the review of income and wealth

Review of Income and Wealth
Series 63, Number 4, December 2017
DOI: 10.1111/roiw.12243

GROWTH IN EARNINGS AND HEALTH: NOTHING IS AS PRACTICAL AS A GOOD THEORY

A NOTE ON EARNINGS GROWTH AND MOVEMENTS IN SELF-REPORTED HEALTH BY HALLIDAY

BY DANIEL AVDIC AND MARTIN KARLSSON

CINCH, University of Duisburg-Essen

We discuss some issues associated with the empirical analysis of the relationship between socioeconomic status and health. We point out that, in addition to elaborate empirical modeling and good data, a conceptual framework is helpful both for making sense of one's own results and for the purpose of reconciling results across studies. We find that when we align the empirical specification with the Grossman model, a negative effect of income on health emerges. Even though this unexpected finding can be rationalized, we think that some caution regarding standard dynamic panel data techniques is warranted in this context.

JEL Codes: C33, I10, I14

Keywords: dynamic panel data, health production, socioeconomic disparities

1. INTRODUCTION

The empirical observation that individuals from lower socioeconomic backgrounds report poorer health, the so-called socioeconomic gradient in health, is one of the more thoroughly investigated topics in the history of the social sciences. Perhaps the most influential contribution in this field of research is the seminal article by Marmot and coauthors using the famous Whitehall study on British Civil servants in which a strong correlation between occupation grade and the risk of cardiovascular disease was documented (Marmot *et al.*, 1978). Numerous articles have since then been written attempting to unravel the relationship between social class and health outcomes. However, despite the immense importance for explaining the causes and consequences of social inequalities across the world, no consensus has been reached regarding the causal relationships between health and socioeconomic status.

There are several reasons why such a consensus has not been reached. One problem relates to how health and social status are measured since both

Note: We have borrowed this expression from Kurt Lewin, a German-American psychologist (1890–1947). Lewin's motto was "Nothing is as practical as a good theory," a statement implying that while theory is a prerequisite for good practice, good theories should also have sound practical foundations. We thank Ieva Sriubaite for excellent research assistance and Daniel Kamhöfer and Matthias Westphal for valuable comments on the draft.

*Correspondence to: Martin Karlsson, Chair of Health Economics, University of Duisburg-Essen, Weststadttürme Berliner Platz 6-8, 45127 Essen, Germany (martin.karlsson@uni-due.de).

concepts are not easily operationalized and it could be challenging to compare studies using different definitions. A related problem concerns the use of different data sets and, most importantly, different empirical strategies to disentangle causal relationships. The main empirical issue is that, while social class may affect health outcomes, the opposite relationship is also likely to hold. Whether the focus has been on the former or the latter relationship, many authors have developed and applied elaborate identification strategies to isolate and quantify the effects of interest. Furthermore, reporting heterogeneity have been found to bias the relationship in the sense that individuals from lower social backgrounds tend to report poorer health than other individuals (Etilé and Milcent, 2006). As a consequence, we have witnessed significant improvement in the understanding of inherent problems with measurement errors, selection and other issues related to the interpretation of estimated parameters. Unfortunately, however, as in other mainly empirically driven research fields, theory has been largely left out of the analysis (Deaton, 2010). This leads to the problems that the interpretation of results, and comparison across studies, may often be confusing in the absence of a conceptual framework to base the motivation of the empirical strategy and its results on.

This note discusses the above-mentioned empirical issues in the context of the contribution "Earnings Growth and Movements in Self-Reported Health" written by Halliday. The paper uses a dynamic panel data approach to the relationship between earnings and self-reported health of individuals sampled from the Panel Study of Income Dynamics (PSID) between the years 1984 to 1993. The paper diligently scrutinizes the empirical problem of selection and reverse causality and estimates variations of a regression model in which the focus is on quantifying the causal link from earnings to self-rated health. To address these challenges, the author sets up a dynamic first-differenced model where changes in self-reported health is regressed on earnings growth under the assumption that earnings are pre-determined; with identification relying on moment restrictions on the econometric model. The author finds some mixed results, but concludes that the overall pattern seem to indicate a positive causal effect running from income to self-assessed health for certain population groups.

Dynamic panel data models like those used in the study by Halliday have delivered an extremely powerful tool for estimating causal relationships in situations where the dependent variable, as in this case health, is inherently persistent. By putting some often very plausible restrictions on the intertemporal correlations between left- and right-hand side variables and the error term, it is possible to consistently estimate parameters representing the dynamic relationship between different variables, while at the same time allowing for unobserved heterogeneity to be absorbed by the individual fixed effect. However, the initial optimism regarding the methods as they were developed during the 1990s has waned somewhat in recent years. Several authors have argued that dynamic panel data methods offer little guidance on how to pick the appropriate identifying assumptions—which leaves a great degree of discretion with the individual researcher (Roodman, 2009; Bontempi and Mammi, 2012; Biørn, 2015).

Our main departure and motivation for writing this text is the observation that, in many applications, it is equally important to discuss the *economics* of the research question as it is to statistically model the relationship. We exemplify how a simple conceptual framework can assist in rationalizing assumptions; support interpretation; allow for comparisons of the obtained results; as well as give guidance on which empirical specifications to choose and implement. In the current study, the identifying assumption that earnings are predetermined is presented as an improvement over the alternative assumption that earnings are strictly exogenous. This is indeed a considerable improvement, since strict exogeneity would rule out any influence of health on future earnings. But what if we judge the assumption by its own merits? We argue that it would be desirable to include theory to guide us in the choice of which exclusion restrictions to maintain in order to isolate the parameter of interest. While a theoretical framework naturally does not eliminate the identification problem, it could nevertheless aid in the interpretation of obtained results as well as putting them in a relevant context. Perhaps most importantly, such results might also allow for more accurate predictions of impacts of prospective social policies.

Hence, this note does not argue specifically against the identification strategy used in the paper but merely suggests that, in the absence of an economic model for any statistical relationships found in the analysis, it is difficult to judge and compare the meaning and implications of obtained results. Just as a theoretical model stands void of any practical meaning without a corresponding system in real economic life, statistical relationships are of limited utility unless paired with appropriate economic models of human behavior (Heckman, 2010).

2. Theoretical Framework

We will discuss the hypothetical relationship between income and health under a simplified form of the well-known Grossman model of the demand for health investments (Grossman, 1972). According to Grossman, health has the dual role of being an investment good and a consumption good at the same time. By investing in their health, individuals may spend more time in productive employment and thus earn a higher income. At the same time a higher income may be spent on health-improving activities. Thus, even though neither health nor income is a direct choice variable in Grossman's model, the model does explain how the two are determined. This can be illustrated in the simple two-equation model

$$y = wt(h), \tag{1}$$

$$h = m(y). \tag{2}$$

In equation (1), income (y) is a function of the wage (w) and inputs of time (t) (hours of work). We assume, for simplicity, that only the latter is determined by health (through reductions of days on sick leave).[1] In equation (2), health (h) is a

[1]An extended model could also include wages as a function of health as well as the probability of job loss (and the reception of unemployment benefits).

function of inputs, such as medical expenditures m (e.g. doctor visits, medical drugs, and health-improving activities) and paid for out of earnings. In this static model the stock of health affects productivity and therefore income. This simultaneity between income and health creates an identification problem for researchers that want to study the impact of one factor on the other and any hope of a solution must be motivated through means of exclusion restrictions; i.e. assumptions on the dependency structure between the variables in the model.

Introducing dynamics, we follow Grossman (1972) in including a flow component to health which depreciates over time and therefore needs to be replaced by health-improving investments in order to keep productivity and, consequently, income from declining. Medical expenditure and other health investments then enter as a means of reducing unhealthy time (on sick leave) and therefore serves to uphold earnings. Grossman assumes that investments in health (e.g. medical expenditures) affect the stock of health in the next period,

$$y_{t+1}=wt_{t+1}(h_{t+1}), \tag{3}$$

$$h_{t+1}=h_t+[I(y_t)-\ \delta_t]. \tag{4}$$

Consequently, we assume that negative changes in health at time t, measured by the health deprecation rate δ_t, reduce health and therefore also earnings in the following time period $t+1$. We assume no such instantaneous reverse relationship running from income to health as we, following Grossman, think of health as an *investment* good. The assumptions on the dynamics of the model are paramount for our choices of empirical specification and identifying assumptions when we turn to estimation of the model.

3. Data and Empirical Specification

To estimate the impact of income on health, we specify an empirical model based on the assumptions of the process as outlined in the previous section. Hence, introducing heterogeneity, the following process for health is assumed (for individual i at time t)

$$h_{it}=\alpha_i+\gamma h_{i,t-1}+\beta_g y_{it}+\phi_g a_{it}+v_{it},\ \ g\in G, \tag{5}$$

where α_i is an individual-specific intercept, a_{it} is the age of the individual and β_g is the group-specific parameter of interest measuring the impact of earnings on health. The index g corresponds to the sex and marital status of the individual for which separate regressions are subsequently estimated.[2] To eliminate the individual fixed effect, the model is first-differenced,

$$\Delta h_{it}=\gamma\Delta h_{i,t-1}+\beta_g\Delta y_{it}+\phi_g+\Delta v_{it.} \tag{6}$$

[2]Marital status is included in order to follow the original article as closely as possible.

Finally, to identify the parameters of interest, β_g, the author makes the assumption that health only affect earnings with a lag while earnings may affect health instantaneously. Formally,

$$E[\upsilon_{it}|h_{it-1}, y_{it}]=0, \tag{7}$$

implying that any instantaneous effect of health is excluded from the earnings equation. This assumption is implemented in a GMM dynamic panel data estimation framework in which lagged versions of earnings and health are used as instruments for the first differenced factors in the model.

The crucial point to make here is how to rationalize the exclusion restriction in (7), which identifies the effect of earnings on health. To this end, the author refers to other papers in which strictly stronger identifying assumptions have been maintained and proposes this alternative parametrization as an improved way of conducting inference. However, and as the author also is well aware of, relaxing assumptions comes at a cost as it typically requires more of the data and may lead to problems of small sample inference; in particular the issue of weak instruments. As we argue throughout, an alternative way of motivating identifying assumptions is to assess them by the logic of an appropriate theoretical framework.

In our case, the simplified Grossman framework outlined above suggests that health investments affect health with a lag. This is not the assumption made in equation (6) where contemporaneous earnings affect health. Furthermore, the model states that health *instantaneously* affects earnings though reduced working hours (due to less healthy days). Hence, this feature of the Grossman model *violates* the moment restriction in (7) in which the author assumes that health can only affect earnings with a lag, which could be true for, e.g. subsequent wage cuts, job loss and other more slow-moving effects of poor health. However, the main point is that it is quite likely, both theoretically and empirically, that poor health has an immediate effect on earnings for working individuals through increased sickness absence and general productivity.

Specifically, the Grossman model stated in equations (3)–(4) above would suggest that we should estimate the following model for the evolution of health,

$$h_{it}=\alpha_i+\gamma h_{i,t-1}+\beta_g y_{it-1}+\phi_g a_{it}+\upsilon_{it} \tag{8}$$

where the depreciation term δ_t is proxied by the age of the individual and allowed to vary (linearly) across age categories. To identify the parameters of interest we assume, based on the theoretical model, that

$$E\left[\upsilon_{it}|h_{i,t-1}, y_{i,t-1}\right]=0. \tag{9}$$

That is, compared to the situation in the paper, where health does not impact earnings instantaneously, we now assume that it may while income in the *previous* period is predetermined, given health in the previous period. The reason is that we assume that individuals *invest* in their health and the return shows up only in the following period.[3] Together with differencing out the individual fixed effect using the first difference transformation this assumption may allow us to consistently estimate β_g.

To estimate the model we make use of the PSID as in the original article. Of course, the theoretical model outlined above may be just as misspecified as any other model. We have not solved the identification problem by assuming a different model structure, but we have made the analysis more transparent and generated a number of testable hypotheses deduced from our theory. These hypotheses provide an opportunity to check whether our theoretical model matches the data or not.

First, given the model framework, we would expect γ to be positive and ϕ_g to be negative (higher health in period t increases health in period $t+1$ and health depreciation should increase with age). One important analysis would be to compare the effect across subgroups and the main effect from including the full sample; i.e. compare β with the different group-specific parameters $\beta_g,\ g \in G$. Starting with the full sample reduces multiple testing concerns, and a comparison is meaningful in the sense that it makes visible which group(s) contribute to the main effect.

However, splitting the sample by subgroup also possibly introduces a sample selection issue whenever the stratification variable is an individual choice. This is clearly the case for marital status, which requires two individuals to agree to form a relationship or to stay in a relationship.[4] If there is no change in marital status over the entire observation period, we may hope that this selection bias is absorbed by the individual effect. But movements into and out of a certain relationship status will inevitably lead to bias whenever unobservables in the health and relationship dimensions are correlated.[5] In terms of the estimating equations (5) and (9) above, the expected value of the residual υ will not be zero. Instead, in each subsample specification, υ is conditioned on the actual marital status being chosen by the individual, and it may thus well be a function of covariates and lagged values of health. This problem has been demonstrated convincingly by Heckman (1979) and in the notation used here it would imply

$$E\left[\upsilon_{it} | h_{\mathrm{i},t-1}, y_{i,t-1}, g_t\right] = E\left[\upsilon_{it} | h_{i,t-1}, y_{i,t-1}, U_g \geq U_{-g}\right] \neq 0 \qquad (9)$$

where g_t now represents current marital status, and U_g represents the utility associated with that marital status. However, the GMM framework possibly offers a way of credibly addressing the sample selection issue. The endogeneity problem is essentially identical to that of the lagged value of health. Thus, just as for health, we may use the twice-lagged value of the marital status variable as an instrument for its change between *t-1* and *t*.

[3]It should be noted that the required exposure time could be longer than one period. We disregard that possibility here, even though there are many examples of longer exposure times in the literature (van den Berg *et al.*, 2006).

[4]We disregard bereavement here, but it has similar empirical implications.

[5]There is evidence in the literature suggesting that this is indeed the case (Kohn and Averett, 2014).

TABLE 1
Sample Summary Statistics

	Aggregate	SM	SW	MM	MW
Health	0,86	0,85	0,79	0,90	0,73
	(0,35)	(0,35)	(0,40)	(0,30)	(0,45)
Earnings	8,68	8,48	7,49	9,29	6,79
	(3,85)	(3,80)	(4,25)	(3,51)	(4,42)
Age	38,39	35,16	39,01	38,91	39,04
	(9,58)	(8,67)	(10,46)	(9,30)	(9,99)
Married	0,65	0,00	0,00	1,00	1,00
	(0,48)	(0,00)	(0,00)	(0,00)	(0,00)
Male	0,75	1,00	0,00	1,00	0,00
White	0,60	(0,53)	0,41	0,71	0,23
#Observations	52,807	7,706	10,641	31,856	2,604
#Individuals	8,844	1,968	2,255	5,208	830

Note: The table reports average values (standard deviations) for the included variables. Earnings are measured in logs. Health is measured as a dummy variable derived from a 5-point scale taking the value one if health is regarded as Excellent = 5, Very good = 4 or Good = 3, and zero if health is regarded as Fair = 2 or Poor = 1.

We sample all household heads in the family files of the PSID family dataset for years 1984–1993. We merge this data with the PSID individual file and end up with a total of 8,844 individuals or 52,807 individual-year cells. Table 1 reports summary statistics of the sample and included variables by gender and marital status. The number of observations is fewer in the model estimations because of the first difference transformation and the inclusion of lagged instruments.[6]

4. Results

Table 2 reports the results from the baseline case in the paper (where assumption (7) is assumed to hold) under the assumption of exogeneity of earnings (panel A) and when current earnings are assumed to be predetermined (panel B) for the aggregate sample (column (1)) and for each gender-marital status subgroup individually (columns (2)–(5)). The models are estimated using differenced GMM with three lags as instruments for the endogenous variables. As expected, previous health is positively related to current health, both for the aggregate sample and for all subgroups. Furthermore, income is, whenever significantly different from zero, typically estimated to have a positive effect on health.

Panel C and D of Table 2 show results of pooled regressions when earnings (Panel C) and earnings and health (Panel D) are interacted with the four marital status-gender combinations. The main advantage of this approach is that the endogeneity of a person's relationship status can be addressed. If results in panels C and D differ a lot from the results in panel B, we should be concerned that endogeneity of marital status is an issue. Results in Table 2 are indeed quite robust, and contemporaneous earnings seem to have a positive effect on health

[6]The sample does not exactly correspond to the original article as it was difficult to assess the sampling procedure form the information given. We have therefore followed the sampling in Meghir and Pistaferri (2004), dropping all individuals who were never household head in the individual files as well as the Latino sample introduced in 1990.

TABLE 2
BASELINE CASE

	(1) Aggregate	(2) SM	(3) SW	(4) MM	(5) MW
A. Exogenous earnings					
γ	0.067***	0.087***	0.084***	0.082***	0.100**
	(0.013)	(0.027)	(0.021)	(0.011)	(0.048)
β	0.002**	0.002	0.000	0.002**	−0.008*
	(0.001)	(0.002)	(0.002)	(0.001)	(0.005)
#Instruments	*39*	*62*	*62*	*62*	*62*
F-stat	*226.06*	*32.47*	*31.49*	*114.1*	*11.25*
B. Predetermined earnings					
γ	0.088***	0.109***	0.075***	0.088***	0.097**
	(0.012)	(0.033)	(0.025)	(0.013)	(0.048)
β	0.001	0.006	−0.002	0.003*	−0.011
	(0.002)	(0.004)	(0.004)	(0.002)	(0.008)
#Instruments	*61*	*61*	*61*	*61*	*61*
F-stat	*224.06*	*32.18*	*30.79*	*107.36*	*10.95*
C. Predetermined earnings interacted					
γ	0.091***				
	(0.011)				
β		0.004**	−0.005	0.011**	−0.004
		(0.002)	(0.005)	(0.005)	(0.004)
#Instruments		*130*	*130*	*130*	*130*
F-stat		*240.31*	*240.31*	*240.31*	*240.31*
D. Both interacted					
γ		−0.063*	−0.033	0.099***	0.059*
		(0.036)	(0.026)	(0.013)	(0.033)
β		0.003**	−0.006	0.008*	−0.005
		(0.002)	(0.004)	(0.005)	(0.003)
#Instruments		*193*	*193*	*193*	*193*
F-stat		*244.09*	*244.09*	*244.09*	*244.09*
#Obs.	35,963	4,888	6,981	22,760	1,334
#Ind.	6,573	1,258	1,539	4,138	380

Note: Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

for males. However, the effect size fluctuates somewhat and it is in general smaller than in the original paper.

Table 3 reports the results from estimating the Grossman model where current earnings are assumed not to affect current health but one-year lagged earnings are assumed predetermined conditional on lagged health. Otherwise all specifications are the same as in Table 2. The results for the autoregressive parameter γ are as expected similar to the previous estimates but the results for earnings growth are quite different. In fact, some of the parameters are estimated with negative signs implying that earnings changes affect health negatively. The overall effect on the full sample is estimated to be negative, irrespective of whether lagged earnings are assumed to be exogenous or predetermined. There is limited agreement between specifications as regards the subgroups that are driving this effect. Panel D, which takes the endogeneity of marital status into account and yet allows for a lot of flexibility, attributes the negative effect to singles, both males and females. A similar result is found in panel B, where sample is split

TABLE 3
GROSSMAN CASE

	(1) Aggregate	(2) SM	(3) SW	(4) MM	(5) MW
A. Exogenous lagged earnings					
γ	0.066***	0.086***	0.078***	0.081***	0.105**
	(0.013)	(0.028)	(0.021)	(0.012)	(0.048)
β	−0.001**	−0.006***	−0.002	−0.001*	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.005)
#Instruments	*39*	*60*	*60*	*60*	*60*
F-stat	*224.58*	*43.16*	*31.93*	*109.11*	*9.13*
B. Predetermined lagged earnings					
γ	0.071***	0.063*	0.039	0.093***	0.100*
	(0.014)	(0.036)	(0.029)	(0.014)	(0.051)
β	−0.005***	−0.010**	−0.010**	0.001	−0.001
	(0.002)	(0.004)	(0.004)	(0.002)	(0.009)
#Instruments	*59*	*59*	*59*	*59*	*59*
F-stat	*239.07*	*42.56*	*36.68*	*110.26*	*8.60*
C. Predetermined lagged earnings interacted					
γ	0.073***				
	(0.013)				
β		0.000	−0.018***	0.015**	−0.011***
		(0.002)	(0.005)	(0.006)	(0.004)
#Instruments		*122*	*122*	*122*	*122*
F-stat		*249.72*	*249.72*	*249.72*	*249.72*
D. Both interacted					
γ		−0.064*	−0.048*	0.091***	0.069*
		(0.038)	(0.029)	(0.014)	(0.036)
β		−0.014***	−0.010***	−0.000	0.008
		(0.004)	(0.004)	(0.002)	(0.005)
#Instruments		*185*	*185*	*185*	*185*
F-stat		*252.63*	*252.63*	*252.63*	*252.63*
#Obs.	35,963	4,888	6,981	22,760	1,334
#Ind.	6,573	1,258	1,539	4,138	380

Note: Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

according to sex and marital status. However, all four subgroups have a negative and significant estimate in at least one specification.

What is the message to take home from all these results? The first message is a positive one: it appears that the sample selection problem mentioned above hardly matters: results in panels B and D of Tables 2 and 3 tend to agree, suggesting that the endogeneity of marital status hardly matters. Second, when we impose the realistic assumption that it is the lagged income that matters to health, we get perverse income effects. There are different possible explanations for this unexpected result. In a series of papers, Ruhm has shown that health tends to be countercyclical (Ruhm, 2012). The negative coefficient might thus be picking up a business cycle effect. A related literature shows that transitory income shocks tend not to improve health outcomes (Andersson *et al.*, 2015; Apouey and Clark, 2015). Alternatively, it could be argued that the parameter picks up subjective elements of self-assessed health which are orthogonal to the person's actual physical health. Finally, one could attribute the finding to limitations in the method. The reported F statistics suggest that we do not have a weak instrument problem.

However, the unexpected sign of the income effect in combination with our difficulties to replicate the original study and the lack of stability between relatively similar specifications do indeed call for some additional persuasion from proponents of the method.

5. Conclusions

The relationship between socioeconomic status and health has been extensively analyzed for decades without a consensus emerging regarding the causal links between the two factors. In this note we discuss some issues associated with an empirical analysis of these relationships and point out that, in addition to elaborate empirical modelling and good data, a hypothesized conceptual framework is helpful for both making sense of one's own results as well as for the purpose of reconciling results across studies. Using as an example a paper applying the PSID in a thorough empirical framework, we suggest complementing this with a simple Grossman model framework of investments in health. Such a framework allows the researcher to deduce an empirical specification in which assumptions made on the data generating process are based on a priori imposed restrictions on the theoretical model. We demonstrate how this conceptual framework may guide the choice of empirical specification and how it can ease the interpretation and validity of the estimated parameters of the regression model.

Our empirical results are not very uplifting. We find that when we align the empirical specification with the Grossman model, a negative effect of income on health emerges. Even though this unexpected finding can be rationalized with reference to previous findings in the literature, we think that some caution regarding standard dynamic panel data techniques is warranted in this context. Since the dynamic panel data estimators are very simple yet powerful econometric tools, it would be highly desirable to gain a better understanding of when they are appropriate, and when not.

References

Andersson, E., P. Lundborg, and J. Vikström," Income Receipt and Mortality—Evidence from Swedish Public Sector Employees," *Journal of Public Economics*, 131, 21–32, 2015.

Apouey, B. and A. E. Clark, "Winning Big but Feeling no Better? The Effect of Lottery Prizes on Physical and Mental Health," *Health Economics*, 24, 516–538, 2015.

Biørn, E., "Panel Data Dynamics with Mis-Measured Variables: Modeling and GMM Estimation," *Empirical Economics*, 48, 517–535, 2015.

Bontempi, M. E. and I. Mammi, "A Strategy to Reduce the Count of Moment Conditions in Panel Data GMM,"Available at SSRN 2131333, 2012.

Deaton, A., "Instruments, Randomization, and Learning about Development," *Journal of Economic Literature*, 424–455, 2010.

Etilé, F. and C. Milcent, "Income-Related Reporting Heterogeneity in Self-Assessed Health: Evidence from France," *Health Economics*, 15, 965–981, 2006.

Grossman, M. "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80, 223–255, 1972.

Heckman, J. J., "Sample Selection Bias as a Specification Error," *Econometrica*, 153–161, 1979.

——— "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 48, 356–398, 2010.

Kohn, J. L. and S. L. Averett," The Effect of Relationship Status on Health with Dynamic Health and Persistent Relationships," *Journal of Health Economics*, 36, 69–83, 2014.

Kruiniger, H., "GMM Estimation and Inference in Dynamic Panel Data Models with Persistent Data," *Econometric Theory*, 25, 1348–1391, 2009.
Meghir, C. and L. Pistaferri, "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72, 1–32, 2004.
Ruhm, C. J., "Understanding the Relationship Between Macroeconomic Conditions and Health," *The Elgar Companion to Health Economics*, 1, 2012.
Van den Berg, G. J., M. Lindeboom and F. Portrait, "Economic Conditions Early in Life and Individual Mortality," *American Economic Review*, 96, 290–302, 2006.