UNC CHARLOTTE ECONOMICS WORKING PAPER SERIES

IS THE SIMPLE LAW OF MOBILITY REALLY A LAW? TESTING CLARK'S HYPOTHESIS

Kelly Vosters

Working Paper No. 2016-003

THE UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE
BELK COLLEGE OF BUSINESS
DEPARTMENT OF ECONOMICS
9201 University City Blvd
Charlotte, NC 28223-0001
June 2016

ABSTRACT

Recent work by Gregory Clark and coauthors uses a new surnames approach to examine intergenerational mobility, finding much higher persistence rates than traditionally estimated. Clark proposes a model of social mobility to explain the diverging estimates, including the crucial but untested assumption that traditional estimates of intergenerational persistence are biased downward because they use only one measure (e.g., earnings) of underlying status. I test for evidence of this using an approach from Lubotsky and Wittenberg (2006), incorporating information from multiple measures into an estimate of intergenerational persistence with the least attenuation bias. Contrary to Clark's prediction, I do not find evidence of substantial bias in prior estimates.

Is the simple Law of Mobility really a law? Testing Clark's hypothesis

Kelly Vosters*

November 4, 2015

Abstract

Recent work by Gregory Clark and coauthors uses a new surnames approach to examine intergenerational mobility, finding much higher persistence rates than traditionally estimated. Clark proposes a model of social mobility to explain the diverging estimates, including the crucial but untested assumption that traditional estimates of intergenerational persistence are biased downward because they use only one measure (e.g., earnings) of underlying status. I test for evidence of this using an approach from Lubotsky and Wittenberg (2006), incorporating information from multiple measures into an estimate of intergenerational persistence with the least attenuation bias. Contrary to Clark's prediction, I do not find evidence of substantial bias in prior estimates.

Keywords: Intergenerational Mobility

JEL Classification Codes: J62

*Department of Economics, Michigan State University, 110 Marshall Adams Hall, East Lansing, MI 48824. Email: *vosters@msu.edu*

The author is grateful to Gary Solon for his guidance and suggestions. Thanks also to Jeff Wooldridge, Steven Haider, Leslie Papke, Margaret Brehm, and seminar participants at Michigan State University. The author acknowledges support from Institute of Education Sciences Grant R305B090011 to Michigan State University. The opinions expressed are those of the author and do not represent the views of the U.S. Department of Education.

1 Introduction

There has been long-standing interest in the persistence of outcomes across generations, from earlier theoretical work by Becker and Tomes (1976, 1979), to the development of intergenerational datasets enabling expansions of empirical work. These studies aim to describe, for instance, the extent to which inequalities are passed on from one generation to the next, or the extent to which opportunities or outcomes have been equalized for children from various family backgrounds. The typical approach to studying intergenerational mobility begins with a basic model relating children's outcomes to parents' outcomes:

$$y_{it+1} = \beta y_{it} + \epsilon_i \tag{1}$$

where i indexes family, t indicates parent's generation and *t+1* indicates the child's generation.[1] Generally, y_{it+1} and y_{it} represent a measure such as income, wealth, or education. The regression coefficient, β, then provides a measure of persistence, or *immobility*, in the outcome from the parent's generation to the child's generation. Hence, the quantity 1-β can be interpreted as a measure of mobility. For the U.S., the persistence parameter relating a child's log income to parent's log income (hence, an income elasticity) is estimated to be about 0.4 to 0.6 (Solon, 1999; Mazumder, 2005; Lee and Solon, 2009; Black and Devereux, 2011), while for Nordic countries the estimate is lower at 0.1 to 0.3 (Black and Devereux, 2011).[2] These estimates are taken to be summary statistics, describing the extent to which income differences persist from one generation to the next in a country or society. Among the explanations for the lower persistence observed in Nordic countries relative to the U.S. is one that highlights why so much attention is given to such differences in mobility: higher mobility may reflect policy differences, such as more redistributive tax structures and generous social welfare programs.

In a recently published book, though, Gregory Clark makes the provocative claim that these

[1]In equation (1), along with the remaining equations in the paper, the intercept is suppressed by considering the variables in deviation-from-mean form.

[2]This paper uses intergenerational income regressions as a point of departure, thus extending the income mobility literature, but there is a broader literature that looks at other outcomes. For example, Hertz et al. (2007) is an oft cited recent example providing intergenerational correlation and regression coefficients in educational attainment for 42 countries; Björklund and Salvanes (2011) also provide a succinct review of related literature. Additionally, another subset of the literature is concerned with intergenerational persistence in occupation or occupational prestige. Hodge (1966) is an early example studying intergenerational occupational mobility in the U.S., while Long and Ferrie (2007, 2013) are more recent examples; see also Black and Devereux (2011) for a brief discussion of related studies.

estimates are substantially biased downward, and that 'true' persistence in social status is much higher—approximately 0.75—and is uniform across all countries and over time (Clark, 2014). The latter part of Clark's assertion, regarding lower mobility, draws on a body of work by Clark and his coauthors, including an article in this journal, that uses innovative methods and a variety of creative names data sources covering many societies over several centuries.[3] The methods exploit the information content of *rare* surnames in these societies to explore social mobility, *without* having actual intergenerational family links.[4] The basic idea is that if inheritance matters, then rare surnames contain information on economic status, and they also indicate some family lineage given naming conventions and the inheritance of paternal surnames (or in some countries both maternal and paternal surnames).[5]

The first part of Clark's controversial claim—regarding bias in prior estimates—is based on a model proposed to explain the discrepancies between mobility estimates. Clark (2014) postulates that the higher persistence rate (0.75) governs a law of social mobility, and summarises the general intuition underlying the hypothesised downward bias in traditional estimates as:

> 'Families turn out to have a general social competence or ability that underlies partial measures of status such as income, education, and occupation. These partial measures are linked to this underlying, not directly observed, social competence only with substantial random components. The randomness with which underlying status produces particular observed aspects of status creates the illusion of rapid social mobility using conventional measures.' (Clark, 2014, p.8)

More formally, Clark and Cummins (2015) and Clark (2014) present a simple model for mobility:

$$x^*_{it+1} = bx^*_{it} + e_{it} \tag{2}$$

where x^* represents underlying social status, and b the 'true' persistence rate. The hypothesised

[3]See Clark (2014) for a comprehensive list of these studies, as well as the more recent papers Clark and Cummins (2015) and Clark et al. (2015).

[4]For the data sources containing explicit socioeconomic measures, such as probated wealth at death, equation (1) is estimated using the group averages of wealth for rare surnames. For data without such measures, the approach instead looks at persistence in the representation of the rare surname in an 'elite' group relative to representation in the population as a whole.

[5]Güell et al. (2014) show that rare surnames do contain such information, and propose a method using the joint distribution of surnames and economic status to explore intergenerational transmission of status in Spain.

attenuation bias in prior estimates is thought to arise from the focus on a single 'noisy' measure, y_{it} (e.g., income, wealth, or education), of the underlying social status, x^*_{it}, where this relationship is assumed to be of the form:

$$y_{it} = x^*_{it} + u_{it} \tag{3}$$

where u_{it} is idiosyncratic error.[6] Additionally, Clark claims to be able to measure the 'true' persistence rate by using surname group averages in equation (1), or $\bar{y}_{zt+1} = b\bar{y}_{zt} + \bar{u}_{zt}$, where z indexes surname (instead of i indexing family). The argument relies on classical measurement error assumptions so that $\bar{y}_{zt} \simeq \bar{x}^*_{zt}$ because $\bar{u}_{zt} \simeq 0$ when the surname samples are sufficiently large.[7]

In a recent article in this journal, Clark and Cummins (2015) present both traditional and surname estimates of social mobility in England using wealth measures to illustrate the discrepancies in mobility estimates, and also test implications of one dimension of the proposed model—the AR(1) form of the law of motion for social mobility in equation (2). However, they do not test the proposed explanation for the discrepancies:

> '... if we were to measure the social status of families as an aggregate of earnings, wealth, education, occupation, and health, then observed social mobility even in parent child studies would decline. For such an aggregation would reduce the variance of the error component in measured status. Thus the measured rate of persistence, even in one generation, will be much closer to that of the underlying latent variable.'
> (Clark and Cummins, 2015)

I fill this gap by exploring this hypothesis that when information from multiple measures is aggregated and then used to obtain traditional estimates, the lower mobility rates will be revealed.[8]

[6] Specifically, the assumption is that traditional estimates are biased downward by the usual classical measurement error attenuation factor $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}$, where σ_x^2 is var(x^*) and σ_u^2 is var(u).

[7] Clark (2014) notes that any group averaging over individuals would similarly reduce measurement error and reveal true status, thus resulting in much higher estimates of persistence (Clark, 2014, p.110). As noted by Solon (2015) however, many of the intergenerational mobility studies that use group averages do not actually find such results. For example, Chetty et al. (2014) show in Appendix D that using surname group averages from administrative U.S. income tax data results in estimates similar to the individual-level regressions.

[8] Other recent papers have been testing other hypotheses put forth in Clark's work. For example, in footnote 7, I mentioned the estimates from Chetty et al. (2014) that do not support Clark's assertion that mobility estimates based on any group averages over individuals will result in higher estimates. Clark (2014) also advocates that his results explain why findings from multigenerational regressions indicate a positive grandparent coefficient. In fact, as Solon (2015) points out, the papers do not all find positive coefficients. For instance, Lucas and Kerr (2013) find

This paper empirically tests the proposed existence, magnitude, and nature of a downward bias in traditional estimates. Conveniently, the theoretical setup for the law of social mobility laid out in equations (2) and (3) translates nicely into a latent variables framework, and the attenuation bias portion of this intriguing theory can be easily tested using publicly available data. Considering x^* the latent status, equation (2) can be interpreted as the structural equation. For each of the particular measures mentioned in the first quotation above (i.e., income, education, and occupation) we can write a separate measurement equation of the form presented in equation (3). Under the strong classical measurement error assumptions maintained in Clark's theory, instrumental variables (IV) using one noisy measure to instrument for another noisy measure produces a consistent estimate of the intergenerational coefficient (IGC), b. If the classical assumptions are relaxed to allow for slope coefficients in the measurement equations as well as unrestricted correlations among the measurement errors, IV estimation is inconsistent. The magnitude and direction of the inconsistency is potentially unknown, depending on the assumptions and measures used. However, an approach proposed by Lubotsky and Wittenberg (2006) is particularly well suited for addressing the case of multiple noisy measures, and under less stringent assumptions. While not identifying b, the method allows one to obtain an estimate with the least attenuation bias—so in this case a greatest lower bound on b—by incorporating information from all of the suggested measures (i.e., income, education, occupation) into a single estimate of b.

In this paper, I employ these approaches using a sample of fathers and sons from the Panel Study of Income Dynamics to test the attenuation bias assumption underlying the law of social mobility. I find little evidence supporting the hypothesised downward bias in prior estimates, and show that incorporating additional measures such as education and occupation has no meaningful impact on the estimated persistence rates obtained from traditional models focused on single measures. Considering intergenerational persistence in this more comprehensive sense does not reveal higher persistence estimates, but rather confirms the picture of mobility obtained from prior studies that focused on a single measure of socioeconomic status. The paper is organised as follows. Section 2 describes the data and sample. Section 3 outlines the empirical approach, and then Section 4

little evidence of non-zero grandparent coefficients in multigenerational regressions using administrative income data for Finland. Similarly, Braun and Stuhler (2015) use survey data on education and occupation in Germany and find that after controlling for parents' outcomes, they cannot reject a zero coefficient for the grandparents' outcome.

presents the results. Section 5 summarises the results and concludes.

2 Data

I use data from the Panel Study of Income Dynamics (PSID), as this data is ideally suited for my study. The data contains the requisite intergenerational links and also includes information on *multiple* measures of socioeconomic status, which is crucial for testing the attenuation bias claim.[9] Further, I am able to select a sample of individuals very similar to prior PSID studies about which the attenuation bias claims are made, thereby facilitating an appropriate comparison.[10]

The PSID is a longitudinal study that began in 1968 with a sample of approximately 5,000 families in the U.S., with interviews conducted annually through 1997, and biennially since then. Children from these original families are followed when they start their own households, and one can observe family links and follow multiple generations, which is key for traditional intergenerational mobility studies. This paper focuses on the Survey Research Center (SRC) part of the sample[11], in particular during the 1968-1972 surveys for fathers and 1992 survey for sons.[12] While more recent years are available, this time period allows for more direct comparability to prior estimates targeted by the proposed bias, lessens concerns about deterioration of data quality in later years, and still allows sons' ages to be appropriate for measuring earnings outcomes.

My analysis sample is comprised of sons who were members of the original 1968 sample and are male heads of their household in the 1992 survey, restricted to those who were born in 1951-1961. The lower bound on birth year ensures that the sons were 17 years of age or younger in 1968, avoiding selecting older children still living at home. Further, the sons' birth year restrictions

[9] Although administrative datasets such as the income tax records used by Chetty et al. (2014) have much larger samples, the data would not suffice for the tests conducted in this paper because information on other status measures such as educational attainment or occupation is not available.

[10] For example, Solon (1992) and Chadwick and Solon (2002) use similar father-son samples. Their sample selections differ in that son's earnings is observed starting at age 25. I restrict my sample to sons for whom I observe earnings starting at age 30 (up to age 40), to minimize life-cycle bias, as discussed below.

[11] The SRC sample was designed to be nationally representative in 1968, while the other component—the Survey of Economic Opportunity (SEO) sample—oversampled low income households.

[12] Focusing on father-son persistence in status rather than parent-child (or mother-daughter, etc.) is more straightforward given female labour force participation patterns, and the resulting issues with defining and measuring earnings and occupation outcomes. The surnames work also focused primarily on patrilineal lines of inheritance, given naming conventions (Clark, 2014, p.15), but still posited this general *law*. Hence, the proposed *law* of mobility should be just as evident using only fathers and sons as would be the case if mothers or daughters were included.

minimise life-cycle bias in annual earnings by ensuring that sons are 30 to 40 years old for the 1991 earnings measure (reported during the 1992 survey).[13] Fathers are identified as the male heads of the household in which the son lived in 1968. The earnings outcome for both fathers and sons is measured as log annual earnings, so the sample excludes any observations with non-positive earnings or earnings which were imputed by major assignment (for sons, this refers to earnings in 1991, and for fathers, earnings in each of the years 1967–71). Fathers missing data on educational attainment are also excluded. The earnings exclusions apply to 24 sons and 28 fathers, with 11 additional fathers excluded due to missing education, amounting to excluding a total of 46 father-son pairs, and leaving a final sample of 415 sons from 293 fathers.[14]

Table 1 provides summary statistics describing this sample. The sample is predominately white, with only five percent black. Given the age exclusions for sons (and lack thereof for fathers), the fathers are observed, on average, at an older age than sons, with fathers' average age just over 40 in 1967 and sons' average age approximately 35 in 1991. Average annual earnings are slightly lower for sons than fathers, and are also more variable for sons, consistent with the well-documented life-cycle profile in earnings.[15] Approximately 25 percent of the fathers have at least a four-year college degree.

For the empirical analysis, I define the education measure of father's latent status as father's educational attainment as of the 1968 survey, coded as 1-16 for years of schooling up to a 4 year college degree, with a value of 18 indicating any graduate school completed. The occupation measure refers to the main job discussed in the 1969 survey, and is incorporated in the form of occupational category indicators. As listed in Table 1, there are seven categories: 1) professional, technical; 2) managers, businessmen, self-employed; 3) clerical, sales; 4) craftsman, foreman; 5) operatives; 6) labourers, service workers, farmers, and farm managers; 7) miscellaneous (includes armed services

[13] Haider and Solon (2006) show that the measurement error in men's current earnings as an indicator of lifetime earnings is non-classical at younger and older ages, causing intergenerational persistence estimates to be biased downward (as also illustrated in Lee and Solon (2009)). They find that observing men's earnings from the early thirties through the early forties best avoids this life-cycle bias, as this is when the measurement error is approximately 'classical'. Findings presented by Nybom and Stuhler (*forthcoming*) show similar results using Swedish earnings data.

[14] It is possible to construct larger PSID samples, but I choose a sample similar to those in prior intergenerational studies since these were used to produce the U.S. estimates which Clark purports are biased downward, and are thus germane to the explorations in this paper. Further, Nybom and Vosters (2015) use Swedish administrative data to conduct similar tests as well as supplementary analyses examining the robustness of the results in this paper, showing that the results are not unique to this sample or the measures used.

[15] All earnings variables are expressed in 1991 dollars (adjusted for inflation using the CPI-U) for illustration purposes, but this transformation does not affect IGC estimates since the log of earnings is being used.

members, protective services workers, those not currently employed, and those missing an occupation category). To further illustrate the composition of the occupation categories for fathers, Table 2 provides average education and earnings by category. Average earnings and education are generally monotonically decreasing from occupation categories 1 to 6. The final category, 7–miscellaneous, is similar to categories 3 and 4, though with few observations and substantial variability in earnings. Hence, I take a flexible approach in the analysis, incorporating the occupation measure as a vector of indicators for each of the first six occupation categories (with category 7 the omitted reference group), taking no stance on the relative social status of the categories, but rather assuming that each contains some information on the underlying latent status.

3 Empirical approach

To test the hypothesis that traditional estimates of intergenerational persistence suffer from attenuation bias, I begin by providing a baseline traditional estimate from this PSID sample. I use the five-year average of log earnings from 1967-71 as the measure of father's status in equation (1), similar to previous studies (e.g., Solon, 1992; Zimmerman, 1992; Chetty et al., 2014).[16] Given that the proposed attenuation bias is thought to come from the focus on a single noisy measure of an underlying latent social status, and that incorporating additional measures such as education and occupation should reveal greater persistence in status, I extend the model by adding these other measures of father's status. I then estimate these intergenerational regressions using the typical ordinary least squares (OLS) approach, an instrumental variables (IV) approach, and the approach proposed by Lubotsky and Wittenberg (2006), to look for evidence of attenuation bias. In all estimations, I control for a quadratic in father's age and a quadratic in son's age to account for the life-cycle profile in earnings.[17]

[16]With classical noise in annual earnings measures, estimating equation (1) using OLS results in an IGC estimate that is biased downward by the well-known attenuation factor of $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}$, where σ_x^2 is var(x^*) and σ_u^2 is var(u). Taking the five-year average of earnings mitigates the attenuation bias, reducing the attenuation factor to $\frac{\sigma_x^2}{\sigma_x^2+(\sigma_u^2/5)}$. The attenuation factor becomes more complicated when one incorporates serial correlation in earnings from one year to the next.

[17]Including quadratics in both father's and son's age as controls arises from taking models of current earnings of the form $y_{it} = y_i + a_{i0} + a_{i1}Age_{it} + a_{i2}Age_{it}^2 + v_{it}$, for i = father or son and t = time (e.g., year), then solving for the long run component of earnings y_i, and substituting each into equation (1). Taking the five-year average of log earnings implies using the five-year average of age. See Solon (1992) for explicit derivations.

To more clearly illustrate Clark's theory discussed above in the context of my empirical approach, I present a more formal latent variables framework, with the intergenerational equation (1) now represented by the so-called structural equation:

$$y_{it+1} = \beta x^*_{it} + \epsilon_i \tag{4}$$

where y_{it+1} is son's log earnings, and x^*_{it} is father's underlying social status. Then we can consider equation (3) expanded to comprise the system of measurement equations:

$$y_{1it} = \rho_1 x^*_{it} + u_{1it} \tag{5}$$
$$y_{2it} = \rho_2 x^*_{it} + u_{2it} \tag{6}$$
$$\vdots$$
$$y_{jit} = \rho_j x^*_{it} + u_{jit} \tag{7}$$

In these measurement equations, y_{1it} represents the average of father's log annual earnings in 1967-71, y_{2it} is father's education, and y_{3it} is father's occupation (specifically, a vector of occupation category indicators).[18] Further, this framework allows for slope coefficients in the measurement equations, relaxing the theory presented earlier, which took these ρ_j to be equal to 1.[19]

This notation reflects the fact that I do not directly address the latent status for sons. If we were to take literally the simple law's assumption of classical measurement error on the left-hand side, there would be no concern of this limitation inducing bias. More generally, with any status measure on the left-hand side, we should still see growth in the intergenerational coefficient towards 0.75 as we add measures for fathers on the right-hand side if the proposed attenuation bias argument holds. Further, addressing status for sons is tricky, as there is no basis for obtaining optimal weights (discussed below) for son's measures on the left-hand side. Even so, I perform a robustness check by applying the weights determined for fathers' measures to those for sons to obtain a more comprehensive status measure for sons, and get very similar results.[20]

[18] The occupation indicators are generally referred to as one measure—occupation—even though occupation is flexibly accounted for by including an indicator for each occupation category. This implementation is similar to the drinking water proxy for wealth used in one of the examples presented in Lubotsky and Wittenberg (2006).

[19] Intercepts are omitted because the outcome, measures, and latent variable should all be considered to be demeaned, which is consistent with the implementation of the Lubotsky and Wittenberg (2006) approach discussed below.

[20] As discussed below with the results on robustness checks, the intergenerational coefficient obtained from this

Also under the perhaps unwise assumptions of classical measurement error, one method for consistently estimating β is instrumental variables (IV). One can use any y_j to instrument for another measure y_k and consistently estimate β, provided that $\sigma_{jk}\equiv cov(u_j, u_k) = 0$ and ρ_k=1 (otherwise, the estimate converges to β/ρ_k). Hence, this IV approach is slightly robust to failure of classical assumptions, allowing some $\rho_j\neq 1$. In the case where $\sigma_{jk} = 0$ fails, the IV estimator is no longer consistent for β, but the direction of bias may be intuitively inferred based on belief about the sign of $cov(u_j, u_k)$.[21] Although Clark's simple law assumes the measurement errors are uncorrelated (i.e., $cov(u_j, u_k) = 0$) there are obvious reasons to believe this assumption is violated in the setting considered here.[22] Thus, I next turn to my preferred approach which allows for this correlation.

The approach proposed by Lubotsky and Wittenberg (2006) (henceforth LW), not only produces a single estimate of β while incorporating multiple measures, but does so in an optimal way such that the estimate asymptotically provides the greatest lower bound on β. The approach results in the least attenuation bias by extracting the strongest combined signal out of all of the measures.[23] Hence, I can directly test the attenuation bias argument by observing whether estimates are converging to the hypothesised persistence rate of 0.75 as I incorporate additional noisy measures of father's status. Not only does the method allow for incorporating all available measures, it also relaxes the strong assumptions that $cov(u_j, u_k) = 0$ for all $j\neq k$ and ρ_j=1 for all j, allowing these to be mostly unrestricted (subject to a normalisation on ρ).[24]

The normalisation on ρ is needed to identify this vector of slope coefficients in the measurement equations. I normalise ρ_1 to equal 1, which simply sets the scale of the latent x^* to that of y_1 (earnings). Clearly latent status has no scale, but given that I am positioning this paper using

regression based on using income, education, and occupation for sons and fathers is 0.433, which is not significantly different from the estimate of 0.473 based on only income for sons.

[21]When $\rho = 1$, β_{IV} converges to $\beta\frac{\sigma_x^2}{\sigma_x^2+\sigma_{jk}}$, where σ_x^2 is var(x^*), implying upward bias if σ_{jk}<0 or downward bias if σ_{jk}>0. When $\rho\neq 1$, β_{IV} converges to $\beta\frac{\rho_j\sigma_x^2}{\rho_k\rho_j\sigma_x^2+\sigma_{jk}}$, with a more complicated inconsistency factor.

[22]For example, it is plausible that an idiosyncratic shock may affect father's income and occupation, inducing correlation among the measurement errors. Allowing for unrestricted correlation among the measurement errors thus permits error structures that contain a common factor, so that shocks may affect all observable measures for an individual.

[23]Specifically, the LW estimate achieves a greatest lower bound among a class of estimators, but other estimates can simply be mapped into this class for comparing magnitudes.

[24]The approach assumes that $cov(u_j, \epsilon) = 0$, although very small deviations from this will not substantially alter the results.

intergenerational income regressions as the point of departure, the natural normalisation to adopt is to the scale of father's income. With this normalisation, the equation for the remaining ρ_j can be shown to be:

$$\rho_j = \frac{cov(y_{it+1}, y_{jit})}{cov(y_{it+1}, y_{1it})} \tag{8}$$

This ratio can be estimated directly using IV estimation, instrumenting for y_{1it} (father's income) using y_{it+1} (son's income), with y_{jit} (the measure we are estimating ρ_j for) as the dependent variable. LW show that an auxiliary ordinary least squares regression of y_{it+1} on the measures y_{1it}, y_{2it}, $\ldots$, y_{jit}, produces the vector of coefficient estimates, $\hat{\phi}$, which provides information on the noisiness of the measures and on the conditional covariance of each measure with y_{it+1} (conditional on the other measures). Then, these coefficient estimates, $\hat{\phi}$, combined with the estimates, $\hat{\rho}$, form an optimal linear combination of the information from the j measures.[25] This optimal linear combination provides a greatest lower bound on β. Explicitly, the LW estimator is[26]:

$$\beta_{LW} = \hat{\rho}_1\hat{\phi}_1 + \hat{\rho}_2\hat{\phi}_2 + \cdots + \hat{\rho}_j\hat{\phi}_j \tag{9}$$

To control for other covariates (namely the quadratics in father's and son's age), these covariates are included in the auxiliary regression of son's earnings, y_{it+1}, on father's status measures, y_{1it}, y_{2it}, $\ldots$, y_{jit}, (to obtain ϕ) as well as in the IV estimations of the ρ_j.[27] Standard errors for the β_{LW} estimates are bootstrapped with 1,000 repetitions, using a block/panel bootstrap to account for clustering within family.

[25] Linearity is adopted throughout this discussion and is relied upon for the LW approach, but this is a reasonable approximation for the measures considered here and the hypothesis being examined.

[26] Note that each element of β_{LW} (i.e., $\rho_j\phi_j$) can be considered as a product of ratios $\frac{cov(y_{it+1},y_{jit})}{cov(y_{it+1},y_{1it})}\frac{cov(\tilde{y}_{it+1},\tilde{y}_{jit})}{var(\tilde{y}_{jit})}$, where $\frac{cov(\tilde{y}_{it+1},\tilde{y}_{jit})}{var(\tilde{y}_{jit})}$ is conditional on the other measures in y_t, so the estimated $\hat{\beta}_{LW}$ will be monotonically increasing in magnitude as measures are added only in the case where the conditional covariance has the same sign as the unconditional covariance.

[27] This implementation strategy is theoretically (and numerically) equivalent to that suggested in Lubotsky and Wittenberg (2006)—to first regress each measure and the dependent variable on the other covariates and use these residualised variables for estimation of ρ and ϕ.

4 Results

4.1 Main results

First, I establish a baseline estimate using the traditional approach with this PSID sample of 415 father-son pairs. Next, I explore the sensitivity of this estimate to including other measures of status in the regression. Panel A of Table 3 provides the results from this series of ordinary least squares (OLS) regressions. The baseline estimate of the intergenerational coefficient (IGC) is 0.439, using the traditional approach of regressing son's log earnings on the five-year average of father's log earnings (hence this can also be interpreted as an income elasticity). As expected, this is in the range (0.4–0.6) of traditional IGC estimates for the U.S. (Solon, 1999; Black and Devereux, 2011). Moving along columns 2-4 of Table 3, I present the OLS results from the augmented models. Adding education to the model, the coefficient on father's earnings falls slightly to 0.398, but the coefficient on education is essentially zero. Similarly, when I add occupation categories instead of education, the coefficients on these occupation category indicators are not jointly significant (F=0.54, p-value=0.779); in this case, however, the coefficient on earnings rises slightly to 0.480. When education and occupation are both incorporated, the coefficient on earnings is similar to the baseline estimate. Again, neither the coefficient on education nor the coefficients on the occupational category indicators (F=0.66, p-value=0.682) are significant.[28]

The next panel in Table 3 shows the results from an IV approach, which is commonly used to address classical measurement error. With two 'noisy' measures of status (earnings and education)

[28] Similar to my results, other studies also find that when the variable used for the parent is the same as that used for the offspring in the dependent variable, then additional variables for the parents do not have practically or statistically significant coefficients. Sewell and Hauser (1975, p.86) find this result in analysis based on the Wisconsin Longitudinal Study. With son's earnings as the dependent variable, they note that the coefficients on father's education and occupation are not statistically significant after conditioning on father's income. Using the PSID, Corcoran et al. (1992) also use son's earnings as the dependent variable and similarly find that after accounting for parental income, the coefficients for several other family or community background characteristics are not practically or statistically significant. Duncan et al. (2005) find similar results for intergenerational associations for 17 outcome measures (traits and behaviors) in the National Longitudinal Survey of Youth (NLSY). After accounting for the same measure for parents, the coefficients on the other trait or behavioral measures are not statistically significant in 84 percent of the 272 cases. Further, two very recent studies find this result using large administrative datasets: Boserup, Kopczuk, and Kreiner (2014) estimate the wealth elasticity in Denmark, and upon adding parental and child income find that these coefficients are not practically significant; Nybom and Vosters (2015) perform analyses analogous to those in this paper using Swedish administrative data and show that, with son's income as the dependent variable, after conditioning on father's income the coefficient on father's education is not practically or statistically significant.

I use education to instrument for earnings.[29] The estimated IGC is 0.497, which still falls in the range of traditional estimates for the U.S. and does not indicate substantial attenuation bias in the baseline estimate.

Finally, in Panel C, I present the estimates of the intergenerational persistence coefficient obtained using the LW approach to minimise the attenuation bias from using multiple noisy measures of status. All of the IGC estimates themselves are statistically significant, so I focus the discussion on changes in the estimates across specifications. The first estimate is simply the OLS estimate (0.439), as this is a special case of the LW approach when one uses a single measure. Adding father's education as an additional measure of status produces only a slight increase in the estimated IGC to 0.445. When occupation information is added instead of education, the IGC estimate is larger, at 0.465. And, when both education and occupation measures are simultaneously included, the IGC estimate increases slightly to 0.473, but again there is not a substantial increase in the estimated persistence.[30] Note that the OLS coefficient estimates presented in Panel A are identical to the auxiliary coefficient estimates, $\hat{\phi}_j$, used in the LW approach. Given the lack of practical or statistical significance of these estimates discussed above, it is unsurprising that we do not see large changes in the LW estimates of the intergenerational correlation. Attempting to incorporate additional information on social status causes the IGC to fluctuate some, but all estimates remain in the range of prior estimates for the U.S. Figure 1 shows that even when considering the precision of the estimates and looking at the 95 percent confidence intervals (the bars) around the estimates (the dots), neither indicate IGC estimates increasing to the hypothesised underlying persistence rate of 0.75. The plots show the estimates and confidence intervals for each specification listed in Table 3, beginning with the baseline estimate, then adding education, occupation, and both. The upper bounds on the confidence intervals are, respectively, 0.585, 0.585 0.622, and 0.629, still falling short of the hypothesised persistence rate. The precision of these estimates is hampered by the PSID sample size, but Nybom and Vosters (2015) find strikingly similar—and more precise—results

[29]As noted above, instrumenting in this fashion produces an IV estimate that converges to β/ρ_1 where ρ_1 is the coefficient in the earnings measurement equation (and assuming the measurement errors are uncorrelated), thus enabling the comparability to our LW estimate based on latent status set to the scale of father's income.

[30]When a more flexible approach is taken using the five annual earnings years as separate variables as well as separate education category variables (high school graduate, some college, four-year degree, at least some graduate school), the LW estimate of intergenerational persistence is still quite similar at 0.485 but less precise with a standard error of 0.095.

for Sweden, with similarly small increases in persistence estimates, even after using more detailed measures and incorporating analogous measures for mothers.

4.2 Robustness checks

In the main analysis, I focus on adding measures of status for fathers, but do not directly address son's latent status. As discussed in the empirical approach section, this should not substantially alter the results. However, I still perform a sensitivity check in which son's latent status is explicitly addressed. I apply the weights determined by the LW approach for father's latent status to the measures for both fathers and sons, creating index measures of status for each generation. Then I regress the composite measure for sons on the composite measure for fathers. This results in an IGC estimate of 0.433 with a (bootstrapped) standard error of 0.071, which is similar to, albeit slightly smaller than, the main LW estimates reported in Table 3.

My LW results are also robust to adjusting several of the sample restrictions, as shown in Table 4. The first row of results, with the estimates in bold and standard errors in italics underneath, simply provides the main results from Panel C of Table 3 for comparison. Allowing sons who are 25-29 years old at their 1991 earnings measure to also be included in the sample, the sample grows to 582 father-son pairs (with sons aged 25-40 years old). The IGC estimates of 0.402–0.464 are slightly smaller relative to the main results, consistent with the life-cycle effects literature (Haider and Solon, 2006; Nybom and Stuhler, *forthcoming*), but the pattern of minimal increases remains unchanged as additional measures of status are included. The same pattern is revealed when instead of adjusting the restrictions on son's age, I do so for father's age, limiting the fathers to those aged 30-50 in 1968 and obtaining IGC estimates ranging 0.457–0.494. Incorporating both of these sample adjustments at the same time also produces the same pattern, as expected, which is shown in the next row of results with estimates ranging 0.420–0.452. Returning to the original sample restrictions, except now including mother-son pairs from female-headed 1968 households (so single mothers) in the sample, the IGC estimates are smaller in magnitude (0.360–0.410) but still follow the same pattern as more measures of status are added. Finally, the last row of estimates in Table 4 presents results from changing the functional form of the father's earnings measure from the *average of log earnings* for 1967–71 to the *log of average earnings*. These results yet again exhibit the same pattern as the main results, with IGC estimates ranging 0.463–0.495.

5 Conclusions

Several recent studies by Gregory Clark and coauthors have examined intergenerational mobility using a new method based on surnames and newly developed datasets, finding higher persistence rates (i.e., lower mobility) than previously estimated (e.g., Clark, 2014; Clark and Cummins, 2015). In these studies, the hypotheses presented to explain the discrepancy use a simple measurement error argument that is consistent with the proclaimed higher persistence rate of approximately 0.75 from surname methods and the smaller estimates from traditional studies. I am the first to empirically test the proposition that prior estimates are attenuated from focusing on a single measure such as income and should rise when additional information is incorporated.

I use Lubotsky and Wittenberg's (2006) approach designed for scenarios such as this, where multiple measures of a latent variable (i.e., status) are available, but the measurement errors are likely correlated. The method combines the information from available measures of the latent variable in a way that produces a single persistence estimate with the least attenuation bias. I aggregate information from income, education, and occupation—three recommended measures of father's social status—using the LW method, yet I see no indication of the persistence rates approaching 0.75 as the additional measures are added. There are small increases in the persistence estimates as additional measures are incorporated, but these changes are not meaningful in a statistically significant or practical sense. In fact, all of the estimates presented in the main results, as well as in robustness checks, range from 0.360 to 0.491, quite similar to the prior estimates for the U.S. The pattern of small increases with additional measures is robust to adjusting sample restrictions as well as measure definitions. And, although my sample size is not conducive to assessing the statistical significance of these small changes in the point estimates, the sample I use facilitates relevant comparisons to prior literature. I am able to obtain a baseline estimate analogous to the prior studies about which the attenuation bias claims are made, which is an appropriate starting point for then incorporating information from other measures. I find no evidence that adding information from other status measures produces estimates that are converging to a substantially greater level of intergenerational persistence.

My findings reject Clark's measurement error interpretation of his results relative to those from prior literature, but they do not shed light on why his estimates based on surnames are higher than traditional estimates. Averaging over surnames does not always produce higher persistence estimates, as shown with U.S. income tax data in Appendix D of Chetty et al. (2014). Further work is needed to gain a more nuanced understanding of discrepancies between Clark's estimates using the surname-average method and traditional methods, and what each method might be identifying. As noted by Solon (2015) and Chetty et al. (2014), the traditional approach may be correctly identifying individual-level mobility, while the surnames method may be identifying group-level mobility for these particular groups of surnames. This is further developed in a recent exposition by Torche and Corvalan (2015), which shows that estimating surname-level regressions captures between-group persistence in average outcomes for the particularly 'elite' or 'underclass' surnames chosen, rather than Clark's interpretation of using group averages to eradicate measurement error and reveal individual-level mobility.

References

Becker, G. and Tomes, N. (1976). 'Child endowments, and the quantity and quality of children', *Journal of Political Economy*, vol. 84(4)2, pp. S143-S162.

Becker, G. and Tomes, N. (1979). 'An equilibrium theory of the distribution of income and intergenerational mobility', *Journal of Political Economy*, vol. 87, pp. 1153-189.

Björklund, A., and Salvanes, K. G. (2011). 'Education and family background: Mechanisms and policies', in E. Hanushek, S. Machin, and L. Woessmann (eds.), *Handbook of the Economics of Education*, vol. 3(3), pp. 201-247.

Black, S. E. and Devereux, P. J. (2011). 'Recent developments in intergenerational mobility', in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4, pp. 1487-1541, Amsterdam: Elsevier.

Braun, S., and Stuhler, J. (2015). 'The transmission of inequality across multiple generations: Testing recent theories with evidence from Germany'. Mimeo.

Boserup, S.H., Kopczuk, W., and Kreiner, C.T. (2014). 'Intergenerational wealth mobility: Evidence from Danish wealth records of three generations.' Working paper, October 2014.

Chetty, R., Hendren, N., Kline, P. and Saez, E. (2014). 'Where is the Land of Opportunity? The geography of intergenerational mobility in the United States', *Quarterly Journal of Economics*, vol. 129(4), pp. 1553-1623.

Clark, G. (2014). *The son also rises: surnames and the history of social mobility.* Princeton University Press.

Clark, G. and Cummins, N. (2015). 'Intergenerational wealth mobility in England, 1858-2012. Surnames and social mobility', *Economic Journal*, vol. 125, pp. 61-85.

Clark, G., Cummins, N., Hao, Y. and Vidal, D.D. (2015). 'Surnames: a new source for the history of social mobility', *Explorations in Economic History*, vol. 55, pp. 3-24.

Corcoran, M., Gordon, R., Laren, D., and Solon, G. (1992). 'The association between men's economic status and their family and community origins', *Journal of Human Resources*, vol. 27(4), pp. 575-601.

Duncan, G., Kalil, A., Mayer, S. E., Tepper, R., and Payne, M. R. (2005). 'The apple does not fall far from the tree', in Samuel Bowles, Herbert Gintis and Melissa Osborne Groves (eds.), *Unequal Chances: Family Background and Economic Success*, Princeton University Press., pp. 23-79.

Güell, M., Rodríguez Mora, J. V. and Telmer, C. (2014). 'The informational content of surnames, the evolution of intergenerational mobility and assortative mating', *Review of Economic Studies*, Advance Access published online December 10, 2014, doi:10.1093/restud/rdu041.

Haider, S. J. and Solon, G. (2006). 'Life-cycle variation in the association between current and lifetime earnings', *American Economic Review*, vol. 96(4), pp. 1308-1320.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2007). 'The inheritance of educational inequality: International comparisons and fifty-year trends', *The BE Journal of Economic Analysis and Policy*, vol. 7(2).

Lee, C. I. and Solon, G. (2009). 'Trends in intergenerational income mobility', *The Review of Economics and Statistics*, vol. 91(4), pp. 766-772.

Long, J., and Ferrie, J. (2007). 'The path to convergence: Intergenerational occupational mobility in Britain and the U.S. in three eras,' *The Economic Journal*, vol. 117 (519), pp. C61-C71.

Long, J., and Ferrie, J. (2013). 'Intergenerational occupational mobility in Great Britain and the United States since 1850', *The American Economic Review*, vol. 103(4), pp. 1109-1137.

Lubotsky, D. and Wittenberg, M. (2006). 'Interpretation of regressions with multiple proxies', *The Review of Economics and Statistics*, vol. 88(3), pp. 549-562.

Lucas, R. E., and Kerr, S. P. (2013). 'Intergenerational income immobility in Finland: Contrasting roles for parental earnings and family income', *Journal of Population Economics*, vol. 26(3), pp. 1057-1094.

Mazumder, B. (2005). ‘Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data’, *The Review of Economics and Statistics*, vol. 87(2), pp. 235-255.

Nybom, M. and Stuhler, J. (*forthcoming*).‘Heterogeneous income profiles and life-cycle bias in intergenerational mobility estimation’, *Journal of Human Resources*.

Nybom, M. and Vosters, K. (2015). ‘Intergenerational persistence in latent socioeconomic status: Evidence from Sweden’. SOFI Working Paper 3/2015.

Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Institute for Social Research, University of Michigan, Ann Arbor, MI (accessed Dec 2013).

Sewell, W. H., and Hauser, R. M. (1975). *Education, occupation, and earnings. Achievement in the early career.* New York: Academic Press.

Solon, G. (1992). ‘Intergenerational income mobility in the United States’, *The American Economic Review*, vol. 82(3), pp. 393-408.

Solon, G. (1999). ‘Intergenerational mobility in the labor market’, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, pp. 1761-1800, Amsterdam: North-Holland.

Solon, G. (2015). ‘What do we know so far about multigenerational mobility?’ NBER Working Paper No. w21053.

Torche, F. and Corvalan, A. (2015). ‘Estimating intergenerational mobility with grouped data: a critique of Clark’s *The Son Also Rises*’. Working Paper 15-22, NYU Population Center.

Zimmerman, D. J. (1992). ‘Regression toward mediocrity in economic stature’, *The American Economic Review*, vol. 82(3), pp. 409-429.

Table 1

Summary statistics

	Mean	*Std. Dev.*	*Min*	*Max*
Race - black	0.05	0.22	0	1
Sons' age in 1991	34.92	3.14	30	40
Sons' 1991 individual earnings	35,695	26,251	300	335,000
Fathers' age in 1967	40.47	6.81	27	67
Fathers' Individual earnings				
Annual earnings 1967	39,684	24,409	1,101	244,671
Log annual earnings 1967	10.43	0.60	7.00	12.41
5-year-avg of log earnings, 1967-71	10.46	0.59	7.79	12.65
Fathers' Educational attainment				
Less than HS graduate	0.33	0.47	0	1
High school graduate	0.32	0.47	0	1
Some college	0.11	0.31	0	1
Bachelor's degree	0.14	0.34	0	1
At least some graduate school	0.11	0.31	0	1
Fathers' 1969 Occupation categories				
1 - Professional, technical	0.23	0.42	0	1
2 - Manager/businessmen	0.14	0.35	0	1
3 - Clerical, sales	0.09	0.29	0	1
4 - Craftsman, foreman	0.23	0.42	0	1
5 - Operatives	0.17	0.38	0	1
6 - Laborers, service, farmers	0.12	0.32	0	1
7 - Not currently employed/missing	0.02	0.15	0	1

*Notes***.** The sample includes 415 sons and 293 fathers. All earnings are expressed in 1991 dollars.

Table 2

Occupation categories

	Earnings in 1969		Educational attainment		
Occupation Category	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*N*
1 Professional, technical	61,382	40,129	15.67	1.76	66
2 Manager/businessmen	54,983	41,871	12.83	2.73	42
3 Clerical, sales	38,379	10,920	12.96	1.89	27
4 Craftsman, foreman	38,212	15,203	10.79	2.62	67
5 Operatives	30,044	11,108	9.76	2.53	50
6 Laborers, service, farmers	20,614	10,468	9.97	2.55	34
7 Not employed or missing	38,379	31,544	10.00	3.61	7
Overall	42,419	30,209	12.09	3.27	293

Notes. The sample includes 293 fathers. All earnings are expressed in 1991 dollars.

Table 3

OLS, IV, and LW results

Fathers' noisy measures of status		[1] Earnings	[2] Earnings, education	[3] Earnings, occupation	[4] Earnings, education, occupation
Panel A: OLS results					
Five-year average of log earnings: 1967-71		0.439	0.398	0.480	0.438
		0.075	*0.098*	*0.100*	*0.120*
Educational attainment			0.010		0.016
			0.013		*0.016*
Occupation categories					
1 - Professional, technical				0.002	-0.077
				0.228	*0.236*
2 - Manager/businessmen				-0.029	-0.064
				0.233	*0.222*
3 - Clerical, sales				0.001	-0.051
				0.256	*0.258*
4 - Craftsman, foreman				0.066	0.052
				0.233	*0.218*
5 - Operatives				-0.027	-0.032
				0.229	*0.211*
6 - Laborers, service, farmers				0.181	0.152
				0.254	*0.244*
Panel B: IV results (education to IV for 5-yr-avg earn)					
First stage	0.105				
	0.006				
Second stage		0.497			
		0.090			
Panel C: LW estimates of IGC					
		0.439	0.445	0.465	0.473
		0.075	*0.072*	*0.080*	*0.080*
N		415	415	415	415

Notes. All specifications use log of son's 1991 earnings as the dependent variable and include as controls a quadratic in son's earnings and a quadratic in father's age (the average age during the five years of earnings observations). The omitted occupation category is '7 - Not employed or missing'. The sample size for all estimations is 415 father-son pairs, from 293 families. LS & IV standard errors are clustered by family. LW standard errors are computed using a block bootstrap to account for within family correlation (1,000 repetitions).

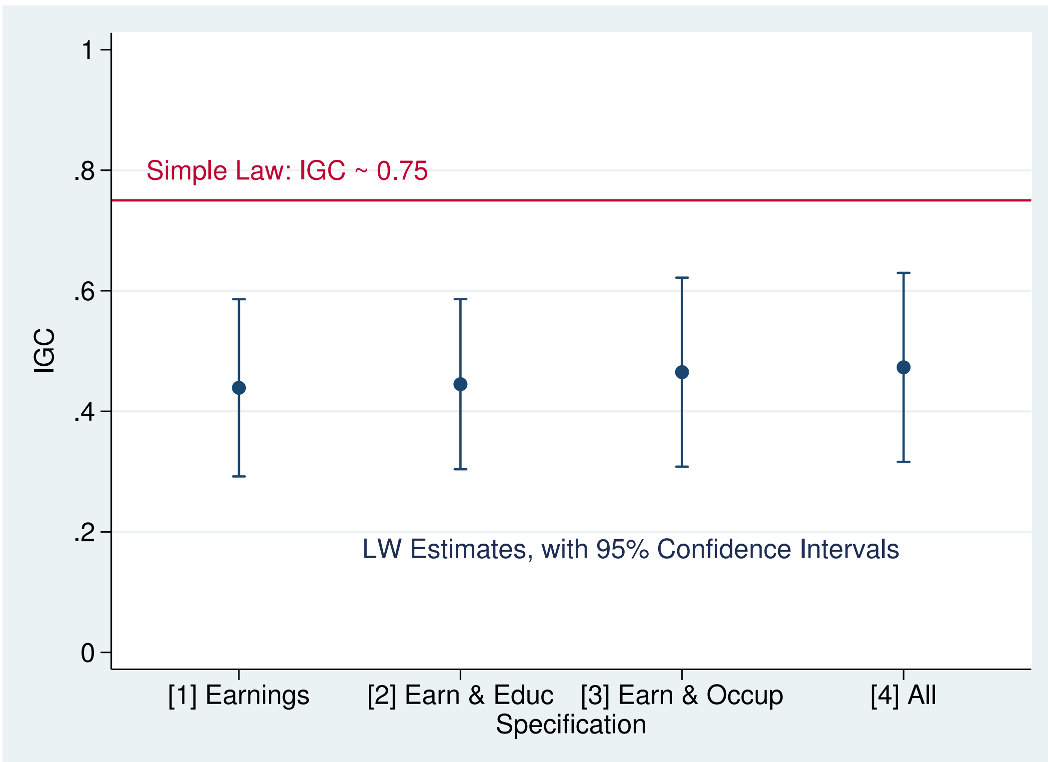
Table 4

Robustness of LW results

	N	[1] Earnings	[2] Earnings, education	[3] Earnings, occupation	[4] Earnings, education, occupation
Main results	415	**0.439**	**0.445**	**0.465**	**0.473**
		0.075	*0.072*	*0.080*	*0.080*
Adjusting sample exclusions:					
Son's age 25-40	582	**0.402**	**0.422**	**0.446**	**0.464**
		0.065	*0.061*	*0.075*	*0.077*
Father's age 30-50	380	**0.457**	**0.463**	**0.484**	**0.494**
		0.083	*0.083*	*0.090*	*0.089*
Son's age 25-40 and Father's age 30-50	483	**0.420**	**0.426**	**0.444**	**0.452**
		0.076	*0.073*	*0.083*	*0.082*
Include 1968 female-headed households	444	**0.360**	**0.375**	**0.392**	**0.410**
		0.072	*0.064*	*0.072*	*0.069*
Adjusting earnings measure:					
Log of father's 5-yr avg of annual earnings 1967-71	415	**0.463**	**0.466**	**0.490**	**0.495**
		0.075	*0.073*	*0.081*	*0.081*
Father's status measures					
Earnings (5-yr-avg)		x	x	x	x
Educational attainment			x		x
Occupational categories				x	x

*Notes***.** The dependent variable is log of son's 1991 earnings, and the measure of father's earnings is the 5-year average of log earnings from 1967-71. All specifications include as controls a quadratic in son's earnings and a quadratic in father's age (the average age during the five years of earnings observations). The omitted occupation category is '7 - Not employed or missing'. Standard errors are computed using a block bootstrap to account for within family correlation (1,000 repetitions).

Figure 1: *LW results*



Notes. The sample includes 415 fathers and 293 fathers. Standard errors are computed using a block bootstrap to account for within family correlation (1,000 repetitions).