.303	0.369	0.547	0.253	0.246	0.327	0.508
Limiting Illness	0.110	0.162	0.230	0.385	0.104	0.115	0.174	0.278
Smoking Y/N	0.452	0.437	0.400	0.377	0.315	0.283	0.294	0.235
Cigarettes (GHS)	8.233	8.766	8.054	6.268	4.821	4.380	5.539	3.958
Alcohol Quantity	3.399	3.346	3.259	3.068	3.408	3.331	3.308	3.149
Bloodpressure Condition	0.243	0.278	0.488	0.654	0.243	0.220	0.423	0.624
High Bloodpressure	0.250	0.247	0.436	0.574	0.244	0.217	0.381	0.557
Cardiovascular Condition	0.017	0.041	0.100	0.218	0.007	0.037	0.080	0.176
Respiratory Condition	0.075	0.060	0.056	0.091	0.069	0.077	0.059	0.043
Mental Cond. (GHQ 12)	1.271	1.284	1.285	1.358	1.551	1.373	1.190	0.959
Mortality Rate / 1000 persons (both education groups)	0.908	1.802	5.180	15.736				

Women	Low Education				High Education			
	Age 30	Age 40	Age 50	Age 60	Age 30	Age 40	Age 50	Age 60
Education (FES) *)	0.513	0.471	0.491	0.542	0.487	0.529	0.509	0.458
Education (GHS) *)	0.532	0.456	0.534	0.591	0.468	0.544	0.466	0.409
Education (HSE) *)	0.493	0.436	0.479	0.510	0.507	0.564	0.521	0.490
Log Income (HH)	4.836	5.050	5.160	4.902	5.272	5.343	5.498	5.272
Log Total Expend. (HH)	4.495	4.682	4.788	4.543	4.786	4.912	5.038	4.903
Log Food Expend. (HH)	3.011	3.130	3.248	3.224	3.091	3.227	3.274	3.332
Log FruitVeg Expend. (HH)	0.767	0.907	1.139	1.084	1.105	1.236	1.351	1.422
Cigarettes (FES, HH)	2.134	2.143	2.270	2.271	1.937	1.864	2.029	2.187
Log Alcohol Expend. (HH)	1.521	1.739	1.963	1.671	1.697	1.692	1.993	1.714
Good Health	0.613	0.583	0.498	0.436	0.683	0.705	0.666	0.569
Poor Health	0.093	0.135	0.173	0.184	0.071	0.074	0.095	0.133
Longstanding Illness	0.280	0.314	0.408	0.545	0.242	0.260	0.321	0.510
Limiting Illness	0.124	0.190	0.250	0.341	0.098	0.124	0.175	0.308
Smoking Y/N	0.460	0.443	0.409	0.376	0.253	0.247	0.231	0.242
Cigarettes (GHS)	7.294	7.321	6.526	5.087	3.089	3.194	3.039	3.226
Alcohol Quantity	2.725	2.834	2.644	2.422	2.919	2.906	2.922	2.709
Bloodpressure Condition	0.085	0.131	0.354	0.646	0.064	0.123	0.337	0.548
High Bloodpressure	0.077	0.105	0.294	0.547	0.062	0.108	0.306	0.449
Cardiovascular Condition	0.030	0.046	0.073	0.208	0.012	0.033	0.088	0.183
Respiratory Condition	0.101	0.073	0.101	0.094	0.084	0.061	0.084	0.100
Mental Cond. (GHQ 12)	1.832	1.600	1.812	1.856	1.768	1.578	1.622	1.336
Mortality Rate / 1000 persons (both education groups)	0.475	1.189	3.305	8.865				

Note: *) Data from 1991-2000 only to ensure comparability across datasets. All other numbers are based on total number of observations for each variable, see Table 1. HH: Variable measured at the Household Level, equivalised by OECD scale.

Table 4: Regressions Showing the Correlations Between Health and Income.

Dependent Variable	Coef.	Std. Err.	N
Self-Reported Good Health	0.415**	0.010	36,641
Self-Reported Poor Health	-0.491**	0.014	36,641
Longstanding Illness	-0.213**	0.009	36,639
Limiting Illness	-0.332**	0.009	36,639
Bloodpressure Condition	-0.067**	0.012	23,871
High Bloodpressure	-0.053**	0.011	27,567
Cardiovascular Condition	-0.130**	0.013	36,639
Respiratory Condition	-0.092**	0.012	36,639
Mental Condition (GHQ 12)	-0.535**	0.018	35,313

Note: Coefficients of simple probit regressions of health outcomes on household income in logarithms (Mental Condition: OLS). All specifications further include year dummies as well as sex and a two-dimensional education measure, both interacted with a quadratic age trend. ** significant at the 5% level.

Table 5: The Autocovariances of the Unexplained Growth of Income

Order	Low Education		High Education		NT
	Men	Women	Men	Women	
0	6.50**	7.25**	5.91**	6.12**	944
	(5.18 , 8.04)	(6.08 , 8.48)	(4.67 , 7.03)	(5.37 , 6.68)	
1	-2.38**	-2.70**	-2.22**	-2.54**	888
	(-3.18 , -1.70)	(-3.60 , -1.82)	(-2.95 , -1.50)	(-3.36 , -1.83)	
2	-0.41	0.07	-0.27	0.57	832
	(-2.13 , 1.02)	(-1.32 , 1.22)	(-1.16 , 0.80)	(-0.42 , 1.67)	
3	-0.19	-0.07	0.93	-0.22	776
	(-0.92 , 0.49)	(-0.98 , 0.89)	(-0.15 , 1.86)	(-1.39 , 0.72)	
4	0.89	-0.12	-1.30**	-0.22	720
	(-0.26 , 2.28)	(-1.22 , 0.93)	(-2.19 , -0.30)	(-1.32 , 0.94)	
5	-1.00**	-0.12	0.80	0.07	664
	(-1.75 , -0.20)	(-1.42 , 1.00)	(-0.49 , 1.90)	(-1.06 , 1.16)	

Note: Bootstrapped 95% confidence intervals in parentheses, ** significant at the 5% level. All values scaled up by 1000.

Table 6: The Variance of the Permanent Shock to Income

Time Period	Low Education		High Education		NT
	Men	Women	Men	Women	
1978 − 2003	2.257**	2.027**	1.675**	0.872	832
	(0.60 , 4.09)	(0.43 , 3.72)	(0.75 , 2.57)	(-0.32 , 1.84)	

Note: Bootstrapped 95% confidence intervals in parentheses, ** significant at the 5% level. All values scaled up by 1000.

Table 7: The Effect of Permanent Income Shocks on Mortality.

Variable	Age Trend	Moving Average Parameter (q)			NT
		0	1	2	$q = 0/1/2$
Mortality	Quadratic	-0.786**	-1.108**	-1.013**	320 / 272 / 224
		(-2.118 , -0.076)	(-3.202 , -0.136)	(-2.204 , -0.091)	
Mortality	Cubic	0.721**	0.729	0.527	320 / 272 / 224
		(0.329 , 1.617)	(-0.063 , 3.629)	(-0.090 , 1.561)	
Mortality	Quartic	1.000**	1.151	0.650	320 / 272 / 224
		(0.469 , 1.953)	(-0.082 , 6.448)	(-0.231 , 2.276)	

Note: Mortality refers to number of deaths per 1000 persons with respect to a 1% change in household income. Bootstrapped 95% confidence intervals in parentheses.

Table 8: The Effect of Permanent Income Shocks on Health Outcomes

Variable	Time Period	Moving Average Parameter (q)			NT $q = 0/1/2$
		0	1	2	
Good Health	1978-2003	-0.0647 (-0.2189 , 0.1111)	-0.1683 (-0.5471 , 0.1532)	0.0235 (-0.2835 , 0.3372)	832 / 720 / 608
Poor Health	1978-2003	0.1074 (-0.0165 , 0.2564)	0.1012 (-0.0684 , 0.3454)	-0.0781 (-0.3070 , 0.2196)	832 / 720 / 608
Longstanding Illness	1979-2003	0.0228 (-0.1252 , 0.2361)	0.0595 (-0.2589 , 0.2888)	0.0592 (-0.2525 , 0.3728)	832 / 720 / 608
Limiting Illness	1979-2003	0.0264 (-0.1083 , 0.1739)	0.0750 (-0.2272 , 0.2959)	-0.0999 (-0.3416 , 0.1470)	800 / 688 / 576
Bloodpressure Condition	1991-2003	-0.1342 (-0.9238 , 0.2461)	0.4226 (-0.3116 , 1.4468)	0.1660 (-0.3655 , 0.8034)	352 / 272 / 192
High Bloodpressure	1991-2003	-0.2106 (-1.0594 , 0.4111)	0.4100 (-0.1700 , 1.2800)	0.0361 (-0.4885 , 0.7392)	352 / 272 / 192
Cardiovascular Condition	1991-2003	0.0006 (-0.1917 , 0.3149)	-0.1112 (-0.3034 , 0.1809)	0.1978 (-0.1156 , 0.8175)	352 / 272 / 192
Respiratory Condition	1991-2003	-0.2440 (-0.6579 , 0.0592)	-0.1751 (-0.5663 , 0.2026)	-0.2114 (-0.5878 , 0.0541)	352 / 272 / 192
Mental Health	1991-2003	0.037** (0.0056 , 0.6592)	0.2244 (-0.0052 , 0.8781)	0.0081 (-0.0290 , 0.7814)	352 / 272 / 192

Note: Except for mental health, all values denote changes in percentage points of prevalence rates with respect to a 1% shock to permanent income. Mental health: change in average score on the GHQ12 with respect to a 1% shock to permanent income. Bootstrapped 95% confidence intervals in parentheses, ** significant at the 5% level.

Table 9: The Effect of Permanent Income Shocks on Behaviour

Variable	Time Period	Moving Average Parameter (q)			NT $q = 0/1/2$
		0	1	2	
Total Expenditure	1978-2003	0.4365** (0.2289 , 0.6381)	0.4524** (0.1547 , 0.7756)	0.6491** (0.1230 , 1.0909)	832 / 720 / 608
Food Expenditure	1978-2003	0.2135 (-0.0158 , 0.4572)	0.2494 (-0.0308 , 0.6378)	0.2435 (-0.1693 , 0.6902)	832 / 720 / 608
Fruits and Veg. Expenditure	1978-2003	0.2604 (-0.0590 , 0.8441)	0.4596 (-0.0612 , 1.2263)	-0.0085 (-0.9402 , 1.0431)	832 / 720 / 608
Fraction of Smokers	1978-2003	0.0656 (-0.1485 , 0.3393)	0.3673** (0.0521 , 0.7280)	0.3778** (0.0132 , 0.9539)	832 / 720 / 608
Cigarettes via Self-Rep. Qty.	1978-2003	0.0824** (0.0475 , 0.9420)	0.0966** (0.0589 , 0.2549)	0.1025** (0.0534 , 0.3999)	832 / 720 / 608
Cigarettes via Expenditure	1978-2003	0.0631** (0.0354 , 0.1846)	0.4539 (-0.0183 , 0.9135)	0.0592** (0.0355 , 0.5185)	832 / 720 / 608
Alcohol Self-Rep. Qty.	1978-2003	0.0059** (0.0024 , 0.0125)	0.0044 (-0.0058 , 0.0374)	0.0041 (-0.0078 , 0.2533)	832 / 720 / 608
Alcohol Expenditure	1978-2003	0.5325 (-0.3041 , 1.5006)	0.4306 (-1.0875 , 34.1617)	-0.0502 (-2.0854 , 36.3326)	832 / 720 / 608

Note: Expenditure (Total, Food, Fruit and Vegetables, and Alcohol): Coefficients shown are elasticities. Smoking Fraction: Change in percentage points of prevalence rate with respect to a 1% shock to permanent income. Cigarette variables: Change in daily number of cigarettes smoked (unconditional) with respect to a 1% shock to permanent income. Alcohol Self-Reported Quantity: Change in average reported value with respect to a 1% shock to permanent income. Bootstrapped 95% confidence intervals in parentheses, ** significant at the 5% level.

Table 10: Robustness of Results.

Variable	$\xi^{C,P} = 0$	$\xi^{C,P} = 0.1$	$\xi^{C,P} = 1$
Good Health	-0.0647	-0.0717	-0.1487
Poor Health	0.1074	0.1144	0.1918
Longstanding Illness	0.0228	0.0297	0.0963
Limiting Illness	0.0264	0.0335	0.1191
Bloodpressure Condition	-0.1342	-0.1269	-0.0217
High Bloodpressure	-0.2106	-0.2033	-0.0967
Cardiovascular Condition	0.0006	0.0075	0.0716
Respiratory Condition	-0.244	-0.2372	-0.1781
Mental Health	0.037	0.0454	0.2727

Note: The calculations are described in detail in Section 4.5. For the health variables defined in negative terms, the sign of ξ_2 is reversed.
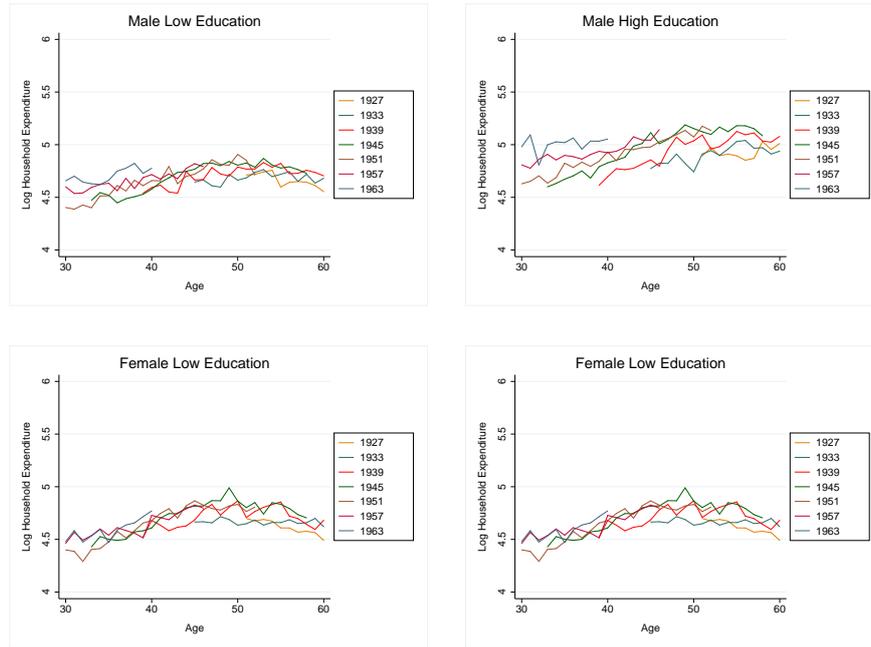
B   Figures

Figure 1: OECD-equivalalised Household Income in Logarithms by Cohort
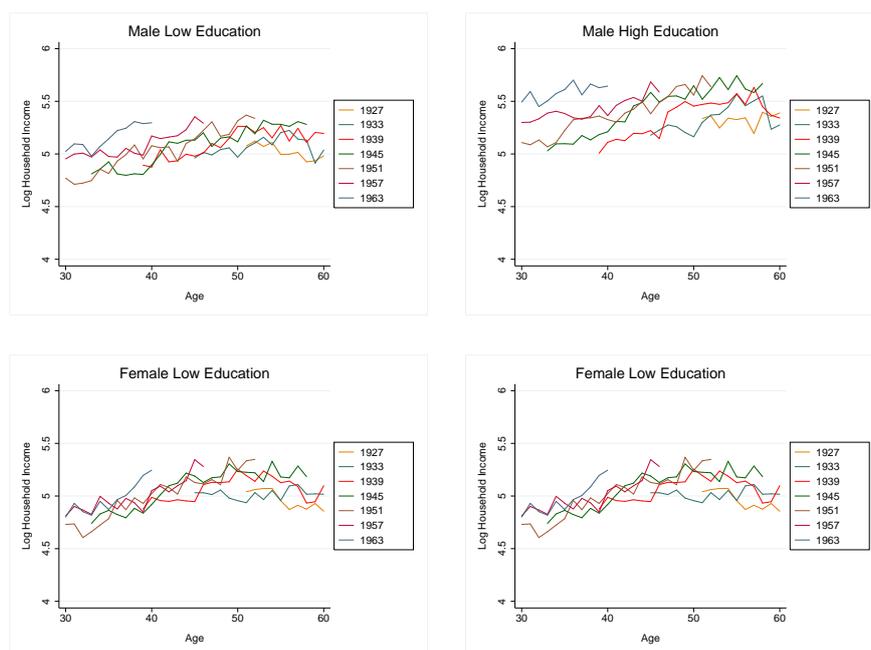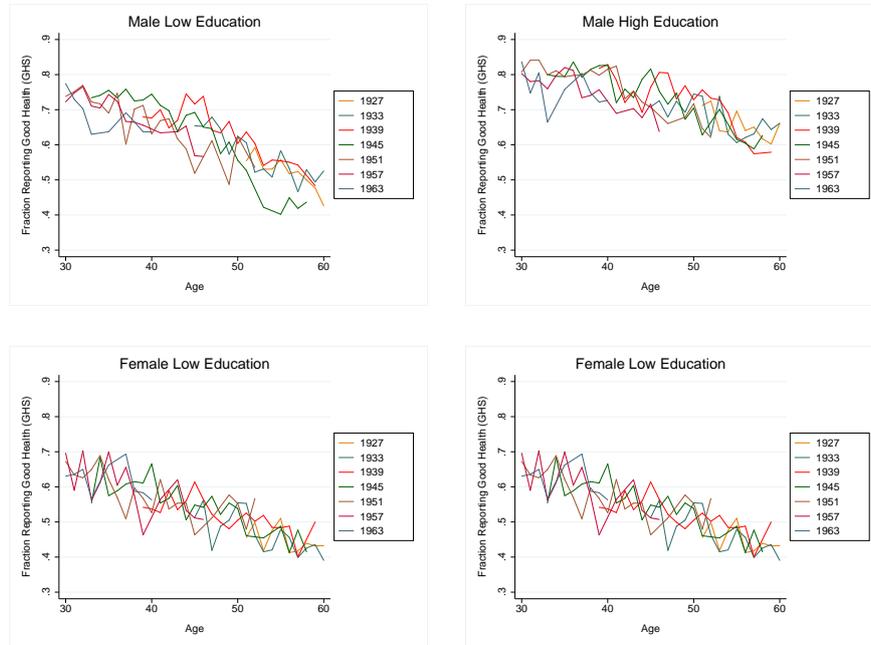

Source: FES 1978-2003, own calculations. Only every other cohort is shown for legibility reasons.

Figure 2: OECD-equivalised Total Household Expenditure in Logarithms by Cohort
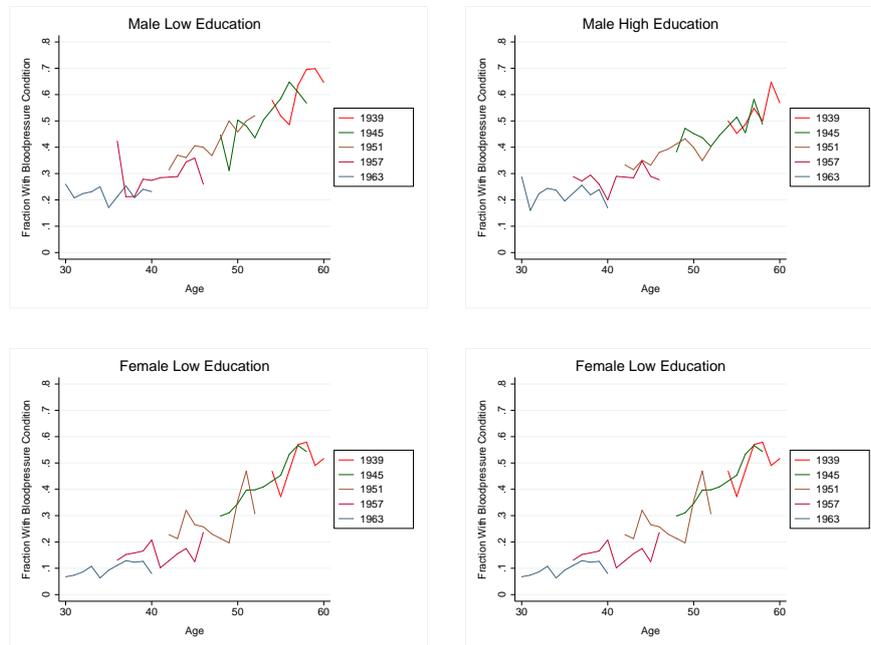
Figure 3: Fraction Reporting Good Health by Cohort

Figure 4: Fraction with Bloodpressure Condition by Cohort

33