Survey Research Methods (2016)
Vol. 10, No. 2, pp. 85-102
doi:10.18148/srm/2016.v10i2.6277

ISSN 1864-3361
http://www.surveymethods.org

Two-Sample Two-Stage Least Squares (TSTSLS) estimates of earnings mobility: how consistent are they?

John Jerrim
Institute of Education
University College
London, U.K.

Álvaro Choi
University of Barcelona and
Barcelona Institute of Economics
Barcelona, Spain

Rosa Simancas
University of Extremadura
Spain

Academics and policymakers have shown great interest in cross-national comparisons of intergenerational earnings mobility. However, producing consistent and comparable estimates of earnings mobility is not a trivial task. In most countries researchers are unable to observe earnings information for two generations. They are thus forced to rely upon imputed data from different surveys instead. This paper builds upon previous work by considering the consistency of the intergenerational correlation (ρ) as well as the elasticity (β), how this changes when using a range of different instrumental (imputer) variables, and highlighting an important but infrequently discussed measurement issue. Our key finding is that, while TSTSLS estimates of β and ρ are both likely to be inconsistent, the magnitude of this problem is much greater for the former than it is for the latter. We conclude by offering advice on estimating earnings mobility using this methodology.

Keywords: Earnings mobility; two sample two stage least squares; imputation

1 Introduction

Academics have shown great interest in intergenerational mobility – the strength of the association between individuals' social origin and social destination. Historically this work was the realm of sociologists (see Erikson & Goldthorpe, 1992, for a review), but economists have added much to this debate over the last twenty years, particularly through their examinations of the link between the earnings (or incomes) of fathers and sons. However, due to data limitations, obtaining consistent estimates of earnings mobility remains a non-trivial task (Black & Devereux, 2011; Blanden, 2013; Solon, 1992). The contribution of this paper is to present new evidence on the consistency of Two-Sample Two-Stage Least Squares (TSTSLS) estimates of earnings mobility; a methodology now widely applied in this literature (Appendix A reviews almost 30 papers where it has been used). Indeed, TSTSLS has proven to be the only way to estimate earnings mobility in a number of countries, including Australia, France, Italy, Spain, Switzerland, Japan, China and South Africa. It has therefore played an important role in cross-national comparisons of earnings mobility; of the 20 countries included in Corak (2012), TSTSLS has been used in more than half.

Contact information: John Jerrim, UCL Institute of Education, 20 Bedford Way, Bloomsbury, London, United Kingdom (email: j.jerrim@ucl.ac.uk)

Yet, despite the important work of Björklund and Jäntti (1997) and Nicoletti and Ermisch (2008), more needs to be known about the consistency of TSTSLS estimates of earnings mobility. We therefore build upon the aforementioned authors' work by extending their framework from the intergenerational elasticity (β) to the intergenerational correlation (ρ) quantifying the inconsistency of TSTSLS estimates when using a range of different instrumental (imputer) variables, and considering a potentially important (yet little discussed) measurement issue.

The TSTSLS estimation procedure can be summarised as follows. Ideally, earnings mobility would be estimated via the following Ordinary Least Squares (OLS) regression model:

$$Y_{\text{True}} = \alpha + \beta X_{\text{True}} + u \quad , \tag{1}$$

where Y_{True} are the (Log) permanent earnings of sons, and X_{True} are the (Log) permanent earnings of fathers.

Two different measures of earnings mobility would then typically be produced. First, the intergenerational earnings elasticity β_{OLS} with

$$\beta_{\text{OLS}} = \frac{\sigma_{X,Y}}{\sigma_X^2} \quad , \tag{2}$$

with $\sigma_{X,Y}$ being the covariance between father's and son's permanent earnings, and σ_X^2 being the variance of father's earnings. Second, the intergenerational correlation ρ_{OLS} with

$$\rho_{\text{OLS}} = \frac{\sigma_{X,Y}}{\sigma_X^2} \cdot \frac{\sigma_X}{\sigma_Y} = \frac{\sigma_{X,Y}}{\sigma_X \cdot \sigma_Y} \quad , \tag{3}$$

with σ_X being the standard deviation of father's earnings, and σ_Y being the standard deviation of son's earnings.

The measure of X_{True} preferred in the literature is a time-average of father's annual earnings across several years (X_{AVG}).[1] However, in many countries, earnings data cannot be linked across generations – i.e. there is no dataset where both father's and son's earnings can be observed. The TSTSLS approach attempts to overcome this problem via imputation – predictions of father's earnings are made based upon other observable characteristics (e.g. their occupation and education level). Equation 1 is then estimated using these predictions of father's earnings ($\widehat{X}$ instead of a measure that has been directly observed (e.g. X_{AVG}). This is often described as an instrumental variable technique in the earnings mobility literature (e.g. Lefranc & Trannoy, 2005; Nuñez & Miranda, 2011), though it can alternatively be viewed as a cold-deck imputation procedure (Nicoletti & Ermisch, 2008) or a "generated regressor" approach (Murphy & Topel, 1985; Wooldridge, 2002, p. 115; Inoue & Solon, 2010).

Solon (1992), Björklund and Jäntti (1997) and Nicoletti and Ermisch (2008) consider the properties of TSTSLS estimates of the intergenerational elasticity (β_{TSTSLS}). They show that consistent estimates can be obtained if either:

- The instrumental (imputer) variables have no direct effect upon son's earnings
- The R^2 of the equation used to predict father's earnings equals one

Yet, as father's education and occupation are the instruments (imputer variables) usually available, it is widely recognised that neither of these conditions hold. (Father's education and social class are likely to independently influence offspring's earnings, while also not being perfect predictors of father's permanent earnings). It is thus often stated that β_{TSTSLS} will be upward inconsistent as a result (Blanden, 2013).

The key issue thus becomes the *magnitude* of this upward inconsistency. Is it small enough to be safely ignored, or is it so large that TSTSLS estimates of earnings mobility become problematic? Likewise, if more detail is added to the model predicting father's earnings, does this significantly reduce the upward inconsistency? Unfortunately, little is currently known about these important issues. Indeed, the only study to quantify the inconsistency of β_{TSTSLS} is Björklund and Jäntti (1997). For one particular imputation model, containing a specific set of predictor variables, they find upward inconsistency of around 30 percent.

We contribute to this evidence base in multiple ways. First, the framework of Björklund and Jäntti (1997) and Nicoletti and Ermisch (2008) is extended from the intergenerational elasticity to the intergenerational correlation (ρ_{TSTSLS}). We use this to explain why ρ_{TSTSLS} is *downward* inconsistent in our empirical analysis (i.e. in the opposite direction of the inconsistency of β_{TSTSLS}). Second, new evidence is provided on the inconsistency of β_{TSTSLS} and ρ_{TSTSLS} using a range of different imputer variables, and thus the extent to which this problem can be reduced through the use of a more detailed first-stage prediction model. Third, we divide β_{TSTSLS} and ρ_{TSTSLS} into components to demonstrate what is driving their inconsistency, and show how this changes when different prediction models are specified. Finally, we note how most studies make predictions of father's current earnings (X_{Single}), whereas permanent earnings (X_{True}) is the actual unobserved variable of interest. We argue that, in this situation, more general expressions for the inconsistency of β_{TSTSLS} and ρ_{TSTSLS} are needed. Our empirical analysis then illustrates how conventional wisdom (e.g. β_{TSTSLS} always being upward inconsistent) no longer necessarily holds.

The paper now proceeds as follows. Properties of TSTSLS earnings mobility estimates are reviewed in section 2. This is followed by an overview of the Panel Survey of Income Dynamics (PSID) dataset and our empirical methodology in section 3. Results are presented in section 4, and conclusions in section 5.

2 TSTSLS estimates of earnings mobility

Our starting point is the framework of Nicoletti and Ermisch (2008). As noted in the introduction, the model of interest is:

$$Y_{\text{True}} = \beta X_{\text{True}} + \mu \quad , \tag{4}$$

with Y_{True} being Log son's permanent earnings, and X_{True} being Log father's permanent earnings.

X_{True} is unobserved in the "main" dataset, but it does contain additional characteristics (Z), such as father's education and occupation, likely to be associated with X_{True}.

Now say a second "auxiliary" sample (i) contains a measure of respondents' permanent earnings[2] (ii) is drawn from

[1] Although five consecutive years of father's earnings is often used (Björklund & Chadwick, 2003; Corak & Heisz, 1999; Hussain, Munk, & Bonke, 2009; Solon, 1992; Vogel, 2008), more than ten may be needed if there is substantial auto-correlation in the transitory component of earnings over time (Björklund & Jäntti, 2009; B. Mazumder, 2005). For recent evidence see (Nybom & Stuhler, 2016).

[2] Time-average earnings would be the preferred measure within

the same population and (iii) contains the same Z variables. The following OLS regression model can be estimated:

$$X_{\text{True}} = \delta Z + v \tag{5}$$

where Z are the instrumental (imputer) variables.

And then used to predict log permanent father's earnings:

$$\widehat{X} = \widehat{\delta} \cdot Z \tag{6}$$

where $\widehat{X}$ is the predicted log father's permanent earnings, and $\widehat{\delta}$ are the estimated regression coefficients from the first-stage prediction model.

Hence (7) can now be estimated rather than (1):

$$Y_{\text{True}} = \beta\widehat{X} + u \tag{7}$$

Estimates of β_{TSTSLS} and ρ_{TSTSLS} then follow from equations 2 and 3 (substituting $\widehat{X}$ for X).

The two most commonly used Z variables are father's education and occupation (see Appendix A). However, both are likely to directly influence son's earnings (i.e. they are likely to be endogenous). Consequently, son's log earnings will actually be given by:

$$Y_{\text{True}} = \lambda_1 X_{\text{True}} + \lambda_2\widehat{X} + u \tag{8}$$

With λ_1 being the direct impact of the father's *actual* permanent earnings on son's earnings and λ_2 the effect of father's *predicted* earnings on son's earnings. (From this point forward, we drop the "True" subscript for notational convenience). Building upon the work of Solon (1992), Björklund and Jäntti (1997) argue that β_{TSTSLS} thus converges in probability to:

$$\begin{aligned} \text{plim}\,\beta_{\text{TSTSLS}} &= \lambda_1 + \lambda_2 \frac{\sigma_{\widehat{X}}}{\eta\sigma_X} \\ &= \beta + \lambda_2\sigma_{\widehat{X}}\frac{(1-\eta^2)}{\eta\sigma_X} \quad , \end{aligned} \tag{9}$$

where $\sigma_{\widehat{X}}$ is the standard deviation of father's *predicted* earnings and σ_X is the standard deviation of father's *actual* long-run earnings. η is defined as

$$\eta = \frac{\sigma_{\widehat{X},X}}{\sigma_{\widehat{X}}\sigma_X}$$

where $\sigma_{\widehat{X},X}$ is the covariance between predicted and actual log father's earnings.

Under the assumption that the limit of the covariance between predicted and actual log father's earnings is asymptotically equal to the covariance between predicted father's earnings and itself:

$$\sigma_{\widehat{X},X} = \sigma_{\widehat{X},\widehat{X}} \tag{10}$$

η becomes:[3]

$$\eta = \frac{\sigma_{\widehat{X},X}}{\sigma_{\widehat{X}}\sigma_X} = \frac{\sigma_{\widehat{X},\widehat{X}}}{\sigma_{\widehat{X}}\sigma_X} = \frac{{\sigma_{\widehat{X}}}^2}{\sigma_{\widehat{X}}\sigma_X} = \frac{\sigma_{\widehat{X}}}{\sigma_X} = R \tag{11}$$

where R is the square root of the variance explained (R^2) in the first-stage prediction model (i.e. of equation 5).

The probability limit of β_{TSTSLS} then becomes:

$$\begin{aligned} \beta_{\text{TSTSLS}} &= \beta + \lambda_2\sigma_{\widehat{X}}\frac{\left(1-\frac{{\sigma_{\widehat{X}}}^2}{{\sigma_X}^2}\right)}{\frac{\sigma_{\widehat{X}}}{\sigma_X}\sigma_X} \\ &= \beta + \lambda_2\sigma_{\widehat{X}}\frac{\left(1-\frac{{\sigma_{\widehat{X}}}^2}{{\sigma_X}^2}\right)}{\sigma_{\widehat{X}}} \\ &= \beta + \lambda_2\left(1-\frac{{\sigma_{\widehat{X}}}^2}{{\sigma_X}^2}\right) \\ &= \beta + \lambda_2(1-R^2) \end{aligned} \tag{12}$$

With the inconsistency of β_{TSTSLS} therefore:

$$\lambda_2(1-R^2) \tag{13}$$

There are a number of important points to note about (11), (12) and (13). First, as $0 \leq R^2 \leq 1$, the variance of father's predicted earnings must be less than or equal to the variance of actual father's earnings:

$$0 \leq \sigma^2_{\widehat{X}} \leq \sigma^2_X$$

Second, if the variance of father's predicted earnings (${\sigma_{\widehat{X}}}^2$) were equal to the variance of father's actual earnings (${\sigma_X}^2$), then $R^2 = 1$ and the inconsistency of β_{TSTSLS} reduces to zero. Hence, in this framework, the inconsistency of β_{TSTSLS} is driven by incorrect estimation of the variability in father's predicted earnings. Third, if the Z variables are indeed exogenous with respect to son's earnings, then λ_2 equals 0, and β_{TSTSLS} is consistent. However, if parental education and occupation are the Z chosen, λ_2 will almost certainly be positive ($\lambda_2 > 0$). Thus, under the reasonable assumption that $\lambda_2 > 0$, and given $R^2 \leq 1$, β_{TSTSLS} will be upwardly inconsistent. Fourth, if everything else remains unchanged, the magnitude of this upward inconsistency will decrease as the variance explained in the first-stage prediction equation increases. Or, to put this another way, the upward inconsistency will decrease as the variance of father's predicted earnings tends towards the variance of father's actual earnings (${\sigma_{\widehat{X}}}^2 \to {\sigma_X}^2$). Fifth, it is important to recognise, however, that including additional variables to increase the R^2 of the first-stage prediction equation may simultaneously influence λ_2. Consequently, adding a particularly endogenous Z variable could increase λ_2 to such an extent that it more

the auxiliary dataset. Unfortunately, this is rarely available, so current earnings are often used as the "first-stage" dependent variable instead. We illustrate how this influences TSTSLS estimates in section 4.

[3] The covariance between a variable and itself is asymptotically equal to the variance of that variable. Hence $\sigma_{\widehat{X},\widehat{X}}$ becomes ${\sigma_{\widehat{X}}}^2$.

than offsets the benefits of any change to the first-stage R^2. Whether adding variables to the prediction equation reduces the inconsistency of β_{TSTSLS} is therefore an (underexplored) empirical issue, representing a gap in the literature that this paper attempts to fill.

Next, we extend the framework of Björklund and Jäntti (1997) and Nicoletti and Ermisch (2008) to the intergenerational correlation (ρ_{TSTSLS}). If one could observe X_{True} and Y_{True}, ρ would simply be:

$$\rho = \beta \frac{\sigma_X}{\sigma_Y} \tag{14}$$

Replacing σ_X with $\sigma_{\widehat{X}}$, and β with β_{TSTSLS}, ρ_{TSTSLS} converges in probability to

$$\begin{aligned} \operatorname{plim} \rho_{\text{TSTSLS}} &= \beta_{\text{TSTSLS}} \frac{\sigma_{\widehat{X}}}{\sigma_Y} \\ &= \left[\beta + \lambda_2(1 - R^2)\right] \frac{\sigma_{\widehat{X}}}{\sigma_Y} \\ &= \beta \frac{\sigma_{\widehat{X}}}{\sigma_Y} + \lambda_2 \frac{\sigma_{\widehat{X}}}{\sigma_Y}(1 - R^2) \end{aligned} \tag{15}$$

The inconsistency of ρ_{TSTSLS} is then given by (15) – (14):

$$\begin{aligned} &\left[\beta \frac{\sigma_{\widehat{X}}}{\sigma_Y} + \lambda_2 \frac{\sigma_{\widehat{X}}}{\sigma_Y}(1 - R^2)\right] - \beta \frac{\sigma_X}{\sigma_Y} \\ &\qquad = \beta \left[\frac{\sigma_{\widehat{X}}}{\sigma_Y} - \frac{\sigma_X}{\sigma_Y}\right] + \lambda_2 \frac{\sigma_{\widehat{X}}}{\sigma_Y}(1 - R^2) \end{aligned} \tag{16}$$

Now define A as the left-hand side of (16) and B as the right-hand side:

$$A = \beta \left[\frac{\sigma_{\widehat{X}}}{\sigma_Y} - \frac{\sigma_X}{\sigma_Y}\right] \tag{17}$$

$$B = \lambda_2 \frac{\sigma_{\widehat{X}}}{\sigma_Y}(1 - R^2) \tag{18}$$

Under the previously stated assumption that ${\sigma_{\widehat{X}}}^2 \leq {\sigma_X}^2$, then $A \leq 0$ (i.e. this will lead to *downward* inconsistency in ρ_{TSTSLS}). In contrast, assuming that $\lambda_2 > 0$ then, as $R^2 \leq 1$, $B \geq 0$ (i.e. this will lead to *upward* inconsistency in ρ_{TSTSLS}). Therefore, unlike β_{TSTSLS}, one does not know the direction of the inconsistency in ρ_{TSTSLS}. Rather, it depends upon the relative magnitudes of A and B. This is again an empirical issue, which we provide the first evidence upon in our analysis.

The derivations presented above have all relied upon the following assumptions:

- The main and auxiliary datasets are random samples from the same population
- The Z variables are independent and identically distributed across the two datasets
- That X_{True} is the first-stage dependent variable, and it is this quantity that we wish to impute into the main dataset.

To meet these assumptions, it would be ideal for the main and auxiliary datasets to be identical (with the exception, of course, that the former does not include X_{True}). In this situation, the consistency of β_{TSTSLS} and ρ_{TSTSLS} is driven solely by the choice of imputer variables (Z) as set out above.

In reality, these assumptions may not be met. For instance, Björklund and Jäntti (1997) note it is common for respondents to report their own education and occupation (Z) in the auxiliary dataset, but for offspring's proxy reports of their father's characteristics to be available in the main dataset. The impact this has upon the consistency of β_{TSTSLS} and ρ_{TSTSLS} will depend upon the nature and extent of this measurement error. We therefore also consider this issue in our empirical analysis.

Moreover, there is the additional complication of how father's earnings are measured in the auxiliary dataset. Returning to equation (5), it has thus far been implicitly assumed that X_{True} (permanent father's earnings) is available within the auxiliary dataset. Yet, in practise, this is almost never the case. Rather, researchers typically have access to data for a cross-section of adults whose earnings are recorded for one particular year (X_{Single}). A common choice is a labour force survey, for example. Therefore the prediction model is often specified as (19) rather than (5):

$$X_{\text{Single}} = \delta Z + \gamma A + \nu \tag{19}$$

where X_{Single} is earnings in a single year for a cross-section of adults, Z are the imputation variables, and A are the age group dummy variables.

Estimates from (19) are then used to generate predictions of father's earnings in the main dataset instead of equation (6), with age set to around 40.[4]

$$\widehat{X}_{\text{Single}} = \widehat{\delta} Z + \widehat{\gamma} \cdot \text{age40} \tag{20}$$

Yet little is known about the consistency of TSTSLS estimates in such situations, where the first-stage dependent variable (X_{Single}) differs from the unobserved construct of interest (X_{True}). Indeed, this issue was not explicitly considered by Björklund and Jäntti (1997) or Nicoletti and Ermisch (2008), and should not be assumed to be an innocuous change to the framework presented above. Moreover, our reading of the literature is that almost all existing income mobility studies applying the TSTSLS estimator use X_{Single} as the first-stage dependent variable rather than X_{True}.

We illustrate this point with an example. First, suppose that X_{true} is contained within the auxiliary dataset, along with a sufficiently rich set of Z so that the first-stage R^2 equals one.

[4]We have chosen age 40 as this is the approximate point when earnings peak. However, we note recent work by Nybom and Stuhler (2016) who suggest that although "lifecycle bias is smallest when incomes are measured around midlife"' is its very difficult to "predict the ideal age of measurement".

Consequently, $\widehat{X}$ will be identical to X_{True}, thus resulting in consistent estimates of β_{TSTSLS} and ρ_{TSTSLS} (e.g. recall equation 13). Now consider the same scenario, but where X_{Single} is the first-stage dependent variable. A first-stage R^2 of one would imply that $\widehat{X} = X_{\text{Single}}$, resulting in rather different estimates of β_{TSTSLS} and ρ_{TSTSLS} (i.e. it is well established in the literature that $X_{\text{Single}} \neq X_{\text{True}}$ – see Solon, 1992 and Blanden, 2013. Specifically, the use of X_{Single} would lead to *downwardly* inconsistent estimates of β_{TSTSLS} and ρ_{TSTSLS}. This highlights how the corollaries presented within the framework above (e.g. β_{TSTSLS} always being upward inconsistent) do not necessarily hold when the first-stage variable being imputed (X_{Single}) differs from the construct actually of interest (X_{True}).

More general expressions for the inconsistency of β_{TSTSLS} and ρ_{TSTSLS} are therefore required, which hold whether either X_{Single} or X_{true} are used as the first-stage dependent variable. First, consistent estimates of β_{OLS} from equation (1) converge in probability to:

$$\text{plim}\,\beta = \frac{\sigma_{X,Y}}{\sigma^2_X} \tag{21}$$

Under TSTSLS, as X is unavailable, $\widehat{X}$ enters in its place:

$$\text{plim}\,\beta_{\text{TSTSLS}} = \frac{\sigma_{\widehat{X},Y}}{\sigma^2_{\widehat{X}}} \tag{22}$$

The inconsistency of β_{TSTSLS} is now given by (22) minus (21):

$$\frac{\sigma_{\widehat{X},Y}}{\sigma_{\widehat{X}}{}^2} - \frac{\sigma_{X,Y}}{\sigma_X{}^2} \tag{23}$$

Note that, in this more general framework, β_{TSTSLS} can be either upwards or downwards inconsistent. Indeed, the direction and magnitude of the inconsistency depends upon one's ability to correctly estimate the ratio of the covariance between father's and son's earnings ($\sigma_{X,Y}$) to the variance of father's earnings ($\sigma_X{}^2$).

Equations (24) to (26) provide analogous expressions for ρ_{TSTSLS}. If X_{true} and Y_{true} were available in the main dataset, ρ could be consistently estimated by:

$$\text{plim}\,\rho = \frac{\sigma_{X,Y}}{\sigma^2_X} \cdot \frac{\sigma_X}{\sigma_Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} \tag{24}$$

Replacing, X with $\widehat{X}$, ρ_{TSTSLS} converges in probability to:

$$\text{plim}\,\rho_{\text{TSTSLS}} = \frac{\sigma_{\widehat{X},Y}}{\sigma^2_{\widehat{X}}} \cdot \frac{\sigma_{\widehat{X}}}{\sigma_Y} = \frac{\sigma_{\widehat{X},Y}}{\sigma_{\widehat{X}} \sigma_Y} \tag{25}$$

with the inconsistency of ρ_{TSTSLS} now given by (25) minus (24):

$$\frac{\sigma_{\widehat{X},Y}}{\sigma_{\widehat{X}} \sigma_Y} - \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} \tag{26}$$

In our empirical analysis we illustrate how the inconsistency of β_{TSTSLS} and ρ_{TSTSLS} can vary depending on whether X_{Avg} (as a measure of X_{True}) or X_{Single} is used as the first-stage dependent variable.

To conclude, we note that generated regressors (e.g. $\widehat{X}$) are also subject to sampling variation. Consequently, second stage standard errors will be underestimated unless this additional uncertainty is taken into account. Murphy and Topel (1985), Wooldridge (2002) and Inoue and Solon (2010) provide formulae to make an appropriate adjustment to the estimated standard errors, while Björklund and Jäntti (1997), Inoue and Solon (2010) and Piraino (2014) suggest bootstrapping as a viable (if computer intensive) alternative. We do not dwell on this issue in this paper, and focus upon the inconsistency of TSTSLS point estimates. Nevertheless, this additional source of sampling uncertainty should also always be taken into account when applying such generated regressor techniques.

3 Data

The Panel Survey of Income Dynamics (PSID) is a nationally representative sample of US households, with the data available for download from http://simba.isr.umich.edu/default.aspx. It began in 1968, with annual follow-ups to 1997, and bi-annual interviews thereafter. Detailed information has been collected at each sweep from the household head and their partner. Offspring are tracked as they leave the initially sampled household. Consequently, the PSID contains earnings information across multiple years for both fathers and sons. Throughout our analysis we restrict the sample to include sons who were household heads aged between 30 and 60 in 2011, and who reported their earnings for the previous year. Moreover, we only include sons whose father can be identified, has reported annual earnings on at least five occasions during their prime working years (between ages 30 and 60), and where both parent and offspring reports of father's education, occupation and industry are available.

After making these restrictions, our working sample equals 1,024 observations. Approximately 80 percent of these individuals have at least 15 reports of father's annual earnings available, with 60 percent having 20 or more. A "permanent" measure of father's earnings is created by averaging across all available reports for each sample member. We call this X_{AVG}, the closest measure to X_{True} available in the PSID. All earnings data have been adjusted to 2010 prices.

As part of each PSID sweep, fathers were asked detailed questions about their educational attainment, occupation and industry (we label father's reports of these variables as Z_{FA}. Education has been recorded using the highest grade ever completed, which we have converted into eight groups. Occupation and industry have been recorded using three digit census codes. These are finely defined categories – separating occupations and industries into approximately 200 groups. We use this detailed information on father's occupation and industry (taken from the year their offspring turned age 15) as the key imputer variables (Z)[5]. At times, we also use more broadly defined "1 digit" occupation (10 categories) and industry (12 categories) groups.

Sons also reported similar information about their father's education, occupation and industry (denoted Z_{Sons}). For instance, in the 2011 sweep, sons were asked:

- How much education did your father complete?
- What was your father's usual occupation when you were growing up?
- What kind of business or industry was that in?

Information on Z is thus available both directly from fathers (Z_{Fathers}) and indirectly via their sons (Z_{Sons}). We exploit this in the following section to examine the robustness of TSTSLS mobility estimates to who reports the Z characteristics.

3.1 Creating an auxiliary dataset

The 1,024 PSID observations described above form our "main" dataset (PSID-MAIN). To create an auxiliary dataset, we sample *with replacement* from these individuals. This generates an auxiliary sample containing 500,000 observations. (Henceforth PSID-AUX). The intuition behind this approach is similar to creating a single bootstrap re-sample. (Indeed, if we were to create an auxiliary dataset of size 1,024, this would be equivalent to us taking a single bootstrap re-sample). Specifically, by randomly re-sampling from PSID-MAIN, we create a second random draw of individuals who belong to the same population. This approach has three important advantages. First, one can guarantee that the main and auxiliary datasets are drawn from the same population. Second, the main and auxiliary datasets contain exactly the same variables measured in exactly the same way. Third, the size of the auxiliary dataset is under our control.

This method for generating the auxiliary dataset differs from the standard procedure used in the literature, where two completely different datasets are typically used for the "auxiliary" and "main" samples. Our justification for following this different approach is that this standard procedure used in the literature actually suffers from *two* problems (i) the potential inconsistency in the TSTSLS estimator and (ii) whether the main and auxiliary datasets chosen are truly comparable (e.g. do they represent the same population, define variables in the same way, both have high response rates). The focus of this paper is point (i) – the inconsistency in the TSTSLS estimator. We therefore want to abstract from point (ii) – the potential additional difficulty of finding two truly comparable datasets. Our method of generating the auxiliary dataset allows us to achieve this important goal, allowing us to rule out non-comparability of data as an explanation for our results. Nevertheless, appreciating that this is a somewhat different approach, section 4.4 describes how we have investigated the robustness of our results to generating the auxiliary dataset in a different way.

We exploit this important advantage to produce TSTSLS mobility estimates under "ideal conditions" (i.e. large auxiliary dataset, identical measurement of key variables across datasets, samples drawn from the same population). This enables us to investigate the consistency of β_{TSTSLS} and ρ_{TSTSLS} under different choices of the Z (imputer) variables. We then add additional complicating factors into the analysis (e.g. measurement of Z differing across datasets) to investigate the robustness of TSTSLS estimates to some of the other challenges researchers face.

3.2 Methodology

PSID-AUX is used to impute father's earnings ($\widehat{X}$) into PSID-MAIN following the TSTSLS approach. The twist, of course, is that PSID-MAIN also contains an actual observed measure of father's long-run earnings (X_{AVG}). One can therefore investigate how intergenerational mobility estimates change when using $\widehat{X}$ to measure father's earnings rather than X_{AVG}.

The first-stage prediction model, estimated using PSID-AUX, takes the form:

$$X_{\text{AVG}} = \alpha + \gamma Z_{\text{Fathers}} + u \tag{27}$$

with X_{AVG} being father's observed time-average earnings, and Z_{Fathers} being father's reports of the imputer variables.

The key decision is then which variables to include in Z_{Fathers}. Appendix A provides an overview of those typically used in the literature. There are four common choices:

1. broad education level – e.g. Dunn (2007)
2. broad education and broad occupation – e.g. Björklund and Jäntti (1997)
3. broad education, occupation and industry – e.g. Piraino (2007)
4. broad education and detailed (3 digit) occupation – e.g. Leigh (2007)

This guides the combination of Z used in this paper. Table 1 illustrates the variables we include in five different first-stage model specifications (henceforth M1 to M5).

These are used to impute father's earnings ($\widehat{X}$) into PSID-MAIN:

$$\widehat{X} = \widehat{\alpha} + \widehat{\gamma} Z_{\text{Fathers}} \tag{28}$$

The following regression model is then estimated six times within PSID-MAIN – once using X_{AVG} to measure father's

[5] We have also created a "modal" value for father's occupation and industry. This is where we take the most often reported occupation and industry of the father when their offspring were growing up. Approximately 75 percent of observations remain within the same occupation and industry category, regardless of whether the age 15 or modal category is used. Likewise, both variables show similar levels of consistency with sons' reports of this information.

Table 1
he imputer (Z) variables used in the first-stage prediction models

Variable	M1	M2	M3	M4	M5
Race	✓	✓	✓	✓	✓
Education	✓	✓	✓	✓	✓
Occupation (1 digit)	✗	✓	✓	✗	✗
Occupation (3 digit)	✗	✗	✗	✓	✓
Industry (1 digit)	✗	✗	✓	✓	✗
Industry (3 digit)	✗	✗	✗	✗	✓

M1 to M5 refers to the five different specifications of the first stage prediction model. All variables refer to characteristics of PSID fathers.

Table 2
Estimates of the intergenerational elasticity (β) and correlation (ρ) using different TSTSLS imputation models

TSTSLS model	First-stage R^2	Elasticity β_{TSTSLS}	Correlation ρ_{TSTSLS}
M1	0.388	0.751	0.261
M2	0.466	0.770	0.293
M3	0.489	0.716	0.279
M4	0.660	0.643	0.291
M5	0.743	0.644	0.310
OLS	-	0.570	0.318

Authors' calculations using the PSID dataset. Sample restricted to the same 1,024 individuals across all specification. SE stands for standard error. M1 to M5 indicate which first-stage TSTSLS imputation model has been used (see Table 1). Estimates using observed time-average father's earnings (OLS) reported in the bottom row.

earnings and five times using the different predictions of $\widehat{X}$:

$$Y_{2010} = \alpha + \beta X + \varepsilon \tag{29}$$

where Y_{2010} are the log annual earnings of sons in 2010, and X are father's earnings (measured using either X_{AVG} or $\widehat{X}$)

We then compare estimates of β_{TSTSLS} and ρ_{TSTSLS} (obtained using $\widehat{X}$) to β_{OLS} and ρ_{OLS} (obtained using X_{AVG}).

In our main analysis, son's earnings (Y) are taken from a single year (2010), when they are aged between 30 and 60. Ideally, to minimize the impact of "life-cycle bias" (Haider & Solon, 2006), a tighter age restriction would have been used (e.g. 35 to 45 year old sons only)[6]. Unfortunately, making such a restriction here would result in a significant reduction in sample size. We nevertheless appreciate the importance of life-cycle bias, and have hence investigated the sensitivity of our results to using a five-year average of son's earnings, with further details provided in section 4.4.

4 Results

This section presents results from our empirical analysis of the PSID. Sub-section 4.1 focuses upon the choice of the instrumental (imputer) variables. Sub-section 4.2 turns to the issue of who reports the information on these Z characteristics (fathers or their sons). Finally, sub-section 4.3 considers the impact of how earnings are measured within the auxiliary dataset.

4.1 The choice of instrumental (imputer) variables

Table 2 compares estimates of β_{TSTSLS} and ρ_{TSTSLS} to β_{OLS} and ρ_{OLS}. Whereas β_{OLS} stands at 0.570[7], TSTSLS estimate M1 equals 0.751, M2 equals 0.770 and M3 0.716. β_{TSTSLS} is thus upward inconsistent by approximately 30 percent. β_{TSTSLS} declines under M4 and M5 ($\approx$0.65) though the upward inconsistency remains non-trivial (15 percent).

To provide further insight into these results, Table 3 panel A presents the components of the inconsistency of β_{TSTSLS}, corresponding to equations (9) to (12) in section 2. For example, why is the upward inconsistency of β_{TSTSLS} not reduced between M1 and M2, despite the notable increase in the first-stage R^2? Table 3 illustrates that the addition of father's occupation (M2) also influences the direct effect of predicted father's earnings on son's earnings (λ_2); it increases from 0.30 to 0.38 as the R^2 moves from 0.39 to 0.46. In terms of consistency, losses due to the former are not offset by gains from the latter. Consequently, the upward inconsistency of β_{TSTSLS} increases from 0.181 to 0.200. This illustrates how simply choosing the instruments in order for the R^2 of the first-stage regression be as high as possible will not necessarily reduce the inconsistency of β_{TSTSLS}. Indeed, the addition of variables which influence λ_2 as well as the first-stage R^2 can actually do more harm than good.

Table 3 Panel A also reveals that two factors drive the big reduction in the inconsistency between M3 and M4. The first is the large increase in the standard deviation of father's pre-

[6]Böhlmark and Lindquist (2006) suggest lifecycle bias is approximately zero when sons are age 38 in the United States. However, Nybom and Stuhler (2016) suggest that putting a precise age on when life-cycle bias is minimised is not possible.

[7]Using a five-year average of father's earnings, Solon (1992) and Björklund and Jäntti (1997) estimate β_{OLS} to be approximately 0.40. However, B. Mazumder (2005) argues that a five-year average of father's earnings may be insufficient to eliminate problems of measurement error and transitory fluctuations. These estimates may therefore be downward inconsistent. Indeed, Mazumder obtains substantially higher values of β_{OLS} (0.61) when averaging father's earnings over 16 years. The fact that we obtain a higher estimate of β_{OLS} (0.56) than Solon and Bjorklund and JÃd'ntti is therefore likely to be due to father's earnings having been averaged over more than 20 years (see Table 1).

Table 3
Estimates of the inconsistency of TSTSLS earnings mobility estimates

	M1	M2	M4	M5	M6
(a) Elasticity					
λ_2	0.300	0.377	0.288	0.221	0.295
$\sigma_{\hat{X}}$	0.346	0.379	0.388	0.451	0.479
σ_X	0.554	0.554	0.554	0.554	0.554
$\beta_{\hat{X}.X}$	0.119	0.143	0.150	0.203	0.228
R^2	0.386	0.464	0.486	0.657	0.742
β_{OLS}	0.570	0.570	0.570	0.570	0.570
β_{TSTSLS}	0.751	0.770	0.716	0.644	0.644
Inconsistency	0.181	0.200	0.146	0.074	0.074
(b) Correlation					
β_{OLS}	0.570	0.570	0.570	0.570	0.570
$\sigma_{\hat{X}}$	0.346	0.379	0.388	0.451	0.479
σ_X	0.554	0.554	0.554	0.554	0.554
σ_Y	0.995	0.995	0.995	0.995	0.995
Inconsistency part A	−0.119	−0.100	−0.095	−0.059	−0.043
λ_2	0.300	0.377	0.288	0.221	0.295
$\sigma_{\hat{X}}$	0.346	0.379	0.388	0.451	0.479
σ_X	0.995	0.995	0.995	0.995	0.995
R^2	0.386	0.464	0.486	0.657	0.742
Inconsistency part B	0.064	0.077	0.058	0.034	0.037
ρ_{OLS}	0.317	0.317	0.317	0.317	0.317
ρ_{TSTSLS}	0.261	0.293	0.279	0.292	0.310
Inconsistency	−0.055	−0.023	−0.037	−0.025	−0.007

Authors' calculations using the PSID dataset. M1 to M5 refer to the TSTSLS imputation model specification used (see Table 1). See equation (11) and (12) for the components of the intergenerational elasticity and equations (16) to (18) for the components of the intergenerational correlation.

dicted earnings ($\sigma_{\widehat{X}}$) from 0.388 to 0.451. This, via equation (11), substantially increases the first-stage R^2. The second is the decrease in λ_2, which falls from 0.29 to 0.22. Why is there then no further reduction of the inconsistency between M4 and M5? Table 3 reveals that although $\sigma_{\widehat{X}}$ (and thus R^2) increase, λ_2 approximately returns back to its level under M3 (0.29). The effect of the former cancels out the latter, meaning no net gain regarding the consistency of β_{TSTSLS}.

Returning to Table 2, ρ_{OLS} equals is 0.318. The TSTSLS M1 estimate is 0.261; *downward* inconsistency of approximately 18 percent. However ρ_{TSTSLS} increases as additional Z_{Fathers} variables are added to the prediction model, with the downward inconsistency standing at 12 percent using M3 (ρ_{TSTSLS} = 0.279), and essentially zero using M5 (ρ_{TSTSLS} = 0.310). The inconsistency of ρ_{TSTSLS} therefore tends to be (a) in the opposite direction (b) smaller in magnitude and (c) less sensitive to the combination of the Z variables than the inconsistency of β_{TSTSLS}. Indeed, Table 2 illustrates how ρ_{TSTSLS} is not usually too far from ρ_{OLS} . This is important given that, of the near 30 studies applying TSTSLS reviewed in Appendix A, only Björklund and Jäntti (1997) report the intergenerational correlation.

Table 3 Panel B splits the inconsistency of ρ_{TSTSLS} into two components: part A (corresponding to equation 17) and part B (corresponding to equation 18). Recall how the former induces downward inconsistency in ρ_{TSTSLS}, while the latter leads to upward inconsistency. It becomes clear that the comparatively small inconsistency of ρ_{TSTSLS} (relative to the inconsistency of β_{TSTSLS}) is due to these two components partially cancelling one another out. However, the downward pressure induced by part A is always slightly greater than the upward pressure from part B, leading to the overall downward inconsistency of ρ_{TSTSLS}.

What happens as additional variables are added to the prediction model? First, the downward pressure induced by part A is always reduced. This is because the standard deviation of father's predicted earnings ($\sigma_{\widehat{X}}$) is the only term within equation A that changes (see equation 17), and can only in-

Table 4
TSTSLS estimates of the intergenerational correlation and elasticity when son's reports of father's Z characteristics

	β_{TSTSLS}		ρ_{TSTSLS}	
	Father's reports Z_{Fathers}	Son's reports Z_{Sons}	Father's reports Z_{Fathers}	Son's reports Z_{Sons}
M1	0.751	0.796	0.261	0.266
M2	0.770	0.824	0.293	0.283
M3	0.716	0.772	0.279	0.281
M4	0.644	0.686	0.292	0.276
M5	0.644	0.661	0.310	0.293
OLS	0.570		0.318	

Authors' calculations using the PSID dataset. Sample restricted to the same 1,024 individuals across all specification. Table illustrates how β_{TSTSLS} and ρ_{TSTSLS} differ when using son's reports (Z_{Sons}) of their father's characteristics (e.g. education, occupation and industry) rather than using father's own reports (Z_{Fathers}). M1 to M5 refer to the specification of the TSTSLS imputation model used (see Table 1). Estimates using observed time-average father's earnings (OLS) reported in the bottom row.

crease towards the "true" value (σ_X) as variables are added to the prediction model. In contrast, part B includes $(\frac{\sigma_{\widehat{X}}}{\sigma_Y})$ and $(1-R^2)$, with a greater value of $\sigma_{\widehat{X}}$ increasing the former but decreasing the latter. Moreover, λ_2 is also found in component B, which fluctuates in value between M1 and M5. Thus, whereas adding information to the prediction model clearly reduces the inconsistency induced by part A, the influence on part B is hard to predict. Our empirical analysis does suggest, however, that gains from the former more than offset any losses from the latter. Consequently, the inconsistency of ρ_{TSTSLS} does generally decline when information is added to the first-stage prediction model.

4.2 Measurement of imputer variables (Z)

The above investigation took place under "ideal conditions", with identical measurement of key variables across main and auxiliary datasets. We now investigate the impact of the imputer variables being measured using son's recall of their father's characteristics (Z_{Sons} in the main dataset, while individuals own reports are used within the auxiliary dataset (Z_{Fathers}).

First, we investigate the uniformity of parent (Z_{Fathers}) and offspring (Z_{Sons}) reports of father's education, occupation and industry. This investiagtion includes the percentage of occasions where father's and son's report the same category ("percentage correct") and Kappa statistics of inter-rater reliability (a statistic which adjusts for agreement occurring by chance). Kappa statistics range from −1 (complete disagreement) to +1 (complete agreement) with Landis and Koch (1977) providing the following rules of thumb:

- 0–0.20 "Slight" agreement (between parent and child reports)
- 0.21–0.40 "fair" agreement
- 0.41–0.60 "moderate" agreement
- 0.61–0.80 "substantial" agreement
- 0.81–0.99 "almost perfect" agreement

We find that fathers and sons report the same education and industry on more than 60 percent of occasions. Kappa statistics (0.52 and 0.55) are towards the top end of Landis and Koch's "moderate" agreement category, with "substantial agreement" when weighted Kappa is used (0.72 and 0.67). (Weighted Kappa is where categories further apart are considered to show greater levels of disagreement than categories closer together). In contrast, just 27 percent of father's and son's report the same category for father's occupation, with Kappa statistics suggesting agreement is "slight" (0.16) to "fair" (0.28). One potential explanation is sons were asked about their father's occupation at a vague time point ('*what was your father's occupation when you were growing up*?') which we have compared to the job father's reported holding when sons were age 15. Consequently, we are unable to establish whether this lack of agreement is due to son's inability to accurately recall their father's occupation, or different interpretation of the questions asked (e.g. son's recalling their father's occupation at a different age). Moreover, the PSID data suffers difficulties with different occupational coding schemas used across different sweeps, which may introduce a degree of uncertainty into the analysis.

Table 4 illustrates how switching to offspring reports of the imputer variables (Z_{Sons}) influences estimates of β_{TSTSLS} and ρ_{TSTSLS}. Overall, this has relatively little impact upon our results. For instance, β_{TSTSLS} (ρ_{TSTSLS}) is estimated to be 0.770 (0.293) when using imputation model M2 and *father's* reports (Z_{Fathers}. This changes to 0.824 (0.283) when using son's reports instead (Z_{Sons}). Similarly, under imputation model M5, estimates of β_{TSTSLS} (ρ_{TSTSLS}) stand at 0.644 (0.301) using father's reports, and 0.661 (0.293) using son's reports. Differences are therefore usually quite small, though on certain occasions are non-trivial. Nevertheless, our empirical analysis overall suggests that TSTSLS estimates are fairly robust to this particular measurement issue.

4.3 Imputation of current versus time-average father's earnings

Does changing the first-stage dependent variable from X_{Avg} to X_{Single} influence β_{TSTSLS} or ρ_{TSTSLS}? Table 5 provides results, with X_{Single} measured using father's earnings in 1980 (or the closest available year).

Key findings remain largely unaltered under M1, M2 and M3; large upward inconsistency in β_{TSTSLS} remains, with

Table 5
TSTSLS estimates of the intergenerational correlation and elasticity when son's reports of father's Z characteristics

	β_{TSTSLS}		ρ_{TSTSLS}	
	X_{Single}	X_{AVG}	X_{Single}	X_{AVG}
M1	0.751	0.794	0.264	0.252
M2	0.770	0.737	0.293	0.278
M3	0.716	0.665	0.279	0.260
M4	0.644	0.479	0.292	0.230
M5	0.644	0.413	0.310	0.231
OLS	0.570		0.318	

Authors' calculations using the PSID dataset. Sample restricted to the same 1,024 individuals across all specification. X_{AVG} where time-average father's earnings is the dependent variable in the first stage imputation model (i.e. "ideal conditions"). X_{Single} where father's 1980 earnings is the dependent variable in the first stage imputation model. M1 to M5 refer to the specification of the TSTSLS imputation model used (see Table 1).

slight downward inconsistency in ρ_{TSTSLS}. However, β_{TSTSLS} is now much smaller under M4 and M5. For instance, under M5 β_{TSTSLS} was 0.644 when using X_{AVG} (upward inconsistency of 15 percent). But, after changing the first-stage dependent variable to X_{Single}, β_{TSTSLS} falls to 0.413 (*downward* inconsistency of 25 percent). Similarly, ρ_{TSTSLS} using M5 is now 0.231 (downward inconsistency of 25 percent) having previously stood at 0.310 (downward inconsistency of one percent).

Table 6 breaks these TSTSLS estimates down into their respective components (corresponding to equation 22 for β_{TSTSLS} and equation 25 for ρ_{TSTSLS}). To begin, the covariance between father's and son's earnings (i.e. the common numerator of β_{TSTSLS} and ρ_{TSTSLS}) is similar – although always marginally smaller – using X_{Single}. For instance, $\sigma_{\widehat{x},y}$ under M3 falls from 0.108 using X_{AVG} to 0.101 using X_{Single}. Likewise, under M1, M2 and M3, the variance of predicted father's earnings ($\sigma_{\widehat{x}}^2$) does not seem sensitive to the choice of the first-stage dependent variable (e.g. for M3, $\sigma_{\widehat{x}}^2$ is 0.151 using X_{AVG} and 0.152 using X_{Single}). Consequently, none of the key components of β_{TSTSLS} or ρ_{TSTSLS} are particularly influenced by the use of X_{Single} rather than X_{AVG} when the first-stage prediction model is relatively sparse. Hence estimates of β_{TSTSLS} and ρ_{TSTSLS} are similar whichever earnings measure (X_{AVG} or X_{Single}) is used.

The same does not hold true, however, under M4 and M5. Specifically, the variance of father's earnings ($\sigma_{\widehat{x}}^2$) is significantly bigger when the first-stage dependent variable is X_{Single}. In contrast, the covariance between father's predicted earnings and son's earnings tends to be slightly smaller. Using M5 as an example, $\sigma_{\widehat{x}}^2$ rises from 0.229 (X_{AVG}) to 0.308 (X_{Single}, while $\sigma_{\widehat{x},y}$ falls from 0.148 (X_{AVG}) to 0.127 (X_{Single}). Thus, while the denominator of β_{TSTSLS} ($\sigma_{\widehat{x}}^2$) has substantially *increased* (and is now almost identical to the denominator of β_{OLS}) the numerator ($\sigma_{\widehat{x},y}$) has slightly *decreased* (and remains around 25 percent below the numerator of β_{OLS}). This causes β_{TSTSLS} to become *downwardly* inconsistent. Whether one uses X_{AVG} or X_{Single} as the first-stage dependent variable therefore seems to have much more influence upon the key components of β_{TSTSLS} when a detailed set of Z characteristics are included in the first-stage prediction model.

Building upon the intuition above, the *standard deviation* of father's predicted earnings ($\sigma_{\widehat{x}}$) also enters the denominator of ρ_{TSTSLS}. The increase in $\sigma_{\widehat{x}}$ from using X_{Single} as the first-stage dependent variable (as opposed to X_{AVG}) therefore also puts downward pressure on ρ_{TSTSLS}. (The impact is less pronounced than for β_{TSTSLS} due to the *standard deviation* of father's predicted earnings being the key term rather than the variance). Indeed, when using X_{Single} , ρ_{TSTSLS} actually moves further away from ρ_{OLS} as Z variables are added to the first-stage prediction model. For instance, the TSTSLS M2 estimate of *rho* = 0.278 is much closer to the OLS value (0.317) than the estimate obtained under M5 (0.230). In other words, the inconsistency of ρ_{TSTSLS} has *increased* in absolute magnitude, driven by the greater variability in father's predicted earnings. This is in direct contrast to results using X_{AVG} (presented on the left hand side of Table 6) where adding Z variables to the prediction model almost typically brought ρ_{TSTSLS} and ρ_{OLS} closer together (i.e. *decreased* the inconsistency).

These results have important implications. First, changing the first-stage dependent variable can lead to rather different estimates of earnings mobility. Second, it is only safe to assume β_{TSTSLS} is upward inconsistent if X_{True} is the first-stage dependent variable (i.e. the earnings measure being imputed into the main dataset). Third, this strengthens the empirical evidence that TSTSLS estimates of the intergenerational correlation are typically downward inconsistent. Finally, even subtle changes to the imputation model can make important differences to β_{TSTSLS} and ρ_{TSTSLS}.

4.4 Robustness tests

We have investigated the robustness of results in two ways. First, we have considered how our estimates change when using a different methodology to produce the "auxiliary" dataset. Specifically, we randomly divide our sample of 1,024 PSID observations into two equal groups. The first set of 512 observations are then designated as the "main" sample and the other 512 as the "auxiliary" sample. Hence this alternative approach does not use replacement sampling to create the auxiliary dataset. Our key findings can be summarised as follows. First, we continue to find the intergener-

Table 6
The numerator and denominator of β_{TSTSLS} and ρ_{TSTSLS} when "current" earnings used as the first-stage dependent variable

	First-stage dependent variable = X_{AVG}					First-stage dependent variable = X_{Single}				
		$\sigma_{X,Y}$		σ^2_X			$\sigma_{X,Y}$		σ^2_X	
	Coef.	Value	%	Value	%	Coef.	Value	%	Value	%
(a) Intergenerational elasticity										
M1	0.751	0.090	−49	0.120	−61	0.794	0.079	−55	0.100	−67
M2	0.770	0.110	−37	0.144	−53	0.737	0.104	−41	0.141	−54
M3	0.716	0.108	−38	0.151	−51	0.665	0.101	−42	0.152	−51
M4	0.644	0.131	−25	0.203	−34	0.479	0.117	−33	0.244	−20
M5	0.644	0.148	−16	0.229	−25	0.413	0.127	−27	0.308	0
OLS	0.570	0.175	-	0.307	-	0.570	0.175	-	0.307	-
(b) Intergenerational correlation										
M1	0.261	0.090	−49	0.346	−38	0.252	0.079	−55	0.316	−43
M2	0.293	0.110	−37	0.379	−32	0.278	0.104	−41	0.375	−32
M3	0.279	0.108	−38	0.388	−30	0.260	0.101	−42	0.390	−30
M4	0.292	0.131	−25	0.451	−19	0.238	0.117	−33	0.494	−11
M5	0.310	0.148	−16	0.479	−14	0.230	0.127	−27	0.555	0
OLS	0.317	0.175	-	0.554	-	0.317	0.175	-	0.554	-

Authors' calculations using the PSID dataset. M1 to M5 refer to the TSTSLS imputation model specification used (see Table 1). X_{AVG} where time-average father's earnings is the dependent variable in the first stage imputation model. X_{Single} where father's 1980 earnings is the dependent variable in the first stage imputation model. "Value" presents the value of the statistic in question. "%" illustrates percentage underestimation relative to OLS results.

ational elasticity to be overestimated under TSTSLS, while the intergenerational correlation is slightly underestimated. Second, there is still relatively little change to results whether one uses parent or offspring reports of their father's characteristics. Finally, we continue to find that estimates of both β_{TSTSLS} and ρ_{TSTSLS} can differ, sometimes quite markedly, depending upon the first-stage imputation model.

Second, the results presented above use information on sons' earnings from the 2011 PSID wave only, and could therefore be subject to "life-cycle bias". We have therefore investigated how results change when using a five-year average of sons' earnings instead. Three features of this supplementary analysis stand out. First, the "preferred" estimate of $\widehat{\rho}$, based upon a time-average of both father's and son's earnings, is 0.405. This is around 20 percent higher than when sons' earnings are based upon the 2011 data alone. Second, consistent with the results reported in previous sub-sections, estimates of $\widehat{\rho}$ using TSTSLS to predict father's earnings (M1 to M5) are always quite close to the value obtained under the time-average approach (typically within 10 to 15 percent). Finally, we continue to find $\widehat{\beta}$ to be overestimated when TSTSLS is used to impute father's earnings. This is consistent with the key findings presented in previous sub-sections; $\widehat{\beta}$ tends to be substantially overestimated when using TSTSLS – though with notable improvement as additional information is added to the first-stage prediction model.

Further details on these additional analyses, along with the Stata code, are available in the supplementary material (see https://zenodo.org/record/49376#.VwfmDeTmrIU).

5 Conclusions

Intergenerational earnings mobility is a topic of great academic and policy concern. However, producing consistent estimates of earnings mobility is not a trivial task. In many countries earnings data cannot be linked across generations. Consequently, several studies estimate earnings mobility using TSTSLS instead. This paper has presented new evidence on the consistency of earnings mobility estimates based upon this methodology.

A summary of our results can be found in Table 7. This illustrates the sensitivity of β_{TSTSLS} and ρ_{TSTSLS} to using different first-stage imputation models and measurement of key variables. Column 1 indicates whether X_{Single} (1980) or X_{AVG} (AVG) is the first-stage dependent variable. Column 2 indicates whether father's (FA) or son's (CH) reports of Z are used, while column 3 provides the specification of the prediction model (to be cross-referenced with Table 1). Columns 4

Table 7

A comparison of TSTSLS estimates using different measures of key variables and different imputation model specifications

First stage dependent variable	Father/son report of Z	Imputer variables (Z)	β	ρ
AVG	CH	M2	0.824	0.283
1980	CH	M1	0.801	0.251
1980	FA	M1	0.794	0.252
1980	CH	M2	0.782	0.256
AVG	CH	M1	0.796	0.266
AVG	CH	M3	0.772	0.281
AVG	FA	M2	0.770	0.293
AVG	FA	M1	0.751	0.261
1980	FA	M2	0.737	0.278
AVG	FA	M3	0.716	0.279
1980	CH	M3	0.706	0.256
AVG	CH	M4	0.686	0.276
1980	FA	M3	0.665	0.260
AVG	CH	M5	0.661	0.293
AVG	FA	M4	0.644	0.292
AVG	FA	M5	0.644	0.310
Time-average benchmark			0.570	0.318
1980	FA	M4	0.479	0.238
1980	CH	M4	0.471	0.213
1980	FA	M5	0.413	0.231
1980	CH	M5	0.346	0.182

Authors' calculations using the PSID dataset. Auxiliary dataset sample size set to 500,000 observations. "Imputer variables" refers to the Z variables used to predict father's earnings (see Table 1). AVG / 1980 refers to the first-stage dependent variable (AVG = time-average; 1980 = single measure of father's earnings in 1980). FA/CH indicates whether father's or son's reports of the Z characteristics used in the main dataset.

and 5 provide estimates of β_{TSTSLS} and ρ_{TSTSLS}. The following findings emerge:

- β_{TSTSLS} is often (although not always) upwardly inconsistent.
- β_{TSTSLS} is particularly sensitive to the choice and measurement of the first stage imputation model. Estimates are up to 50 percent upwardly inconsistent or 40 percent downwardly inconsistent.
- Estimates of ρ_{TSTSLS} tend to be more stable and suffer less inconsistency.
- Although the inconsistency of ρ_{TSTSLS} can in theory be in either direction, our empirical analysis suggests that, in practise, they tend to be below ρ_{OLS}.

Based upon our findings, we provide the following guidance to researchers wishing to estimate earnings mobility using TSTSLS. First, ρ_{TSTSLS} and β_{TSTSLS} should both be reported where possible. But, if a choice has to be made, our empirical analysis suggests there may be reasons to prefer the former over the latter[8]. Second, the auxiliary and main datasets should contain information on educational attainment and *detailed* (3 digit) occupation as a minimum. This means that at least two first-stage specifications can be estimated – a "broad" specification (as per our model M2 or M3) and a "detailed" specification (as per our model M4 or M5). One can then investigate how this changes estimates of ρ_{TSTSLS} and β_{TSTSLS}, including a breakdown into their separate components (as per our Table 6). Third, the auxiliary dataset should ideally contain information on respondents' time-average earnings (X_{AVG}). The use of cross-

[8] At the same time, "classical" measurement error in son's earnings will lead to inconsistent estimates of ρ but not β (Black & Devereux, 2011). A key implication is that estimating the intergenerational correlation is therefore more demanding in terms of the amount of information required. Counter-arguments can therefore be made as to why one may prefer β over ρ In any case, these two measures capture different aspects of mobility, with both providing important information. Hence both β over ρ should be reported whenever possible.

sectional data with respondents' earnings reported at a single time-point (e.g. a labour force survey) should be considered a second-best alternative. Fourth, as briefly discussed in section 2, standard errors should be corrected to account for the sampling variation in the predictions of father's earnings. This can be done via a Murphy-Topel correction (Murphy & Topel, 1985) or appropriate application of a bootstrap technique (Björklund & Jäntti, 1997; Inoue & Solon, 2010)). Fifth, researchers should note that their estimates of earnings mobility may differ from other studies due to methodological rather than substantive reasons. This includes instances where TSTSLS has been used in rather different ways (e.g. different combinations, definitions and measurement of key variables). Finally, we urge great care to be taken when comparing mobility estimates across studies – and across countries – where different methodologies have been applied.

References

Aaronson, D. & Mazumder, D. (2008). Intergenerational economic mobility in the United States, 1940 to 2000. *Journal of Human Resources*, *43*(1), 139–172.

Andrews, D. & Leigh, A. (2009). More inequality, less social mobility. *Applied Economics Letters*, *16*, 1489–1492.

Bauer, P. (2006). *The intergenerational transmission of income in Switzerland: a comparison between natives and immigrants.* WWZ Discussion Paper 06/ 01. Retrieved from http://ideas.repec.org/p/bsl/wpaper/2006-01.html

Bidisha, S., Das, A., & McFarlane, A. (2013). Microeconometric analysis of earnings mobility of immigrants and ethnic minorities: evidence from the UK. *Journal for Labour Market Research*, *46*(1), 35–42.

Björklund, A. & Chadwick, L. (2003). Intergenerational income mobility in permanent and separated families. *Economics Letters*, *80*, 239–246.

Björklund, A. & Jäntti, M. (1997). Intergenerational income mobility in Sweden compared to the United States. *The American Economic Review*, *87*(5), 1009–1018.

Björklund, A. & Jäntti, M. (2009). Intergenerational economic inequality. In W. Salverda, B. Nolan, & T. Smeeding (Eds.), *The oxford handbook of economic inequality* (pp. 491–521). Oxford: Oxford University Press.

Black, S. & Devereux, P. (2011). Recent developments in intergenerational mobility. In O. Ashenfelter & D. Card (Eds.), *The handbook of labor economics*. North-Holland: Elsevier.

Blanden, J. (2013). Cross-national rankings of intergenerational mobility: a comparison of approaches from economics and sociology. *Journal of Economic Surveys*, *1*(1), 38–73.

Böhlmark, A. & Lindquist, M. (2006). Life-cycle variations in the association between current and lifetime income: replication and extension for Sweden. *Journal of Labor Economics*, *1*(4), 879–896.

Cavaglia, C. (2014). *A cross-country investigation on the patterns of intergenerational mobility.* IZA working paper. Retrieved from http://www.iza.org/conference_files/SUMS_2014/cavaglia_c9961.pdf

Cervini-Pla, M. (2015). Intergenerational earnings and income mobility in Spain. *Review of Income and Wealth*, *61*(4), 812–828.

Corak, M. (2012). Inequality from generation to generation: the United States in comparison. University of Ottawa working paper. Retrieved from http://milescorak.files.wordpress.com/2012/01/inequality-from-generation-to-generation-the-united-states-in-comparison-v3.pdf

Corak, M. & Heisz, A. (1999). The intergenerational income mobility of Canadian men: evidence from longitudinal tax data. *Journal of Human Resources*, *1*(3), 504–533.

Dunn, C. (2007). The intergenerational transmission of lifetime earnings: evidence from Brazil. *B.E. Journal of Economic Analysis and Policy*, *7*(2), # 2.

Erikson, R. & Goldthorpe, J. (1992). *The constant flux–a study of class mobility in industrial societies*. Oxford: Clarendon.

Ferreira, S. & Veloso, F. (2006). Intergenerational mobility of wages in Brazil. *Brazilian Review of Econometrics*, *26*(2), 181–211.

Fortin, N. & Lefebvre, S. (1998). Intergenerational income mobility in Canada. In M. Corak (Ed.), *Labour markets, social institutions, and the future of Canada's children* (pp. 51–63). Ottawa: Statistics Canada.

Gong, H., Leigh, A., & Meng, X. (2012). Intergenerational income mobility in urban China. *The Review of Income and Wealth*, *58*(3), 481–503.

Grawe, N. (2004). Intergenerational mobility for whom? the experience of high and low earning sons in international perspective. In M. Corak (Ed.), *Generational income mobility in north america and europe* (pp. 269–310). Cambridge: Cambridge University Press.

Haider, S. & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, *1*(4), 1308–1320.

Hussain, M., Munk, M., & Bonke, J. (2009). Intergenerational earnings mobilities: how sensitive are they to income measures? *Journal of Income Distribution*, *1*(3), 79–92.

Inoue, A. & Solon, G. (2010). Two-sample instrumental variables estimators. *The Review of Economics and Statistics*, *1*(3), 557–561.

Landis, J. & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *1*(1), 159–174.

Lefranc, A. (2011). *Educational expansion, earnings compression and changes in intergenerational economic mobility: evidence from French cohorts, 1931-1976.* Paper presented to the ESPE 2010 conference. Retrieved from http://espe.conference-services.net/resources/321/2017/pdf/ESPE2010_0583_paper.pdf

Lefranc, A., Ojima, F., & Yoshida, T. (2010). The intergenerational transmission of income and education: a comparison of Japan and France. In J. Dronkers (Ed.), *Quality and inequality in education: cross-national perspectives* (pp. 229–253). New York: Springer.

Lefranc, A., Ojima, F., & Yoshida, T. (2013). Intergenerational earnings mobility in Japan among sons and daughters: levels and trends. *Journal of Population Economics*. doi:10.1007/s00148-012-0464-2

Lefranc, A. & Trannoy, A. (2005). Intergenerational earnings mobility in France: is France more mobile than the us? *Annales d'Economie et Statistique*, *78*(2), 57–77.

Leigh, A. (2007). Intergenerational economic mobility in australia. *The B.E. Journal of Economic Analysis & Policy*, *7*(2), # 6.

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*, *1*(2), 235–255.

Mocetti, S. (2007). Intergenerational earnings mobility in Italy. *The B.E. Journal of Economic Analysis & Policy*, *7*(2), # 5.

Murphy, K. & Topel, R. (1985). Estimation and inference in two step econometric models. *Journal of Business and Economic Statistics*, (4), 370–379.

Murtazashvili, I., Liu, D., & Prokhorov, A. (2013). Two-sample non-parametric estimation of intergenerational income mobility. Retrieved from http://alcor.concordia.ca/~aprokhor/papers/TSNPGMM.pdf

Nicoletti, C. & Ermisch, J. (2008). Intergenerational earnings mobility: changes across cohorts in Britain. *The B.E. Journal of Economic Analysis & Policy*, *7*(2), # 9.

Nuñez, J. & Miranda, L. (2011). Intergenerational income and educational mobility in urban Chile. *Estudios de Economia*, *38*(1), 195–221.

Nybom, M. & Stuhler, J. (2016). Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *Journal of Human Resources*, *1*(1), 239–268.

Piraino, P. (2007). Comparable estimates of intergenerational income mobility in Italy. *The B.E. Journal of Economic Analysis & Policy*, *7*(2), # 1.

Piraino, P. (2014). *Intergenerational earnings mobility and equality of opportunity in South Africa.* Southern Africa Labour and Development Research Unit working paper 131. Retrieved from http://opensaldru.uct.ac.za/bitstream/handle/11090/696/2014_131_Saldruwp.pdf?sequence=1

Roccisano, F. (2013). On intergenerational mobility in Italy: what a difficult future for the young. *Review of Applied Socio-Economic Research*, *6*(2), 203–216.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, *1*(3), 393–408.

Ueda, A. (2009). Intergenerational mobility of earnings and income in Japan. *The B.E. Journal of Economic Analysis & Policy*, *9*(1), # 54.

Ueda, A. (2013). *Education and intergenerational earnings transmission: the case of Japan and South Korea.* Institute for Research in Contemporary Political and Economic Affairs working paper. Retrieved from http://www.waseda-pse.jp/file/genseiken/WP/No.E1305

Vogel, T. (2008). *Reassessing intergenerational mobility in Germany and the United States: the impact of differences in lifecycle earnings patterns.* SFB Discussion Paper 2006-055, Humbolt University, Berlin. Retrieved from http://www.wiwi.hu-berlin.de/professuren-en/vwl/wtm2/Employees/vogel/vogelmobility

Wooldridge, J. (2002). *Econometric analysis of cross-section and panel data*. Massachusetts: MIT Press.

Appendix
Table

(Appendix table follows on next page)

Table A1
Intergenerational mobility papers imputing father's earnings using TSTSLS

Study	Country	Sample size (Main data)	Offspring' income	Sample size (Auxiliary)	Imputer variable and 1st stage R^2
Aaronson and D. Mazumder (2008)	United States	Men, 25-54 years old, born btw 1921 and 1975.	Earnings	1940-1970: 1% sample; 1980-2000: 5% sample	State of birth R2: Not reported
Andrews and Leigh (2009)	16 countries	Not reported	Son's log hourly wage.	Not reported	192 Occupation dummies (off-spring reported) R^2: Not reported
Bauer (2006)	Switzerland	2,138	Average earnings from work	41,362	Occupation (9 dummies); Education (7 dummies); Swiss citizen dummy R^2 = 0.27
Bidisha, Das, and McFarlane (2013)	United Kingdom	3,823	Average log wages of full time workers and earnings of self-employees over the panel	935	Education (3 dummies); Occupation (3 dummies); immigrant status; ethnic group; professional level (4 dummies); cohort (2 dummies); Hope-Goldthrope score; R^2=0.323
Björklund and Jäntti (1997)	Sweden and USA	Sweden: 327 US: Not reported	Annual log earnings and capital market income	Sweden: 540 (US: Not reported)	Education (2 dummies); Occupation (8 dummies); Living in Stockholm (Note: Children reports) R^2: Not reported
Cavaglia (2014)	Germany, Italy, UK, US	Germany = 27,442 Italy = 6,860 UK = 14,363 US = 7,530	Labour income	Germany = 4,534 Italy = 1,516 UK = 4,989 US = 7,918	Education, occupation and industry R^2 Germany = 0.47 R^2 Italy = 0.34 R^2 UK = 0.31 R^2 US = 0.22

Continues on next page

Continued from last page

Study	Country	Sample size (Main data)	Offspring' income	Sample size (Auxiliary)	Imputer variable and 1st stage R^2
Cervini-Pla (2015)	Spain	3,520 sons 3,995 daughters	Annual log earnings of sons. For daughters: log family income and log couples' earnings.	5, 929	Education (6 dummies); Occupation (9 dummies). R^2: 0.40
Dunn (2007)	Brazil	14,872	Annual log "earnings from all jobs"	37,396	Father's education (10 categories) R^2: Not reported.
Ferreira and Veloso (2006)	Brazil	25,927	Log wages.	59,340	Father's education (7 dummies), Father's occupation (6 dummies) R^2: Not reported
Fortin and Lefebvre (1998)	Canada	Father-son: 3,400 (1986) and 2,459 (1994) Father-daughter: 2,474 (1986) and 2,308 (1994)	Annual income	ca 500,000 each year	Father's occupation (15 groups) R^2: Not reported
Gong, Leigh, and Meng (2012)	China	5,475	Annual log income.	Varies depending on UHIES sample.	Father's education; Father's occupation; Industry. R^2: Not reported
Grawe (2004)	Ecuador, Nepal, Pakistan, and Peru	Ecuador: 1,461 Nepal: 229 Pakistan: 171 Peru: 98	Total wage income	Ecuador: 685 Nepal: 239 Pakistan: 441 Peru: 166	Father's education.
Lefranc, Ojima, and Yoshida (2010)	France and Japan	Japan: 987 France 13,487	Japan: Individual primary income (labour + assets) before tax or transfer France: Annual earnings from labour.	Fathers btw 25 and 54, in Japan Fathers btw 24 and 60 in France.	Linking variables: year of birth; 3 educational levels and occupation (Japan) year of birth; 6 levels of education (France) R^2: N.R.

Continues on next page

Continued from last page

Study	Country	Sample size (Main data)	Offspring' income	Sample size (Auxiliary)	Imputer variable and 1st stage R^2
Lefranc (2011)	France	29,415	Annual wages	48,245	Father's education (6 groups); Note: Offspring reports R^2: Not reported
Lefranc and Trannoy (2005)	France and USA	1977: 2,023 1985: 2,114 1993: 771	Wages	2,364–6,488 depending on the year.	Father's education (8 groups); Father's occupation (7 groups); (Note: Offspring reported) R^2: 0.49–0.54
Lefranc, Ojima, and Yoshida (2013)	Japan	2,273	Gross individual income	7,170	Father education (3 groups); Father occupation (8 groups) Firm size (2 groups); Self-employment; Residential area (3 groups). R^2: 0.46
Leigh (2007)	Australia	1965: 946 1973: 1,871 1987: 243 2004: 2,115	Hourly wages	1965: 946 1973: 1,871 1987: 243 2004: 2,115	Father's occupations (78 to 241 groups depending on survey); Offspring reported. R^2: Not reported
Murtazashvili, Liu, and Prokhorov (2013)	US and Sweden	US:467 Sweden: 324	Annual earnings	US: 1,613 Sweden: 565	Father's education; Father's occupation
Mocetti (2007)	Italy	3,200	Gross income from all sources but financial assets.	4,903	Father's education (5 groups; Work status (5 groups); employment sector (4 groups); geographical area (3 groups). R^2: 0.30
Nicoletti and Ermisch (2008)	UK	8,832	31-45 years old sons, with positive income (employed or self-employed) in at least one wave of the panel	896	Father's occupation (4 groups); Father's education (5 groups). R^2: 0.31

Continues on next page

Continued from last page

Study	Country	Sample size (Main data)	Offspring' income	Sample size (Auxiliary)	Imputer variable and 1st stage R^2
Nuñez and Miranda (2011)	Chile	11,186	25 to 40 years old log earnings of sons working at least 30hs x week	1987: 19,192 1990: 20,378	Father's occupation (4 groups); Father's education (5 groups). R^2: 0.29–0.37.
Nuñez and Miranda (2011)	Chile (Greater Santiago)	649	Log income	1,736–2,700 (depending on the year)	Father's education (3 groups); Father's occupation (5 groups) R^2: 0.48–0.66
Piraino (2007)	Italy	1,956	Gross income from all sources bar financial assets.	953	Father's education (5 groups); work status (4 groups); employment sector (4 groups); geographical area (2 groups) $R^2 = 0.33$.
Piraino (2014)	South Africa	1,241–2,590	Monthly gross employment earnings	1,355	Education (5 groups); Occupation (5 groups) $R^2 \approx 0.40$
Roccisano (2013)	Italy	786	Earnings	3,203	Education; Occupation; Industry $R^2 \approx 0.20$
Ueda (2009)	Japan	1,114 married sons; 906 single daughters; 1,390 married daughters	Gross annual earnings and income from all sources.		Father's years of education; Father's occupation and firm size (7 groups). R^2: Not reported.
Ueda (2013)	Korea and Japan	Both countries: size varies depending on civil status of the sons and daughters	Annual earnings	Korea: Fathers btw 25 and 54 Japan:	Korea: education and occupation Japan: parental income . R^2: Not reported
Ueda (2013)	Taiwan	745	Annual income	745	Father's education (6 groups); Father's occupation (11 groups). R^2: Not reported.