The Outlook for U.S. Labor-Quality Growth *

Canyon Bosler
University of Michigan

Mary C. Daly
Federal Reserve Bank of San Francisco, IZA

John G. Fernald
INSEAD, Federal Reserve Bank of San Francisco

Bart Hobijn
Arizona State University

November 20, 2017

Abstract

Over the past 15 years, labor-quality growth has been very strong—defying nearly all earlier projections—and has added around 0.5 percentage points to an otherwise modest U.S. productivity picture. Going forward, labor quality is likely to add considerably less and may even be a drag on productivity growth in the medium term. Using a variety of methods, we project that potential labor-quality growth in the longer run (7 to 10 years out) is likely to fall in the range of 0.1 to 0.25 percent per year. In the medium term, labor-quality growth could be lower or even negative, should employment rates of low-skilled workers make a cyclical rebound towards pre-recession levels. The main uncertainties in the longer run are whether the secular decline in employment of low-skilled workers continues and whether the Great Recession pickup in educational attainment represents the start of a new boom or is simply a transitory reaction to a poor economy.

JEL classification codes: J24, O47, O51.

Keywords: Demographic change, growth accounting, labor quality, wages.

*Prepared for NBER/CRIW conference on "Education, Skills, and Technical Change: Implications for Future U.S. GDP Growth." We thank Doug Elmendorf, Chuck Hulten, Valerie Ramey, Todd Schoellman, and the CRIW participants for helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

1 Introduction

Economists have long recognized the importance of human capital accumulation for economic growth. And since the seminal analysis of Jorgenson & Griliches (1967), which provided a straightforward measurement framework, indices of human capital, or labor quality, have become standard in growth-accounting studies for many countries. In this paper, we assess alternative methods for estimating U.S. labor quality and provide projections for the future. We also identify key uncertainties that will determine the actual path of U.S. labor quality in the medium and longer run. In almost all scenarios we consider, labor quality adds less to growth over the next decade than it has historically—in some scenarios, much less.

We begin by reviewing commonly used methods for measuring labor quality. Since labor quality is not directly observable, measuring it requires researchers to find an observable proxy. Not surprisingly, the best proxy is wages, which should move closely with marginal products. For example, a neurosurgeon is likely to have a higher marginal product than a grocery clerk. This difference in marginal products is, in turn, arguably the main reason why the neurosurgeon is paid more.

The question is how best to impute the relative marginal products of workers based on different characteristics. We develop a novel statistical metric that evaluates the reliability of alternative approaches to imputing relative marginal products. Specifically, we examine the trade-off that each approach implicitly makes between (a) the share of the productivity-related variation in observed wages that is explained, and (b) the precision of the imputed estimates of relative marginal products of different workers.[1]

In our statistical assessment, the best-performing model is a parsimonious Mincer specification that includes experience, education and, when accurate data are available, occupation. Experience and education are clearly related to productivity differentials across workers, and are empirically important for explaining the patterns of wages in the data. Other commonly used variables raise challenges. For example, both occupation and gender add explanatory

[1] For example, adding an additional variable might add explanatory power for wages but at the cost of sharply reducing precision of imputed marginal products.

power with little cost in terms of precision. But, historically, occupation has been challenging to forecast with any degree of accuracy so, for the purpose of projections, we exclude it. For gender, it is unclear to what degree gender-related wage differentials reflect marginal products, so we again prefer to exclude it. (In any case, including gender turns out to make little difference empirically to our estimates of labor quality). Other variables (such as industry or race) add little to explanatory power while substantially reducing the precision of estimated marginal products.

We then use our preferred parsimonious Mincer specification to estimate labor-quality growth from 2002 to 2013 across three alternative datasets.[2] We find that labor quality grew about 0.5 percent per year—somewhat faster than its post-war average of about 0.4. Indeed, labor quality arguably explains a bit under one-third of labor-productivity growth of 1.8 percent per year over the 2002-2013 period.[3] This finding is robust across data sources.

Strikingly, the growth and acceleration of labor quality since 2002 has a very different source than it did in the half century before that. In the 20th century, the primary driver of labor quality increases was rising educational attainment (Ho & Jorgenson, 1999; Goldin & Katz, 2009; Fernald & Jones, 2014). In contrast, since 2002, the source of labor-quality growth has been a shift in the composition of employment away from lower skilled and toward higher skilled workers. This change owed to ongoing secular changes in the labor force as well as cyclical adjustments associated with the Great Recession.

Building on this analysis, we provide alternative scenarios for the evolution of labor-quality growth over the medium and longer run. Our work reinforces the view that labor-quality growth will add less to growth in productivity and output than it has historically. That said, the actual path of labor-quality growth is sensitive to uncertainties about trends in

[2] The time period is constrained by our desire to compare results across three publicly available data sources.

[3] This contribution is calculated assuming that growth in output per hour rises one-for-one with growth in labor quality. That is, the growth in labor quality is *not* multiplied by labor's share, which would give the proximate growth-accounting contribution. The one-to-one mapping comes from standard economic models, where there is an indirect effect from endogenous growth in capital. The reason is that capital deepening in the models is typically in terms of "effective labor." Fernald & Jones (2014) discuss this accounting and estimate that increases in labor quality explained 0.4 percent per year of the 2.0 percent annualized growth in U.S. GDP per hour between 1950 and 2007.

employment rates and, to a lesser extent, educational attainment. These differences will show up in productivity growth, but whether they matter for output growth depends on the degree to which they are offset by hours growth. This highlights a takeaway from our analysis, namely that labor-quality growth and hours growth are often negatively correlated. An important implication of this is that forecasts of overall labor input growth, or quality-adjusted hours, are preferable to independent projections of labor quality and hours.

Section 2 reviews the growth-accounting definition of labor quality that we apply in this paper. Section 3 then discusses the practical challenges involved in empirically applying our conceptual framework and assesses alternative approaches and data sets. Section 4 examines the evolution of labor quality since 2002, and compares approaches and data sets. Over this period, labor-quality growth was boosted by disproportionate declines in employment rates among low-skilled workers, especially during and after the Great Recession.

With a framework in place, Section 5 turns to projections of labor-quality growth over the medium- to long-term. We forecast that labor-quality growth is likely to slow to somewhere in the range of 0.1 to 0.25 percentage points a year over the next 10 years. Should employment composition return to its pre-recession levels, medium-term labor-quality growth will fall below this baseline and could even turn negative. In the longer run, trends in education and employment rates are central. To generate labor-quality growth at close to its historical pace requires not just a continuing shift in the composition of employment from low-skill towards high-skilled workers, but also a resumed upward trend in educational attainment. Although such a scenario is possible, we think it unlikely. In particular, although educational attainment has picked up since 2007, our preferred interpretation is that the rise represents a transitory reaction to a poor economy, not a new upward trend.

2 Definition of labor-quality growth

Indices of labor quality are based on standard neoclassical production theory.[4] Consider a neoclassical value-added production function of the form

$$Y = F\left(A, K, H_1, \ldots, H_n\right). \tag{1}$$

Output, Y, is produced by combining the n types of labor inputs, $H_1, ..., H_n$, with a capital input, K; A denotes the level of technological efficiency with which the inputs are combined.[5]

To quantify how changes in inputs affect output growth, we apply a first-order logarithmic Taylor approximation. Small letters denote the natural logarithms of the capitalized variables such that y is the log of output, Y. Applying the first difference operator, Δ, we can write

$$\Delta y = \ln Y_t - \ln Y_{t-1}. \tag{2}$$

This is simply the growth rate of output, as measured by the change in the logarithm of output. The Taylor approximation then reads

$$\Delta y = \frac{\partial F}{\partial A}\frac{A}{Y}\Delta a + \frac{\partial F}{\partial K}\frac{K}{Y}\Delta k + \sum_{i=1}^{n}\frac{\partial F}{\partial H_i}\frac{H_i}{Y}\Delta h_i. \tag{3}$$

Output growth depends on technology growth plus the contribution of the various factors of production. The final term in this expression is the effect of changes in labor inputs on output growth, where growth in each type of labor is multiplied by its respective output elasticity.

The contributions of labor inputs can be further decomposed into the effect of growth in total hours (i.e., growth in $\sum_{i=1}^{n} H_i$) and changes in the composition of total hours. To do this

[4] Ho & Jorgenson (1999) survey the history of labor quality measurement and discuss several semantic and/or conceptual confusions.

[5] Assuming a single capital input is for simplicity and does not affect the results that follow for labor input.

we rewrite (3) as

$$\Delta y = \frac{\partial F}{\partial A}\frac{A}{Y}\Delta a + \frac{\partial F}{\partial K}\frac{K}{Y}\Delta k + \left(\sum_{j=1}^{n}\frac{\partial F}{\partial H_j}\frac{H_j}{Y}\right)\left(\Delta h + \sum_{i=1}^{n}\frac{\frac{\partial F}{\partial H_i}H_i}{\sum_{j=1}^{n}\frac{\partial F}{\partial H_j}H_j}\left(\Delta h_i - \Delta h\right)\right). \quad (4)$$

Growth in total hours is Δh and the change in the composition of hours worked is

$$\sum_{i=1}^{n}\frac{\frac{\partial F}{\partial H_i}H_i}{\sum_{j=1}^{n}\frac{\partial F}{\partial H_j}H_j}\left(\Delta h_i - \Delta h\right). \quad (5)$$

The change in the composition of hours worked in (5) amplifies or attenuates growth in total labor input relative to growth in total hours. This wedge between growth in labor input and growth in hours is commonly interpreted as labor-quality growth. Intuitively, if all type of labor inputs, H_i, grow at the same rate, then the composition of total hours does not change and labor-quality growth is zero. But if, instead, hours of relatively more productive workers (with high $\frac{\partial F}{\partial H_i}$) grow more quickly than hours of less productive workers, then labor-quality growth will be positive.

Empirically, the marginal products of labor, $\frac{\partial F}{\partial H_i}$, in (5) are not observed. Under standard neoclassical conditions, the $\frac{\partial F}{\partial H_i}$ are proportional to the nominal hourly wage earned by workers of type i, denoted W_i. We assume that the proportionality constant is equal across types of labor.[6] If this is the case then

$$\frac{\frac{\partial F}{\partial H_i}H_i}{\sum_{j=1}^{n}\frac{\partial F}{\partial H_j}H_j} = \frac{W_i H_i}{\sum_{j=1}^{n}W_j H_j}, \quad (6)$$

which is the share of total compensation that gets paid to workers of type i.

Under these assumptions, labor-quality growth, denoted by g_{LQ}, is the compensation-share-

[6] In competitive markets, standard neoclassical assumptions imply that real (output-price-deflated) wages equal marginal products, so the assumption holds (with proportionality given by the output price). Imperfect competition in the output market allows firms to charge a markup of price over marginal cost; but the markup is constant across types of workers so the assumption again holds. It also holds if firms have some monopsony power in the labor market, as long as the wedge is constant across types of labor.

weighted average deviation of labor input from total hours growth by type, i.e.

$$g^{LQ} = \sum_{j=1}^{n} \frac{W_i H_i}{\sum_{j=1}^{n} W_j H_j} \left(\Delta h_i - \Delta h\right). \tag{7}$$

This is the measure of labor-quality growth that we analyze. It is the same as the one used in range of growth-accounting datasets for many countries.[7]

Note that growth in total labor input, or "quality-adjusted" hours, is simply the share-weighted growth in hours:

$$g^{LQ} + \Delta h = \sum_{j=1}^{n} \frac{W_i H_i}{\sum_{j=1}^{n} W_j H_j} \Delta h_i. \tag{8}$$

3 Measurement of labor-quality growth

To implement (7) and obtain an empirical estimate of labor-quality growth requires three things:

i **Definition of worker types:** decision regarding the specific types of workers, $i = 1, \ldots, n$, the labor-quality index will distinguish between.

ii **Estimate of wage by worker type:** estimate of average hourly earnings for each worker type, W_i, used to construct the share of each worker type in total compensation.

iii **Measure of hours:** measure of hours worked by worker type, H_i, used to calculate the deviation of hours growth by worker type, Δh_i, from overall hours growth, Δh.

Item (*iii*) is relatively straightforward. Measures of hours worked by individuals are available in many datasets. Once the worker types are defined, calculation of H_i simply involves aggregation of hours across individuals in each of the n groups.

[7] For the United States, examples include Jorgenson *et al.* (1987, 2014); Ho & Jorgenson (1999); Zoghi (2010), and Bureau of Labor Statistics (2015a,b). Notable examples for a wider set of countries include EUKLEMS (O'Mahony & Timmer, 2009), the Conference Board's Total Economy Database (van Ark & Erumban, 2015), and the Penn World Tables (Feenstra *et al.*, 2015).

Items (i) and (ii), are less straightforward than (iii) and we discuss the different options for dealing with them in this section. We are not the first to discuss the choice of worker types and wage measures in the context of the construction of labor-quality indices. For example, see Zoghi (2010). Our contribution relative to that work is to introduce a framework that allows us to make tractable choices for (i) and (ii) and "test" those choices against each other using standard statistical techniques.

In terms of datasets, we focus primarily on the American Community Survey (ACS). The ACS is a smaller, annual version of the decennial census and collects a relatively narrow range of demographic and socioeconomic data on a sample of about 1 percent of the U.S. population (approximately 3 million individuals) each year.[8] We also consider two other datasets. The first is the Current Population Survey's Outgoing Rotation Groups (CPS-ORG), which consists of the outgoing rotation groups from the Current Population Survey (CPS). This is the quarter of the CPS respondents that are asked about their earnings and income in any given month. This results in an annual sample of about 135,000 individuals. The second, the Current Population Survey's Annual Social and Economic Supplement (CPS-ASEC), is the Annual Social and Economic Supplement to the Current Population Survey, also known as the March Supplement. It contains annual earnings and income data from the full March CPS sample (70,000 individuals).

Though based on different samples and sampling methods, each of the datasets allows for the construction of similar hourly wages, as well as the six variables of education, age, sex, race/ethnicity, industry, and occupation, that are our main focus. In all cases, we measure hours as usual hours worked per week, which is available in all three datasets.

[8] The sample of the ACS has been expanded twice and has only been a 1 percent sample of the population since 2006. In its first year, 2000, the sample was just under 400,000 individuals and between 2001 and 2005 the sample was slightly over 1 million.

3.1 Criteria for choosing worker types and wage estimates

Indices of labor quality are built by dividing workers into groups based on their marginal products of labor, $\frac{\partial F}{\partial L_i}$. The decision about how many and which worker types, $i = 1, \ldots, n$, to use depends on: (i) the degree to which the types distinguish between workers with different marginal products, and (ii) the degree to which the different worker types capture the cross-individual variation in wages.

A simple way to quantitatively assess the degree to which these criteria are met for any particular grouping is a regression. To see this, consider j individuals and denote the log of their individual hourly wage by w_j. For each individual we also observe a vector $\mathbf{x}_j$ of individual-level characteristics based on their worker type, i. Under the assumption that relative wages reflect relative marginal products, the extent to which the characteristics in the vector, $\mathbf{x}_j$, capture cross-individual differences in marginal products can be measured as the fraction of individual-level log-wage variation that is explained by the variables in $\mathbf{x}_j$. This measure is equal to the R^2 of the following standard log-wage regression

$$w_j = \mathbf{x}_j' \boldsymbol{\beta} + \varepsilon_j. \tag{9}$$

Here, $\mathbf{x}_j' \boldsymbol{\beta}$ is the part of the wage variation captured by the variables in $\mathbf{x}_j$.

Though simple, this specification is very general. It subsumes the case in which the elements of $\mathbf{x}_j$ are dummy variables that span the set of worker types. In this version, every type is a stratum made up of individuals with the characteristics, as in Jorgenson *et al.* (1987).[9] It also includes the case where $\mathbf{x}_j$ contains polynomial terms of variables affecting workers' marginal product. In this case (9) is a form of a Mincer (1974) regression. This is the model used by Aaronson & Sullivan (2001), among others.

Of course, in practice we do not know the true parameter vector $\boldsymbol{\beta}$ and the log-wage regression (9) is estimated using a sample of workers of finite size. This means that, at best,

[9] Most stratum-based studies use median rather than mean wages. Our results are not sensitive to this choice.

we can obtain an estimate $\hat{\boldsymbol{\beta}}$ of the parameter vector and that we thus infer the part of wages captured by our explanatory variables with error. To formalize this mathematically, we denote the standard deviation of the estimation error of the explained part as

$$\sigma_j = \sqrt{E\left[\left(\mathbf{x}_j'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right)^2\right]}. \tag{10}$$

Since it is important to have a reliable estimate, the smaller σ_j the better. However, for the construction of the labor-quality index, we are not interested in one particular worker, j, but instead in the reliability of the relative marginal product estimate, $\mathbf{x}_j'\hat{\boldsymbol{\beta}}$, across the whole sample. To gauge the reliability of the marginal product estimate across the sample, we consider the p^{th} percentile of the standard errors, σ_j, across individuals. We denote this percentile by $\tilde{\sigma}_p$.

Based on this simple framework, we suggest two statistical criteria for determining the types of workers to distinguish and the method to use when estimating wages.

1. R^2 **of log-wage regression** This measures the share of cross-individual wage variation that is captured by our choice of worker types and specification of the log-wage equation.

2. **Percentile of standard error, $\tilde{\sigma}_p$, of marginal product estimates** This captures how reliably we estimate the (relative) marginal product of labor across workers.

Higher R^2's and lower $\tilde{\sigma}_p$'s are preferred.

Importantly, there is a direct trade off between these two measures. In principle, we can obtain an $R^2 = 1$ in the estimated regression (9) by including as many linearly independent variables in $\mathbf{x}_j$ as we have observations, m. However, this would result in a regression with zero degrees of freedom and $\tilde{\sigma}_p \to \infty$. Alternatively, we can aim for a very low $\tilde{\sigma}_p$ at the expense of a R^2.

Using these tools we can directly compare different choices of (i) worker types and (ii) wage estimates by worker type. We do so using a scatterplots that plot the R^2 and $\tilde{\sigma}_p$ for each choice that we consider. Before we construct the scatterplots, we first describe the choices of worker types and wage regression specifications we consider.

3.2 Choice of worker types and wage regression specifications

So far, we have discussed the choices of worker types, i, and the regression specification, i.e. $\mathbf{x}_j$, as two distinct decisions. In practice, however, they are one and the same. This is because for the variables that are commonly considered in log-wage regressions there are only a countable number of values. Consequently, for a given regression specification in terms of these variables there is only a finite number of permutations of $\mathbf{x}_j$ across individuals. In this context, a worker type, i, corresponds to a permutation of the covariates vector $\mathbf{x}_j$.

With this in mind, two questions remain: (i) which variables should be included in the vector $\mathbf{x}_j$, and (ii) what functional form of these variables works best?

Choice of variables in wage equation

The decision regarding which variables should be included in the regression is guided by the assumption, underlying the labor-quality growth derivation, that wage differentials between worker types reflect differences in relative marginal products of labor. This means that the variables we include in the wage equation should have two properties. First, they should explain a substantial part of the variation in wages across worker types. Second, the part of wage variation they explain should reflect only differences in marginal products.

Whether a variables has the first property is straightforward to verify statistically. The second property—i.e., which variables capture marginal product differentials—is more controversial. This is because certain observable characteristics may be correlated with wedges between wages and marginal products.[10] Though such variables might improve the fit of the wage regression, (9), including them in our measure of labor quality would bias our results.

The most obvious variables to consider for inclusion in (9) are education and experience.

[10] See Boeri & van Ours (2013) for a textbook treatment of many possible sources of such wedges.

Several decades of running Mincer regressions has demonstrated a robust correlation between education and potential experience (or age) and wages (Psacharopoulos & Patrinos, 2004).[11] Although there is some controversy over the degree to which returns to education are derived from improved human capital as opposed to the signaling of unobservable worker characteristics, both perspectives tend to attribute educational wage differentials to differences in marginal products (Weiss, 1995).[12] Overall, there is broad agreement that the correlation between wages and education or experience is driven by real productivity differentials.[13]

A substantial literature, summarized in Altonji & Blank (1999), has also pointed to a role for gender, race, and ethnicity in explaining wage differentials. Here we encounter substantial controversy as to whether, or to what degree, these wage differentials reflect differentials in productivity as opposed to discrimination. On the one hand, gender differentials may capture the fact that women are more likely to work part time or leave the labor force temporarily, which is not captured in the measures of experience available in standard datasets (Light & Ureta, 1995). And ethnic differentials may proxy for unobserved language barriers that have a real impact on productivity (Hellerstein & Neumark, 2008).[14] Yet, there is also a substantial literature documenting the existence of labor market discrimination, particularly on the basis of race and ethnicity, in both hiring and wages (Bertrand & Mullainathan, 2003; Pager *et al.*, 2009; Hellerstein *et al.*, 2002; Oaxaca & Ransom, 1994).

Finally, there is also a body of literature suggesting that there are inter-industry wage differentials that persist even after controlling for education and experience (Dickens & Katz,

[11]Some of the recent Mincer regression literature has suggested that there are important differences in the education-experience return profiles between cohorts (Lemieux, 2006; Heckman *et al.*, 2008). We allow for such cohort effects in that we estimate wage regressions on annual cross-sectional data. Thus, in our analysis cohort and age effects are indistinguishable. This is appropriate for our application, because we are only interested in making robust wage predictions and not in isolating specific returns.

[12]Outside of developing countries there has been little empirical research that even asks the question of whether educational wage differentials might reflect something other than productivity, and the research in developing countries has generally concluded that the differentials are consistent with differences in productivity (Jones, 2001; Hellerstein & Neumark, 1995).

[13]Broad as the agreement is, it is not entirely universal: incomplete labor contracts, labor market segmentation, or cultural factors could potentially drive a wedge between wage premia associated with education and experience and differentials in marginal product (Blaug, 1985).

[14]Skrentny (2013) and Lang (2015) discuss the theoretical and empirical evidence on race and worker productivity.

1987; Krueger & Summers, 1988).[15] Once again, such differentials could originate from genuine differences in productivity (e.g. the matching of a worker to a particular job may reflect differences in social skills (Deming, 2015)) or from non-productivity related features of an industry (such as profit sharing). Interestingly, although similar arguments could apply to occupational differences, there has been little research that considers whether there are persistent inter-occupation wage differentials independent of educational and experience prerequisites. Though not the main purpose of our analysis, our estimates of (9) partially fill this void by including occupation in our analysis.

Thus, the observables we focus on are age, education, gender, race, industry, and occupation. We are aware that there are many other variables that could be interpreted as reflecting differences in marginal product of labor across workers. Examples include marital status, rural-urban location, or family structure. However, given the limited evidence that these variables are of first-order importance in explaining cross-individual variation in wages, we omit them from our analysis.

There is also a wide range of potentially influential unobservable characteristics (such as entrepreneurial talent (Silva, 2007), cognitive and non-cognitive abilities (Heckman *et al.*, 2006), and physical attractiveness (Hamermesh & Biddle, 1994)).[16] Although it would be ideal to include measurements of, or proxies for, these characteristics in our analysis, that is not possible in the datasets available.

Choice of functional form

With the set of variables to include in $\mathbf{x}_j$ in hand, that last thing to consider is the specific functional form imposed on these variables. For example, is the traditional Mincer regression, with a constant, linear years of education, and a quadratic polynomial in experience, the appropriate functional form or should dummies for high school graduation and college graduation

[15] Gibbons *et al.* (2005), however, suggest that sectoral wage differentials can be accounted for by allowing for sector-specific returns to skill.

[16] These characteristics are unobservable in the sense that they are not measured as part of the standard datasets (ACS, CPS-ASEC, and CPS-ORG) that we use for our analysis.

be included to account for sheepskin effects (Hungerford & Solon, 1987)? Are education and experience additively separable, or is there a nonlinear interaction between the two? These questions have been investigated quite carefully for the traditional Mincer regression variables of education and experience (Lemieux, 2006), but less attention has been paid to the other variables.

Given this uncertainty around the appropriate functional form, one approach is to allow for the maximum flexibility in the log-wage regression, (9). To do this, one would treat each possible combination of values of the included variables as a worker type. This boils down to running a fully non-parametric regression in which $\mathbf{x}_j$ is a vector with separate dummies for each worker type. The fitted log wage, $\mathbf{x}_i'\hat{\boldsymbol{\beta}}$, for each worker type in that case is the average log-wage for workers with that combination of values for the included variables. This approach, though flexible, results in a significant loss of degrees of freedom.

For example, if we only consider age and education, restrict the population under consideration to 16-64 year olds, and distinguish 16 educational categories (as is the case with most standard U.S. micro datasets), then this regression has 768 estimated parameters corresponding to the 768 possible permutations of age and education in the data. In practice many of these worker types will contain very few observations in the data. For those worker types for which there is only one observation the standard error of the estimated mean log-wage is infinite, i.e. $\sigma_i = \infty$.

Though such a non-parametric regression might result in a very good fit, the heterogeneity in marginal products of labor across worker types will be estimated with a high degree of uncertainty.

Stratum-based methodologies, which have been used extensively in prior growth accounting exercises that account for labor quality (Gollop & Jorgenson, 1983; Jorgenson *et al.*, 1987; Ho & Jorgenson, 1999; Jorgenson *et al.*, 2014), are a form of this type of dummy regression. Stratum-based studies define worker types by partitioning the population by observable characteristics, with the mean wage of each partition being interpreted as the wage for workers of that type.

In practice, in order not to run into the curse of dimensionality described above, stratum-based studies do not treat each value of a variable as distinct. Instead, they group different values of the variables together. For example, the 16 educational categories are often collapsed into less-than-high-school, high-school, some college, and college categories. Using a less granular partition regains some degrees of freedom but with a loss of some flexibility in the functional form. How granular a partition can be used largely depends on the sample size of the dataset used.

In the context of the regression framework that we use here, this grouping of values imposes multidimensional step functions on the data. Thus, although the most granular partitions result in a non-parametric regression that will have an R^2 that is at least as high as any other regression specification, the partitions used in practice actually impose a restrictive functional form that does not necessarily fit the data better than alternative model specifications.

Concerns about the step functions imposed by partitioned dummy regressions have led some researchers to hew more closely to the Mincer-regression literature (Aaronson & Sullivan, 2001; Bureau of Labor Statistics, 1993). These specifications focus on education and experience as the fundamental drivers of human capital, marginal product, and wages.[17] These regressions generally include education (either as a polynomial in years of education or as a set of dummies indicating levels of educational attainment) and a polynomial in experience.

In addition to the baseline education and experience variables, these human capital specifications often include some interaction between gender and experience to account for women's higher rate of part-time work and temporary withdrawal from the labor force (either as an interaction between gender and experience or by estimating the regression on men and women separately). In some cases (Aaronson & Sullivan, 2001; Bureau of Labor Statistics, 1993, 2015a,b) they also include control variables like part-time status, marital status, veteran's status, race, and rural location. These variables are not included to capture differences in marginal products across workers but instead to reduce omitted variable bias in the education and experience coefficients.

[17] As commonly done, we define experience as the difference between age and years of education (plus six).

Comparison of specifications

Between the question of which variables to include and what functional form to impose, the task of selecting a preferred regression specification for a labor quality measure is quite daunting. Even in the narrowed down set of variables we consider, age, education, gender, race, industry, and occupation, there are several options on how to group their values. For each of the six variables we use, Table 1 lists how many different classifications we consider for our comparison of model specifications. In the last four columns of each row, the table lists how many groups are defined for each classification. For example, for age we consider two classifications: one that splits the individuals up into 9 age groups and another into 13 age groups. The number of permutations across the different classifications of variables is 192. This includes one classification for each of the variables. Once one allows for dropping variables, then the possible number of stratum specifications increases to 1,799. The most detailed one, which includes the most granular classification for all variables, consists of 8,486,400 worker types.[18]

As noted, we apply the statistical tools R^2 and $\tilde{\sigma}_p$ as two clear criteria on which we can base our model specification decision. For our application we use the adjusted R^2, i.e $\bar{R}^2$, as it penalizes for overfitting the data. We consider the 80$^{\text{th}}$ percentile of the standard errors of the estimated relative marginal product of labor across workers, i.e. we use $\tilde{\sigma}_{80}$ as our measure of the reliability of the imputed wages.[19]

We complete our analysis using three different datasets. Our results are qualitatively very similar across datasets. For the sake of brevity, we present results obtained using the ACS, since this is the dataset with the largest sample size.[20]

Figure 1 illustrates the tradeoff between the goodness-of-fit, $\bar{R}^2$, and the precision of the wage imputation, $\tilde{\sigma}_{80}$. Panel 1a shows the scatter plot in $\left(\tilde{\sigma}_{80}, \bar{R}^2\right)$ space for all 1799 stratum-

[18] To put the amount of potential overfitting in perspective, this most granular definition of strata means that, on average, there are less than 20 workers per worker type in the U.S., since civilian employment has never exceeded 150 million.

[19] In principle, the choice of p for the percentile is arbitrary. However, qualitatively all results that we emphasize in this section hold for choices of $p > 75$. The reason we do not use the mean is that, in the case of the stratum-based methods, $\sigma_j = \infty$ for all worker types with one observation. This would also make the sample mean of the σ_j's go to ∞.

[20] See Appendix B for results based on CPS-ORG and CPS-ASEC data.

based model specifications from Table 1. This panel shows how increasing the $\hat{R}^2$ of the model specification comes at the cost of the precision with which the relative marginal products are imputed, i.e. an increase in $\tilde{\sigma}_{80}$. Because a higher $\hat{R}^2$ and lower $\tilde{\sigma}_{80}$ are preferred, we are focusing on specifications that move us to the upper-left in the plotted $\left(\tilde{\sigma}_{80}, \bar{R}^2\right)$ space.

Panel 1b shows the same 1799 points as panel 1a with two sets of points highlighted. The red crosses are the 192 stratum specifications that include all six variables we consider, with the difference being the level of granularity at which the variables are classified. These points are the ones where $\tilde{\sigma}_{80}$ is high, compared to $\bar{R}^2$, and thus correspond to specifications that overfit the data. At the other end of the cloud of points are the ones highlighted as red circles. These are the specifications that do not include age and education. The blue points are specifications that include age and education but not all four of other variables. When we compare the red circles with the blue points we find that, among the blue points, there are several specifications that have a substantially higher $\bar{R}^2$ and not much higher levels of $\tilde{\sigma}_{80}$.

We find that adding occupational dummies to the stratum definitions that already condition on age and education, yields the greatest improvement in fit and a relatively small decline in the precision of the imputed wages. This can be seen from panel 1c, which highlights the specifications that add only occupations as green dots. As can be seen from the figure, adding occupation adds about 0.1 to the $\bar{R}^2$ but increases $\tilde{\sigma}_{80}$ only slightly. In contrast, adding industry alone, depicted by the red squares, does not improve the fit as much as adding occupation and results in lower precision with which the wages are imputed. [21] Adding both industry and occupation results in values of $\tilde{\sigma}_{80}$ well above 0.5. This means that for more than 20 percent of the strata log-wages are imputed with a standard error of more than 0.5 (65 percent).

Panel 1d adds gender and race/ethnicity to the stratum definitions that include education and age. Race/ethnicity only slightly increases the fit at the cost of a substantial reduction in the precision of the marginal product imputation. Gender also increases the fit but it is hard to know whether this reflects marginal product differentials or other factors. Since, in our

[21] Of course, for some purposes, such as estimating industry-specific labor-quality indices, including industry dummies may still be necessary.

analysis, the in- or exclusion of gender does not have a large effect on estimates of labor-quality growth we exclude gender from our specifications in the rest of this paper.

In addition to the stratum-based model specifications, we also consider Mincer-type regressions. In particular, the *baseline* Mincer specification on which we settled includes a quadratic polynomial in experience and 5 education dummies.[22] Because our stratum-based analysis suggests that occupation is an important determinant of wages, we also consider a *baseline-plus-occupation* specification which adds 51 occupation dummies.[23]

Figure 2 compares the regression-based fit and precision of imputed wages for the *baseline* and *baseline-plus-occupation* specifications with the stratum-based specifications. The lower cross in the figure shows the point for the baseline specification and the lower blue dots are the stratum-based points that only include age and education. Because the Mincer-type regression is more parsimonious than the semi-parametric regressions, it results in more precisely imputed marginal product levels across workers, i.e. it has a smaller $\tilde{\sigma}_{80}$. Moreover, the quartic polynomial in experience captures more of the variations in wages across workers than the piecewise linear specifications implied by the stratum-based methods. Consequently the regression results in a higher $\bar{R}^2$. Thus, the flexibility of the semi-parametric specification that Zoghi (2010) emphasizes when she proposes to use stratum-based medians as estimates of wages,[24] is outperformed by the quartic polynomial in experience that we use here. As a result, the Mincer-regression based way of imputing wages dominates the stratum-based methods in terms of both model-selection criteria.

This is not only true for the *baseline* regression specification. It is also true for the one that includes occupational dummies. In Figure 2 the upper cross corresponds to the *baseline-plus-occupations* regression and the upper cloud of blue dots to the corresponding stratum-based regressions that include age, education, and occupation. Again, the Mincer-regression-based

[22] A similar Mincer specification, with the addition of several control variables, was also used by Aaronson & Sullivan (2001),

[23] We focus on this parsimonious *baseline* specification in the main text and illustrate that our main qualitative results are unaltered when additional covariates are included as controls in Appendix B.

[24] The regression framework we use here results in the conditional mean for a stratum to be the imputed wage. In unreported results we redid our analysis with the conditional median as the wage estimate and obtained the same results compared to the Mincer specifications.

specification outperforms the stratum-based ones.

This evidence shows that our *baseline* and *baseline-plus-occupation* specifications perform well in terms of our two model-selection criteria.

3.3 Index formula

Given the choice of the vector $\mathbf{x}_j$ and the period-by-period estimates of the parameter vector $\hat{\boldsymbol{\beta}}_t$, based on (9), the final choice to be made for the calculation of the labor-quality index is the index formula.

In line with the log-linear approximation of (4), the index formula that is used for most labor-quality index calculations is of the Translog form and estimates labor-quality growth as the compensation-share weighted average of log changes in hours across worker types.[25] That is,

$$\hat{g}_t^{LQ} = \sum_{i=1}^{n} \left(\frac{s_{i,t} + s_{i,t-1}}{2} \right) (\Delta h_i - \Delta h)\,, \text{ where } \hat{W}_t(\mathbf{x}_i) = \exp\left(\mathbf{x}_i' \hat{\boldsymbol{\beta}}_t\right) \tag{11}$$

$$\text{and } s_{i,t} = \frac{\hat{W}_t(\mathbf{x}_i)\, H_{i,t}}{\sum_{s=1}^{n} \hat{W}_t(\mathbf{x}_s)\, H_{s,t}}. \tag{12}$$

This translog index formula has the desirable property that it is a so-called superlative index (Diewert, 1978). That is, it is an exact index for a function (the translog) that provides a general second-order approximation of the production function. In other words, the labor-quality index does not rely simply on a first-order approximation (though we used such an approximation in our derivation in Section 2 for expositional clarity).

For labor quality, implementation of the translog formula is complicated by the fact that in some cases the number of hours worked by a worker type, i is zero. In that case Δh_i can not be calculated and such worker types are dropped from the calculations. Though dropping these worker types is a reasonable option, because their compensation share is, presumably, small, one can also use another superlative price index formula that does not suffer from this

[25] Compensation share are averaged across the two periods between which growth rates are calculated.

problem.

This is what we do in this paper. In particular, we follow Aaronson & Sullivan (2001) and use a Fisher Ideal index formula of the form

$$\hat{g}_t^{LQ} = \left\{ \frac{H_{t-1}}{H_t} \right\} \left\{ \frac{\sum_i \hat{W}_t(\mathbf{x}_i) H_{i,t}}{\sum_i \hat{W}_t(\mathbf{x}_i) H_{i,t-1}} \right\}^{\frac{1}{2}} \left\{ \frac{\sum_i \hat{W}_{t-1}(\mathbf{x}_i) H_{i,t}}{\sum_i \hat{W}_{t-1}(\mathbf{x}_i) H_{i,t-1}} \right\}^{\frac{1}{2}} - 1. \tag{13}$$

This formula allows us to include all worker types, i, in our calculations even if $H_{i,t} = 0$ or $H_{i,t-1} = 0$.[26][27]

4 Historical labor-quality growth

Before we consider projections of labor-quality growth, we first examine its behavior over the past 15 years. This is useful for two reasons. First, by comparing historical results for different specifications and datasets, we can assess how sensitive the labor-quality growth estimates are to the different choices discussed in Section 3. Second, and most importantly, the concerns about plateauing educational attainment and the retirement of experienced older workers that many observers currently express were also raised as concerns in the early 2000's. Our historical analysis shows that, contrary to these concerns, labor-quality growth barely slowed over the past 15 years. This realization of labor-quality growth owes much to a reduction

[26]For our benchmark specification, the problem of zeros does not occur, and the Translog and Fisher are virtually identical. It can make a little more difference in cases with extremely large numbers of cells, where there are more zeros.

[27]Note that exponentiating the predicted log-wage would not normally be sufficient to get a predicted wage in levels because

$$E[w_j] = E[\exp(\mathbf{x}_j\beta + \epsilon_j)] = E[\exp(\mathbf{x}_j\beta) + (\epsilon_j)] = \exp(\mathbf{x}_j\beta) \cdot E[\exp(\epsilon_j)]$$

and $E[\exp(\epsilon_j)]$ is not 1. It is, however, a constant if the residuals are assumed to be independently and identically distributed. So if $\hat{W}_i = \exp(\mathbf{x}_j\beta)$ and $c = E[\exp(\epsilon_j)]$, then plugging the predictions into the share of the wage bill calculation from (7) gives

$$\frac{W_i H_i}{\sum_{j=1}^n W_j H_j} \equiv \frac{c\hat{W}_i H_i}{\sum_{j=1}^n c\hat{W}_j H_j} = \frac{c\hat{W}_i H_i}{c\sum_{j=1}^n \hat{W}_j H_j} = \frac{\hat{W}_i H_i}{\sum_{j=1}^n \hat{W}_j H_j}.$$

Therefore we need not make any adjustments to the predictions, nor do we need to impose an assumption on the distribution of the residuals beyond the standard assumption that they are IID.

in the employment rates of less productive individuals, especially during and after the Great Recession. We will return to this point in the projection section.

4.1 Comparison across methods and datasets

As we discussed in Section 3, we construct our benchmark labor-quality index using ACS data based on our *baseline* Mincer specification. The index for labor quality obtained from this specification is plotted as the blue line, labeled *Regression - age and education*, in Figure 3a.

From 2002 through 2013 the cumulative growth in the index was 5.96 percent, which is 0.53 annually. As the figure shows, labor-quality growth has been far from constant at this average during our sample period. Its standard deviation across years is 0.39. From 2002 to 2006 labor quality by this measure grew relatively slowly, about 0.37 percent per year. Subsequently, during the Great Recession from 2008-2010 labor-quality growth logged in at 0.94 percent a year. Since then it has come down to 0.36 percent.

In Section 3 we showed how our *baseline* specification outperformed many others in terms of goodness of fit of the log-wage regression as well as the precision of imputed wages. In terms of labor-quality growth our baseline specification yields an estimate that is very close to those obtained using other specifications that include age and education. This can also be seen in Figure 3a. As the figure plots, the stratum- and regression-based methods give very similar estimates of the labor-quality index when both age and education are included in the vector $\mathbf{x}_i$. Moreover, the index constructed does not change very much when we use the *baseline-plus-occupation* specification instead of the *baseline* specification.

Among the series plotted in Figure 3a there are two clear outliers that exhibit much less cumulative labor-quality growth. The first is the stratum-specification that includes all variables. Such a specification results in large errors in imputed wages, which reduces the correlation between hours growth and wages that drives labor-quality growth. As a result, the overfitted specification yields much less labor-quality growth than our *baseline* model. The other outlier series is the version that excludes age and education entirely (the underfit stratum). That series is flat, confirming that age and education are what drive the series.

Excluding the two outlier series, the cross-specification mean of average annual growth rates of labor quality is equal to the average annual labor-quality growth rate implied by our *baseline* index, namely 0.53 annually. The cross-specification standard deviation in these average annual rates is 0.03. Besides very similar mean growth rates, all these indices also show a very similar qualitative pattern over the sample period: Slow growth from 2002-2006, an acceleration during the Great Recession, and a subsequent slowdown in 2011 and 2012.

The results in Figure 3a are reminiscent of Zoghi (2010)[28] in that she suggests that estimated average annual labor-quality growth rates are fairly robust to the choice of model specification. This robustness of estimated average annual labor-quality growth rates also translates across datasets.

This can be seen from Figure 3b. It plots the *baseline* and *baseline-plus-occupation* results for the three datasets that we consider in this paper, i.e. for ACS, CPS-ASEC, and CPS-ORG. The six indices plotted look very similar.[29] In terms of their summary statistics, the mean average annual labor-quality growth rate across series in the figure is 0.49 percent with a standard deviation of 0.03.

Together, these results suggest that the pattern of labor-quality growth from 2002 through 2013 we find using our *baseline* case is not the result of the particular specification or dataset chosen. Indeed, we find this pattern for all reasonable model specifications and across all data sets. Overall, we conclude that from 2002-2013 labor quality has grown around 0.5 percent a year. This is about the same as the average of about 0.5 percent labor-quality growth between 1992 and 2002 (Bureau of Labor Statistics, 2015a; Fernald, 2015).

[28]See Zoghi (2010) Table 12.2, page 478.

[29]The only exception is the ACS-based indices in 2005-2006. In this year the sample size of the ACS was expanded from 1 to 3 million respondents, which appears to have resulted in a sample with a slightly lower level of labor quality than before.

4.2 Counterfactuals to identify the sources of growth

The fact that we find no substantial deceleration in labor-quality growth since 2002 is surprising, especially given the slow growth of educational attainment and the beginning of retirement among the oldest baby boomers, during the period. Our analysis shows that as these adverse demographic and educational trends were pulling down labor-quality growth, a disproportionate decline in the employment-to-population (EPOP) ratio of lower quality worker types was pushing it up. To illustrate this, we calculate three counterfactual historical indices, which are plotted in Figure 4.

These counterfactuals take advantage of the fact that hours worked by workers of type i, H_i, are the product of (i) average hours worked per year by workers of this type, η_i, (ii) the EPOP of these workers, E_i, and (iii) the population of these workers, P_i. That is,

$$H_i = \eta_i E_i P_i. \tag{14}$$

Using this expression, we can create different counterfactuals by holding one of the three factors, i.e. η_i, E_i, and P_i, fixed at its 2002 level. We then allow the other two factors to change as observed in the data.

Figure 4 shows our baseline estimate, labeled *observed index*, as well as the three counterfactual indices. As can be seen from the figure, changes in average hours worked across worker types have had relatively little impact on labor-quality growth. In contrast, if the composition of the population had not changed since 2002 then labor-quality growth would have been about a third lower. This is because removing population changes eliminates the continued accumulation of experience of the baby boom generation from the calculations.

The most striking of the three counterfactuals, however, is the one for the EPOP ratio. From Figure 4 it is clear that if EPOP ratios by worker type had remained at their 2002 levels labor-quality growth would have been half of what we observed over the past decade. Notably, the wedge between the observed index and the counterfactual with constant EPOP ratios increased most rapidly during the Great Recession. This wedge is consistent with the extensively documented composition effect of recessions on real wages. Many studies, including those by Bils (1985) and Solon *et al.* (1994), find that the incidence of unemployment is more cyclical among low-wage workers.

In growth-accounting terms, this cyclical composition effect means that labor quality has a countercyclical component (Ferraro, 2014). This is reflected in the strong negative correlation of around -0.9 between labor-quality growth and hours growth as measured by our *baseline* specification. This negative correlation is quite robust across specifications: Figure 5 plots the correlations for all of the labor quality specifications plotted in Figure 3a except the overfit and underfit stratum specifications, and all of the correlations are strongly negative. An implication of this negative correlation is that it is important to jointly forecast labor quality and hours worked to get a robust estimate of labor input going forward.

As discussed, our labor-quality index captures the fact that EPOP ratios among lower quality worker types are more cyclical. And our counterfactuals show that the disproportionate decline in employment rates among less-skilled workers led to a recession-driven increase in labor-quality growth. Therefore, an important question for any medium-term forecast of labor-quality growth is to what extent these movements in EPOP ratios by worker types are transitory or permanent. Since a large part of the decline in these EPOP ratios reflects declines in labor force participation rates, this is largely a question of what fraction of recent movements in labor force participation is structural versus cyclical.

If labor force participation rebounds substantially, as Congressional Budget Office (2015) projects, this will put downward pressure on labor-quality growth over our forecast horizon. However, if, as Aaronson *et al.* (2014) suggest, the bulk of the movements in participation rates across groups since 2007 have been structural, then our labor-quality index would be largely unaffected. In that case there would be no downward pressure on labor-quality growth coming from changes in labor force participation by skill level.

This finding highlights an important lesson from our analysis. We should not be misled by the positive sound of "*increases* in labor quality" due to composition effects. Often, labor quality is discussed assuming a path of total hours. But an important factor driving labor-quality growth since the early 2000s has been declines in hours (or a slowdown in hours growth) for lower-skilled workers. From (4) we know that what matters for output growth is the growth rate of the total labor input, which is hours growth plus labor-quality growth. Hence, if labor quality grows as a result of a selection effect among workers when total hours decline, then this is neither necessarily good news for growth of overall labor input nor for output growth.

5 Projecting labor-quality growth

In this section we consider the outlook for labor-quality growth over the next ten years. We begin by reviewing the components of labor-quality growth projections. We then evaluate the performance of our baseline specification for 2002-2013, paying particular attention to the components that have contributed most to historical projection errors. Guided by these findings we provide a range of alternative scenarios for future labor-quality growth in both the medium and longer run.

5.1 Components of labor-quality growth projections

As previously discussed, the index for labor quality, equation (13), is a highly non-linear function of the parameter vector $\boldsymbol{\beta}_t$ and hours worked by worker type H_i. The fact that wages and hours are endogenous to one another further complicates the problem. In practice, producing an optimal forecast of labor-quality growth based on the joint distribution of future log-wage regression coefficients and future hours worked by worker type is not feasible.

In its place, researchers generally project labor-quality growth by projecting independently the log-wage parameter vector, $\boldsymbol{\beta}$, and the hours worked by worker type, H_i, and substitute them into equation (13).[30] Given that time-variation in the $\boldsymbol{\beta}$s accounts for a very small portion of labor-quality growth over time, the convention is to hold log-wage parameters constant (Aaronson & Sullivan, 2001; Jorgenson *et al.*, 2015). We follow this convention and set $\hat{\boldsymbol{\beta}}_{t+h} =$

[30] This gives a joint projection of hours and labor quality, which is important given the negative correlation between hours and quality documented in Section 4.2

$\hat{\boldsymbol{\beta}}_{2013}$.

Turning to hours, recall that hours worked by worker type, $\hat{H}_{i,t+h}$, can be decomposed into the three factors as in equation (14), namely (*i*) average hours, $\hat{\eta}_{i,t+h}$, (*ii*) the EPOP rate, $\hat{E}_{i,t+h}$, and (*iii*) population, $\hat{P}_{i,t+h}$. Historically accounting for heterogeneity in average hours worked by worker type does not make a material difference. This is highlighted in Figure 6 which plots the observed baseline index against an employment-based index constructed under the assumption that all workers work the same number of hours, i.e. $\eta_{i,t} = \eta_t$ for all i. The employment-based index shows average annual labor-quality growth of 0.61 percent, about a tenth of a percentage point higher than the 0.53 obtained from the hours-based index.[31] Given the modest difference, and the significant challenges associated with projecting heterogeneous hours worked, we set $\hat{\eta}_{i,t+h} = \eta_{t+h}$ for all worker types i. We use this 0.61 percent observed average annual growth of labor quality as our baseline for comparing the observed index with forecasts.[32]

5.2 Historical projection accuracy and sources of error

In this section we examine how our baseline projection specification would have performed for the 2002-2013 sample period. Specifically, we compare our projection to observed labor-quality growth and use an informal decomposition to evaluate the sources of forecast errors. Following Aaronson & Sullivan (2001) we build our projections using Census Bureau 2000 ("middle") National Population Projections by age, gender, and race.[33] To obtain population projections for all age and education combinations, we apply a multinominal logit model that estimates

[31] This difference between the hours-worked-based and employment-based indices is even smaller in the CPS-ORG and CPS-ASEC data than in the ACS (see Appendix B).

[32] Note that our baseline specification does not include occupation. Including occupation requires projecting population and EPOP ratios by age, education, and occupation. This turns out to result in very imprecise projections, since projections of employment by occupation, without considerations by age and education, already have large errors. This can be seen in Figure 7, which plots the Bureau of Labor Statistics' actual versus projected growth in employment between 1996 and 2006 by 6-digit SOC codes. To avoid introducing these errors into our projections, we limit ourselves to projections using our baseline specification.

[33] Our projection method is differs from Aaronson & Sullivan (2001) in the following ways: we distinguish five racial groups instead of four, define employment more narrowly to be consistent with our sample selection, and use ACS data.

the probability distribution of our five educational levels, based on age, cohort, gender, and race. We use these estimated probabilities to construct population projections by age and education, i.e. to construct $\hat{P}_{i,t+h}$, for each year. Finally, to project the age and education specific EPOP ratios, $\hat{E}_{i,t+h}$, we estimate the probability that an individual is employed as a function of age, cohort, and education, using logit models that vary by gender and race.[34]

The results are shown in Figure 8. The top line in Panel 8a shows the observed employment-based index of labor quality which grew at an average annual pace of 0.61 percent. The bottom line in Panel 8a shows our projection of labor-quality growth as of 2002. The results are strikingly different, our projection expected average annual labor-quality growth to rise just 0.19 percent, well below the pace observed over the period. This large difference result is consistent with projections by Aaronson & Sullivan (2001) which used a slightly different model specification and CPS-ASEC data rather than ACS.

The remaining lines in the panel 8a plot counterfactual indices that replace (a) projected demographics with observed demographics and (b) projected log-wage regression parameters with observed parameters. The line labeled "2002 betas; observed demographics" is much closer to the actual index than the forecast index, suggesting that errors in the projected demographic variables account for a substantial portion of the forecast error over the period. In contrast, the line labeled "Observed betas; projected demographics" is very close to our baseline projection and far from the observed index. This suggests that time-variation in the log-wage regression parameters accounts for a very small portion of the forecast errors. Panel 8a also shows that the bulk of the forecast errors accumulate during the Great Recession. In other words, deviations in demographics, $H_{i,t}$, from their projections, $\hat{H}_{i,t}$, in the Great Recession account for much of the forecast error.

Panel 8b takes a closer look at the specific demographic variables contributing to the large projection errors. The lines labeled "All observed" and "All projected" are the "Observed

[34] Because the first ACS data were released in 2002, we cannot use ACS data for the estimation of the EPOP and educational attainment models. Instead, we estimate these models using 1992-1997 data from the CPS-ORG for this historical forecast. The full technical details of this projection are provided in subsection A.2 of Appendix A.

index" and "Observed betas; projected demographics" lines from Panel 8a. The line labeled "Observed age & education; proj employment" reflects an alternative index based on observed components of the demographics, less the EPOP ratios, for which we use projections. The difference between this line and the "All observed" index isolates the effect of projection errors in EPOP ratios across worker types. As can be seen from the figure, these errors account for about one-third of the cumulative forecast error in labor-quality growth and are especially important after the onset of the Great Recession in 2008. The line "Observed age; proj education & employment" shows that projection errors in educational attainment also account for about one-third of the forecast error in labor-quality growth. The remaining error owes to misses in Census' population projections.[35] Notably, the projection errors for education and population accumulate relatively smoothly over our sample period.

5.3 Projections of future labor-quality growth

Going forward most commentators project labor-quality growth will be slower than its historical pace. This view stems from the fact that the exceptional increases in U.S. educational attainment during the 20th century seem unlikely to be repeated (Goldin & Katz, 2009). However, as we will show, this oversimplifies the uncertainties surrounding the future path of labor-quality growth both in the medium and the longer run. To illustrate these uncertainties and how they relate to various components of labor-quality growth we consider three potential future paths for educational attainment and employment-to-population rates and assess how these paths affect estimates of future labor-quality growth in the medium and longer run. These alternative paths, which are briefly described below and fully explained in Appendix A.3, illustrate the mechanics of how different economic forces influence future labor quality. Given the limited role of the $\boldsymbol{\beta}$ (which capture relative returns to experience and education)

[35] Since this is not a formal decomposition we are not accounting for the nonlinear contributions associated with interactions between the Census demographics, distribution of education, and employment rates. These interactions, however, appear to be relatively minor compared to the first-order contributions of demographics, education, and employment.

in the accuracy of the historical projections, we hold at them fixed at their 2013 values.[36] To allow sufficient time for the economy to recover from the effects of the Great Recession we define the medium term as 2015-2022. The longer run is 2022-2025.

For employment-to-population, we consider three alternative paths. The paths are meant to illustrate a range of potential outcomes.

(i) *Cyclical rebound, or 'revert'*: Age-education specific EPOP rates return to 2007 values, between 2015 and 2022, and remain there. This scenario corresponds to the view that the changes in EPOP rates for specific age-education groups were cyclical.[37]

(ii) *Structural change, or 'persist'*: Age-education specific EPOP rates remain at 2013 levels. This scenario corresponds to the view that much of the decline in EPOP rates following the Great Recession is permanent.

(iii) *Extrapolated 2002-2007 structural trends in EPOPs*: The final path allows for heterogeneous paths across groups. Specifically, it extrapolates the declining EPOP rates of young people (with heterogeneity across education groups), the increasing EPOP rates of older people (particularly the more educated), and the widening gap between the EPOP rates of more and less educated prime-age people (Dennett & Modestino, 2013; Burtless, 2013; Aaronson *et al.*, 2014).

The paths above illustrate how changes in various EPOP rates affect future U.S. labor-quality growth.

We also consider three alternative paths for educational attainment. Again, these paths are meant to highlight a range of potential outcomes.

(i) *Revert to pre-crisis levels* During the Great Recession enrollment and graduation rates rose. This path assumes that the increase was a temporary cyclical effect and rates will return to their pre-crisis levels.

[36] An alternative approach to projections of education and employment would be to use a statistical model, following Aaronson & Sullivan (2001). Experiments with this methodology produced variable results that appear less reliable, especially in the more distant future, than the methods we employ here.

[37] Both the cyclical (i) and structural (ii) paths allow for a demographically (or educationally) driven structural decline in the aggregate employment-to-population ratio. However, they do not allow a structural decline (or increase) in age-education specific EPOPs.

(ii) *Persist at 2013 educational plateau* This alternative assumes that the uptick in educational attainment in recent years persists through future cohorts. Specifically, 2013 rates of educational attainment carry forward for each cohort over the next decade.

(iii) *Extrapolate 2007-2013 trends in education* The final path assumes that the uptick in educational attainment over the past several years represents a resumed upward trend. Projections are based on age-specific time trends in educational attainment from logistic regressions.

Table 2 shows projections for 2015-22. All scenarios incorporate the Census Bureau's population projections by age group. In addition, the scenarios incorporate differing medium-run cyclical dynamics for employment rates and education. The columns of the table show the three alternative EPOP assumptions. The rows show the three educational attainment assumptions. For each cell, the first number shows growth in labor quality and the second shows growth in hours. (Note that, since we do not model average hours worked, hours grow at the same rate as employment growth.[38])

A notable takeaway from the Table 2 is the potentially negative correlation in the medium run between growth in hours and growth in labor quality. The negative correlation appears in the two "level" EPOP scenarios. The "persist" (structural) scenario results in 0.35-0.36pp faster labor-quality growth than the "revert" (cyclical) scenario. However, this is fully offset by 0.35-0.39pp slower growth in hours. As a result, growth of total labor input grows at 0.59-0.67pp per year in all of the scenarios in columns 2 and 3. This negative correlation highlights the importance of jointly modeling these two variables to obtain a forecast for quality-adjusted hours.

The near-invariance of quality-adjusted-hours growth across the level scenarios seems surprising at first glance. Intuitively, an extra hour of work should add *something* to quality-adjusted hours—albeit more if it involves higher-skilled workers. The reason for the near-invariance in Table 2 is that low-skilled and high-skilled workers have seen an opposite pattern in EPOP ratios since 2007. Employment rates of lower-skilled workers have fallen while rates

[38] This is equivalent to assuming that all workers work the same number of hours, a counterfactual assumption but one which has been relatively innocuous historically (see Figures 6 and B.4).

for higher-skilled workers have risen. Thus, the "revert" scenario includes not only a rise in employment by lower-skilled workers but also a decline by higher-skilled workers.

The first column of Table 2, which extrapolates 2002-07 EPOP trends, looks quite different from the others. In this case, we see markedly stronger growth in both labor quality and hours. For lower-skilled workers, there was little pre-recession trend in EPOP rates. For this group of workers, this extrapolation-based scenario thus looks similar to the "revert" scenario, which boosts hours but holds labor quality down. But for higher-skilled and older workers, the pre-recession trend was to increase employment rates. These workers tended to be below their estimated trend in 2013. Hence, in this scenario, these workers add both hours and skills to the labor force between 2015 and 2022. For both groups, hours increase quickly as employment rates rise. For labor quality, the extra hours of high-skilled workers dominate and labor quality rises more quickly.

Finally, looking down the columns, for none of the cases do the education scenarios matter much between 2015 and 2022. Extrapolating the rising educational trend from 2007-13 (row 1) matters only a few basis points over this time period. The dominant force in the medium run is thus what happens to employment rates.

Turning to the longer run, Table 3 shows projections for 2022-2025. These scenarios assume that all cyclical/transitional dynamics will have taken place by 2022.

In the longer run, educational trends do matter. Looking down the three columns, the educational-extrapolation row implies almost 2/10ths pp faster growth in labor quality than the "revert" or "persist" rows, with minimal difference in hours worked. Of course, this educational-extrapolation path assumes a considerable acceleration in educational attainment relative to what we have seen since WWII. Our reading of the data so far is that there is little indication that such an educational acceleration is actually happening. Rather, we view one of the plateau scenarios for educational attainment as more plausible—either the scenario where educational attainment for entering cohorts reverts to its 2007 levels, or where it persists at its 2013 levels. The CPS data suggest that some of the Great-Recession-induced increase in educational attainment of younger cohorts may already be reversing.[39] The "revert" and "persist" rows of Table 3 are very similar for both labor quality and hours. Relative to the revert or persist scenarios—which are very similar—we take the predictions from the educational extrapolation scenario as an upside risk for labor quality.

Finally, we consider the importance of employment rates for longer-run projections of labor quality. In the longer-run, only *trends* in EPOP rates matter. Indeed, the two 'level' columns look very similar to each other, showing that in the longer run it makes little difference whether we revert to pre-crisis EPOPs or remain at 2013 EPOPs.

5.4 Putting it all together

The previous section highlighted the uncertainties around any forecast of labor-quality growth both in the medium and longer-run. Here we provide a judgmental assessment of the most likely path for labor quality in the longer-run. Looking at the bottom right two cells of Table 3, where education plateaus and EPOPs remain level, we project labor-quality growth of about 0.1 percent per year and hours growth of a little above 0.4 percent per year. Quality-adjusted hours in these scenarios grows a little above 0.5 percent per year.

Although these 'level' scenarios are a reasonable benchmark for the future, continuing shifts in EPOPs also seem plausible. Earlier, we found that these shifts were central to driving labor-quality growth from 2002-2013. This was also the case for the 2002-2007 period, before the employment effects of the Great Recession. Going forward, there is certainly the potential for technological advances to continue to generate job polarization, to displace low- and medium-skill workers, and/or to entice high-skill workers to increase their labor supply. If these trends were all to continue at their 2002-2007 pace, then it would lead to some longer-run boost in labor quality, though the effect on hours is ambiguous.

One particular unknown in this regard is whether older, more educated workers will continue to work longer than they have historically. For example, suppose we extrapolate EPOP

[39] Additional evidence for a reversal comes from Census data on college enrollments relative to the population aged 16-24. That enrollment rate peaked in 2011 and has since retreated somewhat.

trends only for those over 55 years of age—a situation that would boost both labor quality and hours. With that limited extrapolation, we would see hours growth of about 0.55 percent and labor-quality growth of about 0.15 percent, implying quality-adjusted hours growth of about 0.70 percent per year. In the 2022-25 period, these figures are not affected by whether other employment rates revert to pre-crisis levels or remain at 2013 levels.

Trends for individuals under age 55 are more nuanced and challenging to predict. In the extrapolation scenarios, educated prime-aged workers tend to work more, while less-educated prime-aged workers tend to work less. We think it is unlikely that the trends continue at the earlier pace captured by the extrapolation column in Table 3; but, qualitatively, the trends might continue in the same direction. That would suggest that it is plausible labor quality grows a little faster than 0.15 percent per year (the pace in the previous paragraph, where we extrapolate EPOP trends only for those over 55). The effect on hours would be small, since the trends somewhat offset.

Thus, in the longer run, a projection of 0.10-0.25 percent growth in labor quality and perhaps 0.4 to 0.55 for hours is a plausible judgmental baseline.

6 Conclusion

Historically, rising labor quality was an important source of growth in U.S. GDP per hour. Going forward, this source of growth is likely to slow markedly. Indeed, our preferred forecast is that, in the longer run (2022-25), labor quality is likely to rise in the range of 0.10 to 0.25 percent per year. This implies that growth in quality-adjusted hours in the range of 0.5 to 0.8 percent per year is plausible, with a range of 0.7 to 0.8 percent per year seeming perhaps most likely. To see a faster pace of labor-quality growth, closer to its historical average pace, would require a renewed, and sustained, upward trend in educational attainment. In a typical macro model, the slowdown in labor-quality growth passes through one-for-one to slower growth in productivity and GDP.

In the 20th century, the main driver of labor-quality growth was rising educational attainment (Fernald & Jones, 2014; Ho & Jorgenson, 1999). In contrast, in our empirical estimates and forecasts for the 21st century, we find a very different source of labor-quality growth: The diverging trends in employment rates for workers of different skills. Since 2002, employment rates for more educated, older individuals have risen whereas employment rates for less educated, younger individuals have fallen. These diverging trends explain why previous forecasts that labor quality would plateau (Aaronson & Sullivan, 2001) went awry—from 2002-13, labor quality turned out to grow at a pace even faster than it did in the second half of the 20th century because of changing employment dynamics.

These forecast-misses point to a broader lesson: It is essential to jointly examine growth in hours and growth in labor quality. Labor quality and hours are strongly negatively correlated in the short run, which implies that quality-adjusted hours are less variable than either quality or hours alone. Looking at hours or labor-quality growth independently can lead to inaccurate projections of potential output growth.

Going forward, movements in employment-to-population rates for different worker types continue to be central to how future labor quality will evolve. In the medium run (2015-22), an important source of uncertainty is whether the diverging employment-rate movements seen since 2007 are cyclical or structural. If employment rates (based on age and education) revert to 2007 levels, then growth in labor quality is likely to be negative as lower-skilled workers return to employment. In this case, labor quality in the next few years will at least partially offset the strong growth since 2007. In contrast, if the changes since 2007 are structural, then growth in labor quality will be considerably stronger, albeit not at rates seen historically.

But, once again, these alternative paths illustrate the importance of jointly modeling labor quality and hours. Quality-adjusted labor input turns out to grow at remarkably similar rates in the scenarios where employment-to-population rates revert to 2007 values (cyclical), or remain at 2013 values (structural), leaving overall output growth unchanged.

References

Aaronson, Daniel, & Sullivan, Daniel. 2001. *Growth in worker quality*. Economic Perspectives Q IV. Federal Reserve Bank of Chicago.

Aaronson, Stephanie, Cajner, Tomaz, Fallick, Bruce C., Galbis-Reig, Felix, Smith, Christopher, & Wascher, William L. 2014 (Sept.). *Labor Force Participation: Recent Developments and Future Prospects*. Finance and Economics Discussion Series 2014-64. Board of Governors of the Federal Reserve System (U.S.).

Altonji, Joseph G., & Blank, Rebecca M. 1999. Chapter 48 race and gender in the labor market. Handbook of Labor Economics, vol. 3, Part C. Elsevier.

Barrow, Lisa, & Davis, Jonathan. 2012 (October). *The upside of down: Postsecondary enrollment in the great recession*. Research Report 36. Federal Reserve Bank of Chicago.

Bertrand, Marianne, & Mullainathan, Sendhil. 2003 (July). *Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination*. Working Paper 9873. National Bureau of Economic Research.

Bils, Mark J. 1985. Real Wages over the Business Cycle: Evidence from Panel Data. *Journal of political economy*, **93**(4), 666–89.

Blaug, Mark. 1985. Where are we now in the economics of education? *Economics of education review*, **4**(1), 17 – 28.

Boeri, Tito, & van Ours, Jan. 2013. *The Economics of Imperfect Labor Markets: Second Edition*. Economics Books, vol. 1, no. 10142. Princeton University Press.

Bureau of Labor Statistics. 1993. *Labor composition and u.s. productivity growth: 1948-90*.

Bureau of Labor Statistics. 2015a. *Changes in the composition of labor for bls multifactor productivity*. http://www.bls.gov/mfp/mprlabor.pdf [Last accessed:November 20, 2017].

Bureau of Labor Statistics. 2015b. *Multifactor productivity*. http://www.bls.gov/mfp/ [last accessed: November 20, 2017].

Burtless, Gary. 2013 (September). *Can educational attainment explain the rise in labor force participation at older ages?* Tech. rept. 13-13. Center for Retirement Research at Boston College.

Charles, Kerwin Kofi, Hurst, Erik, & Notowidigdo, Matthew J. 2015 (September). *Housing booms and busts, labor market opportunities, and college attendance*. Working Paper 21587. National Bureau of Economic Research.

Congressional Budget Office. 2015. *An update to the budget and economic outlook: 2015 to 2025*.

Deming, David J. 2015 (August). *The growing importance of social skills in the labor market*. Working Paper 21473. National Bureau of Economic Research.

Dennett, Julia, & Modestino, Alicia Sasser. 2013 (December). *Uncertain futures? youth attachment to the labor market in the united states and new england*. Research Report 13-3. Federal Reserve Bank of Boston.

Dickens, William, & Katz, Lawrence F. 1987. Inter-industry wage differences and industry characteristics. *In: Unemployment and the structure of labor markets, basil*. Blackwell.

Diewert, W. E. 1978. Superlative index numbers and consistency in aggregation. *Econometrica*, **46**(4), pp. 883–900.

Feenberg, Daniel, & Roth, Jean. 2007. *Cps labor extracts: 1979-2006*. Tech. rept. National Bureau of Economic Research.

Feenstra, Robert C., Inklaar, Robert, & Timmer, Marcel P. 2015. The next generation of the penn world table. *American economic review*, forthcoming.

Fernald, John G. 2015. *Total factor productivity.* http://www.frbsf.org/economic-research/total-factor-productivity-tfp [last accessed: November 20, 2017].

Fernald, John G., & Jones, Charles I. 2014. The future of us economic growth. *American economic review*, **104**(5), 44–49.

Ferraro, Domenico. 2014. *The Asymmetric Cyclical Behavior of the U.S. Labor Market.* Tech. rept.

Gibbons, Robert, Katz, Lawrence F., Lemieux, Thomas, & Parent, Daniel. 2005. Comparative Advantage, Learning, and Sectoral Wage Determination. *Journal of labor economics*, **23**(4), 681–724.

Goldin, Claudia Dale, & Katz, Lawrence F. 2009. *The race between education and technology.* Belknap of the Harvard University Press, Cambridge MA.

Gollop, Frank, & Jorgenson, Dale. 1983. Sectoral Measures of Labor Cost for the United States, 1948-1978. *Pages 185–236 of:* Triplett, Jack E. (ed), *The measurement of labor cost.* University of Chicago Press.

Hamermesh, Daniel S, & Biddle, Jeff E. 1994. Beauty and the Labor Market. *American economic review*, **84**(5), 1174–94.

Heckman, James J., Stixrud, Jora, & Urzua, Sergio. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of labor economics*, **24**(3), 411–482.

Heckman, James J., Lochner, Lance J., & Todd, Petra E. 2008. Earnings Functions and Rates of Return. *Journal of human capital*, **2**(1), 1–31.

Hellerstein, Judith K., & Neumark, David. 1995. Are Earnings Profiles Steeper Than Productivity Profiles? Evidence from Israeli Firm-Level Data. *Journal of human resources*, **30**(1), 89–112.

Hellerstein, Judith K., & Neumark, David. 2008. Workplace Segregation in the United States: Race, Ethnicity, and Skill. *The review of economics and statistics*, **90**(3), 459–477.

Hellerstein, Judith K., Neumark, David, & Troske, Kenneth R. 2002. Market Forces and Sex Discrimination. *Journal of human resources*, **37**(2), 353–380.

Ho, Mun S., & Jorgenson, Dale W. 1999. *The quality of the u.s. work force, 1948-95*. Tech. rept. Harvard University.

Hungerford, Thomas, & Solon, Gary. 1987. Sheepskin Effects in the Returns to Education. *The review of economics and statistics*, **69**(1), 175–77.

Johnson, Matthew T. 2013. The impact of business cycle fluctuations on graduate school enrollment. *Economics of education review*, **34**, 122 – 134.

Jones, Patricia. 2001. Are educated workers really more productive? *Journal of development economics*, **64**(1), 57–79.

Jorgenson, D. W., & Griliches, Z. 1967. The explanation of productivity change. *The review of economic studies*, **34**(3), pp. 249–283.

Jorgenson, Dale W., Gollop, Frank M., & Fraumeni, Barbara M. 1987. *Productivity and u.s. economic growth*. Harvard University Press. Reprinted- Universe, 1999.

Jorgenson, Dale W., Ho, Mun S., & Samuels, Jon D. 2014. What will revive u.s. economic growth? lessons from a prototype industry-level production account for the united states. *Journal of policy modeling*, **36**(4), 674 – 691. Rapid Growth or Stagnation in the U.S. and World Economy?

Jorgenson, Dale W., Ho, Mun S., & Samuels, Jon D. 2015. Education, participation, and the revival of u.s. economic growth. October.

Krueger, Alan B., & Summers, Lawrence H. 1988. Efficiency wages and the inter-industry wage structure. *Econometrica*, **56**(2), pp. 259–293.

Lang, Kevin. 2015. Racial realism: A review essay on john skrentny's after civil rights. *Journal of economic literature*, **53**(2), 351–59.

Lemieux, Thomas. 2006. The mincer equation thirty years after schooling, experience, and earnings. *Pages 127–145 of:* Grossbard, Shoshana (ed), *Jacob mincer a pioneer of modern labor economics*. Springer US.

Light, Audrey, & Ureta, Manuelita. 1995. Early-career work experience and gender wage differentials. *Journal of labor economics*, **13**(1), pp. 121–154.

Mincer, Jacob A. 1974. *Schooling, experience, and earnings*. National Bureau of Economic Research, Inc.

Oaxaca, Ronald, & Ransom, Michael. 1994. On discrimination and the decomposition of wage differentials. *Journal of econometrics*, **61**(1), 5–21.

O'Mahony, Mary, & Timmer, Marcel P. 2009. Output, input and productivity measures at the industry level: The eu klems database*. *The economic journal*, **119**(538), F374–F403.

Pager, Devah, Western, Bruce, & Bonikowski, Bart. 2009. Discrimination in a low-wage labor market: A field experiment. *American sociological review*, **74**, 777–99.

Psacharopoulos, George, & Patrinos, Harry Anthony. 2004. Returns to investment in education: a further update. *Education economics*, **12**(2), 111–134.

Sherk, James. 2013 (September). *Not looking for work: Why labor force participation has fallen during the recession*. Tech. rept. 2722. Herritage Foundation.

Silva, Olmo. 2007. The jack-of-all-trades entrepreneur: Innate talent or acquired skill? *Economics letters*, **97**(2), 118–123.

Skrentny, John D. 2013. *After civil rights: Racial realism in the new american workplace*. Princeton University Press, Princeton.

SOLON, GARY, BARSKY, ROBERT, & PARKER, JONATHAN A. 1994. Measuring the cyclicality of real wages: How important is composition bias. *The quarterly journal of economics*, **109**(1), pp. 1–25.

VAN ARK, BART, & ERUMBAN, ABDUL. 2015. *Productivity brief 2015.*

WEISS, ANDREW. 1995. Human capital vs. signalling explanations of wages. *Journal of economic perspectives*, **9**(4), 133–154.

WYATT, IAN D. 2010. Evaluating the 1996 - 2006 employment projections. *Monthly labor review*, September, 33–69.

ZOGHI, CINDY. 2010. Measuring labor composition: A comparison of alternate methodologies. *Pages 457–485 of:* ABRAHAM, KATHARINE G., SPLETZER, JAMES R., & HARPER, MICHAEL (eds), *Labor in the new economy.* University of Chicago Press.

A Data details

A.1 ACS, CPS-ASEC, and CPS-ORG

To verify the robustness of our results, we calculate them for three commonly used U.S. datasets that each allow for the construction of measures of labor-quality growth. The first is the American Community Survey (ACS), which is a smaller, annual version of the decennial census and collects a relatively narrow range of demographic and socioeconomic data on a sample of about 1 percent of the population (approximately 3 million individuals) each year.[40] The second, the CPS-ORG, consists of the the outgoing rotation groups from the Current Population Survey (CPS). This is the quarter of CPS respondents that are asked about their earnings and income in any given month. This results in an annual sample of about 135,000 individuals. The final dataset, CPS-ASEC, is the Annual Social and Economic Supplement to the Current Population Survey, also known as the March supplement. It contains annual earnings and income data from the full March CPS sample (70,000 individuals). Though based on different samples and sampling methods, each of the datasets allows for the construction of similar hourly wages, as well as the six variables of education, age, sex, race/ethnicity, industry, and occupation, that are our main focus.

For each dataset we construct the *sample* of workers to cover those in the civilian noninstitutional population ages 16+ that are employed in the private business sector (specifically, excluding anyone with self-employment or government employment earnings) and have both positive earnings and positive hours. The sample period is 2002-2013, because that is the period for which we have a consistent set of occupation and industry crosswalks and data from all three datasets.[41]

We define *wages* as hourly wages. Wages are constructed in slightly different ways in each

[40]The sample of the ACS has been expanded twice and has only been a 1 percent sample of the population since 2006. In its first year, 2000, the sample was just under 400,000 individuals and between 2001 and 2005 the sample was slightly over 1 million

[41]In principle, the CPS-ASEC is available starting in 1962 on and the CPS-ORG from 1979 on if industry and occupation are omitted or approximate crosswalks are used. The ACS is available from 2000 on without any need for adjustments.

of the datasets because of differences in reference period and questions asked. In the CPS-ORG we use the hourly wage as constructed in the National Bureau of Economic Researchs CPS Labor Extracts (Feenberg & Roth, 2007). For the CPS-ASEC and ACS we define hourly wages as total annual earnings divided by the product of usual hours worked per week and weeks worked per year.[42] All wages are deflated into real 2005 dollars using the Consumer Price Index for All Urban Consumers, and wages exclude self-employment, self-owned business, and farm income.

A.2 Projections of educational attainment and employment

The Census Bureau's 2000 National Population Projections provides projections of the age, gender and race/ethnicity distribution of the population, but to forecast labor quality we need to further break these cells down by educational attainment and employment rates. To do so we follow a methodology similar to that used by Aaronson & Sullivan (2001)—our primary adjustments are that we use five race/ethnicity categories instead of four and we define employment more narrowly as being employed exclusively in the private business sector to match the sample selection stated in Section A.1. Given that the methodology is substantively unchanged, this section is largely a restatement of Box 1 from (Aaronson & Sullivan, 2001, p. 65).

Let $p^j_{it} = \mathrm{P}\left[y_{it} = j\right]$ for $j = 1, ..., 5$ by the probability that individual i in year t has educational attainment j, where the five levels of attainment are less than high school, high school graduate (including GEDs), some college (including associates degree holders), college graduates (Bachelor's), and post-graduates, and let $q^j_{it} = \mathrm{P}\left[y_{it} \geq j | y_{it} \geq j-1\right]$ for $j = 2, ..., 5$ be the probability of attaining education j given that the individual has completed the "prerequisite"

[42]In 2008 the ACS switched from collecting weeks worked as a continuous to a categorical value (13 weeks or less, 14-26 weeks, 27-39 weeks, 40-47 weeks, 48 or 49 weeks, and 50-52 weeks). Prior to 2008 the distribution of weeks worked within those ranges was remarkably stable over time, so we imputed a continuous value of weeks worked using the pre-2008 mean of people reporting weeks worked within a given range. We also tested using a more complex regression model on demographic characteristics to impute weeks worked but found that it gave little more variation or precision in predicted weeks worked than using the pre-2008 mean. The same approach is used by the BLS for pre-1975 data, which has the same issue (Bureau of Labor Statistics, 1993, p. 77).

education (e.g. for $j = 4$ this is the probability of an individual having completed college given that they have completed some college). We predict $\hat{q}_{it}^{j}$ using a logistic regression of the form

$$\log \frac{q_{it}^{j}}{1 - q_{it}^{j}} = \sum_{a} D_{it}^{a} \alpha_{ja} + \sum_{b} D_{it}^{b} \beta_{jb} + \mathbf{x}_{it} \gamma_{j}, \tag{15}$$

$$\text{and } \hat{q}_{ab}^{j} = \frac{\exp\left(\alpha_{ja} + \beta_{jb}\right)}{1 + \exp\left(\alpha_{ja} + \beta_{jb}\right)} \tag{16}$$

where D_{it}^{a} and D_{it}^{b} are dummies for being age a and born in year b, and x_{it} is a vector of control variables. From $\hat{q}_{ab}^{j}$ it is possible to calculate $\hat{p}_{ab}^{j} = \prod_{k=2}^{j} \hat{q}_{ab}^{k} \left(1 - \hat{q}_{ab}^{j+1}\right)$, which can be interpretted as the predicted share of people born in year b with education j at age a or, since age, year, and birth year are perfectly collinear, the predicted share of people of age a with education j in year $b+a$. The models for education level j are estimated on the sample of people with at least $j - 1$ education and who are above an education-level specific age threshold.[43] For the projections for the forecast error decomposition exercises in Section 5.2 the models are estimated on the CPS-ORGs from 1992 through 1999, the same period Aaronson & Sullivan used for their forecasts.[44]

The idea behind these models is that educational attainment follows some sort of lifecycle pattern, with the probability of completing a certain level of education increasing rapidly for people under 30 and then more gradually for those who are older. This lifecycle pattern is assumed to be the same for different cohorts, but cohorts born in different years are allowed to have uniformly higher or lower log-odds of completing a given level of education. For high school, some college, and college levels of education the model is estimated separately for each of ten gender-race-ethnicity combinations without any control variables ($\mathbf{x}_{it}$). For post-graduates some of the gender-race-ethnicity samples become quite small, so the model is estimated separately for men and women with race/ethnicity dummies included as con-

[43]The thresholds are 18 for high school, 19 for some college, 22 for college, and 26 for post-graduate.

[44]Ideally this would have been estimated on ACS data to ensure consistency between these projection models and the log-wage regression. However, in order to distinguish age and cohort effects the projection model must be estimated on multiple years of data. Since there is no pre-2000 data for the ACS, this forces us to rely on another dataset to construct the education and employment projections.

trols. The estimated model is then used to predict the fraction of individuals with each level of educational attainment based on the Census Bureau projections of the age, gender, and race/ethnicity distribution of the population.

The projection model is only able to estimate birth year coefficients (β_{jb}) for birth years that are observed in the sample. However, some birth years that are too young to be observed in the sample will be old enough to be in sample by later years of the projections — a child born in 2000 is to young to be in any of our current samples, but by 2025 they will be 25 years old and of critical importance to our forecasts. Therefore we define these unobserved cohort coefficients by a linear extrapolation using the last 15 birth year coefficients (not including the most recent).[45] In effect, this approach extrapolates recent trends in educational attainment into the future.

This process yields projections of the population distribution of age and educational attainment, the key variables for our baseline Mincer specification. However, to construct our forecast of labor quality we must also project the EPOP rates for these worker types. Our EPOP projection model is identical to the educational attainment projection model, except educational attainment is added as a control variable. Rather than using the standard BLS definition of employment we define employment as being employed exclusively in the private business sector — this makes our definition consistent with the sample selection used to construct our labor quality measures.

A.3 Projection scenarios for educational attainment and employment

The Fisher Ideal index does not have the circularity property, so the labor-quality growth calculated from comparing a target year to a base year is not necessarily the same as the growth calculated from cumulated year-over-year changes. However, this is not true for the labor-quality growth projections because our assumption that the log-wage regression coefficients

[45] The most recent coefficient is omitted because it is based on just one year of observations, making the sample size quite small.

are constant over time means that the Fisher Ideal index collapses into the Laspeyres index, which does have the circularity property. This allows us to construct alternative projection scenarios based on assumptions about the education and employment distribution in a target year alone, without having to make assumptions about the path of educational attainment or EPOP between now and then. Therefore our projection scenarios discussed in 5 are based on the Census Bureau age projections for the years 2022 and 2025 and the education and employment assumptions described below.

Baseline labor quality in 2015 is calculated by applying the empirical 2013 education and employment distributions by age from the ACS to the Census Bureau population projections for 2015. That is, we calculate the share of 25 year olds that have a college degree, the share of 25 year olds with college degrees that are employed, and then combine that with the Census projection of the number of 25 year olds in 2015 to estimate the number college educated 25 year old workers in 2015. This same baseline distribution is used in all nine labor quality projections.[46]

Education scenarios All three education scenarios assume that the educational distribution for those over 30 will stay the same as they age. For example, the educational attainment of 52 year olds in 2025 is assumed to be the same as that of 40 year olds in 2013 (the most recent year in our data). Although non-traditional educational attainment, differential mortality rates, and immigration make it unlikely that this assumption strictly holds, those forces are marginal enough that they are unlikely to cause substantial deviations. Where the scenarios differ is in their assumptions on the educational attainment of 1) people 30 and under in the projection year (the "young group"), and 2) the educational attainment of people under 30 in 2013 that will be over 30 in the projection year (e.g. 31-42 year olds in 2025; the "middle group"). The educational attainment of the *young group*, which was in middle school or below during the Great Recession and thus unlikely to have been driven by cyclical factors — their educational attainment can be thought of as representing a "normal" level. Unlike the young

[46] Note that the differences between the growth rates in the different scenarios is completely independent of the baseline, since we report log-growth.

group, those in the middle group were making critical education decisions (such as whether to drop out of high school or college and whether to enroll in college or grad school) during the Great Recession and its aftermath. Therefore, if "educational sheltering" has been a strong force during and after the Great Recession, as posited by Barrow & Davis (2012), Sherk (2013), and Johnson (2013), then their attainment may deviate from the norm.

Revert to pre-crisis educational plateau The first education scenario assumes that the educational attainment of young people reverts to its pre-crisis levels. This reflects the possibility that the uptick in enrollment and graduation rates over the past several years is simply a temporary cyclical effect of the Great Recession. For the young group, this scenario assumes they will have the same distribution of educational attainment as people of the same age in 2007.[47] For those in the middle age group, whose attainment may have been increased by "educational sheltering" effects, this scenario assumes that they will either have the educational attainment of someone that age in 2007 or their current educational attainment, whichever is higher. That is, they will have at least the educational attainment that would have been expected of them before the recession, and they may have a little more if the recession encouraged them to stay in school. Specifically, let $\hat{q}_a^j$ be the probability of someone with age a having at least education j in 2007, let $\tilde{q}_{a-12}^j$ be the probability of someone that will be age a in 2025 having at least education j in 2013, and let $q_a^j = \max\left(\hat{q}_a^j, \tilde{q}_{a-12}^j\right)$. Then for this scenario the share of people of age $a = 31, ..., 42$ with education j will be $p_a^j = q_a^j - q_a^{j+1}$. This is the same for 2022, except using $\tilde{q}_{a-9}^j$.

Persist at 2013 educational plateau The second scenario assumes that the educational attainment of young people persists at its 2013 rate, reflecting the possibility that there was a step increase in educational attainment over the past several years but that attainment has once again reached a plateau. This scenario assumes that people in the young group will have the same distribution of educational attainment as someone of the same age in 2013. For the

[47] Recent research suggests that the housing boom depressed educational attainment by providing good job opportunities to low skill workers, in which case the educational attainment patterns from the boom years would be unusually low Charles *et al.* (2015). That would suggest that this may be a particularly pessimistic implementation of this "cyclical uptick" hypothesis. However, we believe this is still a useful scenario to consider as it provides a plausible worst-case scenario for education trends.

middle group we have to account for the fact that the increase in educational attainment was gradual and had not fully propagated through for those over 30 but people under 30 will often go on to further education, meaning that there is no clear baseline group. To get a baseline for this group we calculate the probability q^j in 2013 of completing at least education j for the five year age group that are young enough to have experienced a sheltering effect but old enough that we would expect them to have completed that level of education already.[48] For this scenario we define the expected educational attainment distribution of the middle group as $p^j = q^j - q^{j+1}$.

Extrapolate 2007-2013 trends in education The final scenario assumes that the uptick in educational attainment over the past several years represents a resumed upward trend in education attainment rather than a temporary cyclical boost or a one-off step increase. Age-specific time trends in educational attainment are estimated from logistic regressions of the form

$$\log \frac{q_{it}^j}{1 - q_{it}^j} = \sum_a \left[year \cdot D_{it}^a \beta_a + D_{it}^a \gamma_a \right]. \tag{17}$$

As in Section A.2, these logits are estimated on the population of people with education $j-1$ or higher, and they are estimated on 2007-2013 data. Let q_a^{2013} be the probability that a person of age a had education j or higher in 2013. Then this scenario assumes that the probability of having at least education j at age a in 2025 is the probability of having education j in 2013 plus the age-specific time trend — that is, they have probability $q_a^j = \text{invlogit}\left[\text{logit}\left(q_a^{2013}\right) + 12 \cdot \beta_a\right]$ of having at least education j at age a in 2025. As in the other cases, we then recover the share of people with education j at age a in 2025 as $p_a^j = q_a^j - q_a^{j+1}$. This is the same for 2022 except $q_a^j = \text{invlogit}\left[\text{logit}\left(q_a^{2013}\right) + 9 \cdot \beta_a\right]$

Employment scenarios The employment scenarios are much more straightforward to construct because there is little to no need to keep track of the stock of employment — the fact that 85 percent of 29 year old college graduates were employed in 2013 does not impose particularly binding constraints on our assumptions about the EPOP rate of 41 year old college

[48]For high school we us 19-23, for some college we use 23-27, for college we use 25-29, and for post-graduate we use 30-34. Less than high school is the residual category.

graduates in 2025. Therefore, our two baseline employment scenarios simply assume that the EPOP rates for specific age-education groups in the projection year will be the same as in some other base year. For the *revert to pre-crisis EPOPs* scenario we assume that the probability of a person of age a with education j being employed in the projection year is the same as it would have been in 2007.[49] This scenario corresponds to the view that the entire decline in EPOP rates for specific age-education groups is cyclical.[50] The second employment scenario is the inverse of this and assumes that the entire change in the EPOP rates of specific age-education groups is structural and will *persist at 2013 EPOPs*.[51]

Extrapolated 2002-2007 structural trends in EPOPs The final scenario extrapolates certain pre-crisis trends in employment patterns out to the projection year. In particular, it extrapolates the declining EPOP rates of young people (with heterogeneity across education groups), the increasing EPOP rates of older people (particularly the more educated), and the widening gap between the EPOP rates of more and less educated working age people Dennett & Modestino (2013); Burtless (2013); Aaronson *et al.* (2014). Given that we have pre-selected the trends that are extrapolated, this scenario can be accused of cherry-picking. We do not deny that vulnerability, and we do not intend this scenario to be understood as a probable outcome. Again, these scenarios are primarily intended to illustrate the mechanics of labor-quality growth and what factors are most critical to the setting expectations about future labor-quality growth in the U.S., as well as to impose certain bounds on plausible forecasts of labor-quality growth.

To implement the third employment scenario we follow an approach similar to that in the education trends scenario above. To extract age and education-specific time trends in

[49] As with the first education scenario, this may be an extreme assumption on what the pre-crisis norm was — if the housing boom boosted EPOP rates to abnormal levels, then this scenario overstates the baseline EPOP rates. Similar to the education case, we believe this remains a useful scenario to consider as it illustrates a sort of best-case scenario for employment rates.

[50] This still allows for a demographically (or educationally) driven structural decline in the aggregate employment-to-population ratio. What it does not allow for is a structural decline (or increase) in EPOP for specific age-education groups. For example, it does not allow for a structural decline in students working part-time, or a structural increase in older people staying employed past the traditional retirement age.

[51] This scenario may be too pessimistic in that the labor market has clearly continued to improve since 2013. Once more recent ACS data becomes available we will revise this scenario to reflect the most recent year of data available. However, this once again provides a sort of outlier case with unusually low EPOP rates.

employment we run the following logistic regression on the sample of 16-24 and 55-69 year olds over the 2002-2007 period

$$\log \frac{p_{it}}{1-p_{it}} = \sum_a \sum_j \left[year \cdot D^a_{it} \cdot D^j_{it} \beta_{aj} + D^a_{it} \cdot D^j_{it} \gamma_{aj} \right]. \tag{18}$$

and to extract education-specific time trends in employment among prime age workers we run the following logistic regression on 25-54 year olds over the same period

$$\log \frac{p_{it}}{1-p_{it}} = \sum_j \left[year \cdot D^j_{it} \beta_j + D^j_{it} \gamma_j \right]. \tag{19}$$

where p_{it} is the probability of individual i being employed in year t, D^a_{it} is an indicator for being age a, and D^j_{it} is an indicator for having education j. Let p^{2007}_{aj} be the probability that a person of age a with education j was employed in 2007. Then this employment trends scenario assumes that the probability of a person of age a and education j being employed in 2025 is the probability of being employed in 2007 plus the relevant age- and education-specific time trend. For 16-24 and 55-69 year olds this is $p_{aj} = \text{invlogit}\left[\text{logit}\left(p^{2007}_{aj}\right) + 18 \cdot \beta_{aj}\right]$ and for 25-54 year olds it is $p_{aj} = \text{invlogit}\left[\text{logit}\left(p^{2007}_{aj}\right) + 18 \cdot \beta_j\right]$.[52] The 2022 projection is the same, except the βs are multiplied by 15 instead.

B Robustness checks

In this section of the appendix we present additional results that illustrate that our qualitative results are unchanged when we change some of the underlying assumptions, specifications, and across datasets.

[52]For people 70 and over there is no time-trend added in and their EPOP rate in the projection year is assumed to be the same as it was in 2007.

B.1 Adding control variables to the baseline Mincer regression

Throughout the main text we limited ourselves to parsimonious baseline Mincer regression. However, prior implementations of such specifications have included control variables to ensure that only productivity-induced wage differentials are reflected in the estimated wages (Aaronson & Sullivan, 2001; Bureau of Labor Statistics, 1993). Here we consider the robustness of our results to including standard control variables, such as part-time status, marital status, veterans status, race, and geographic location.

As discussed in subsection 3.2, it is critical that the variables included in the labor quality specification ($\mathbf{x}_j$) be 1) correlated with wages, and 2) that the correlation is driven by differentials in the marginal product of labor. A desirable property of a regression-based framework like (9) is that it allows for the inclusion of control variables, $\mathbf{z}_j$, that may be correlated with both individual wages, w_j, and the variables meant to quantify marginal product differentials, $\mathbf{x}_j$. The resulting generalized regression framework is

$$w_j = \mathbf{x}_j^{'}\boldsymbol{\beta} + \mathbf{z}_j^{'}\boldsymbol{\gamma} + \varepsilon_j. \tag{20}$$

Because we attribute only the part of wage variations explained by the variables in $\mathbf{x}_j$ to marginal product differentials, we impute the log marginal product of a worker as $\mathbf{x}_j^{'}\hat{\boldsymbol{\beta}}$. The inclusion of these control variables does not alter our definitions of σ_j and $\tilde{\sigma}_p$. They continue to be based on $\mathbf{x}_j$ and $\hat{\boldsymbol{\beta}}$.

What is less clear is the appropriate measure of fit when considering a regression with controls. Consider, for example, a set of controls $\mathbf{z}$ that predict wages ($\boldsymbol{\gamma} \neq \mathbf{0}$) but for which the correlation between any element x of $\mathbf{x}$ and any element z of $\mathbf{z}$ iz zero (corr $(x,z) = 0$). In this case the regression $\bar{R}^2$ will increase, making the specification appear more appealing than the version without $\mathbf{z}$ despite the fact that substantive components of the regression, $\mathbf{x}$ and $\hat{\boldsymbol{\beta}}$, remain unchanged.[53] An alternative approach would be to consider the partial R-squared with respect to $\mathbf{x}$, $\bar{r}_x^2$. However, then maximizing $\bar{r}_x^2$ is not necessarily desirable. For example,

[53]The standard errors will also slightly increase because of the loss of degrees of freedom.

if the association between a control variable z and the core variables $\mathbf{x}$ has the same sign as the association between z and wages w, then the $\bar{r}_x^2$ will decline in the regression with z. But the $\bar{r}_x^2$ declined precisely because z had been a source of omitted variable bias and we are now controling for that.

Ultimately, the selection of $\mathbf{z}$ operates on an orthogonal basis from the selection of $\mathbf{x}$ in a *properly controled regression*. As discussed in subsection 3.2, the desirability of higher $\bar{R}^2$ is entirely conditional on the assumption that $\hat{W}_j \equiv \exp\left(\mathbf{x}_j'\boldsymbol{\beta}\right) = c \cdot W$ — if any omitted variable bias is loaded onto $\boldsymbol{\beta}$ then this assumption is violated. In principle, this means that one should optimize $\mathbf{z}$ for each separate specification of $\mathbf{x}$, at which point we can compare the $\bar{r}_x^2$ of the controlled regressions as is done in subsection 3.2.

Rather than undertaking this highly multidimensional and daunting task, we consider whether it is likely to be of first-order importance to any of our results. Specifically, we consider the impact of including two standard sets of controls in our *baseline* Mincer and *baseline+occupation* specifications. The first set of controls is a set of indicators for part-time employment, marriage, and race, which are the controls included in the specification used by Aaronson & Sullivan (2001).[54] The second set of controls is similar Bureau of Labor Statistics (1993) and includes indicators for part-time employment, veteran status, and which Census division the individual lives in.[55]

Figure B.1a, which is comparable to Figure 2, plots the adjusted R-squared ($\bar{R}^2$) against the 80th percentile standard error of the predictions ($\tilde{\sigma}_{80}$). This shows that, as expected, the inclusion of the additional variables increases both $\bar{R}^2$ and $\tilde{\sigma}_{80}$. The Aaronson & Sullivan (2001) controls improve the fit slightly more and increase imprecision slightly less than the Bureau of Labor Statistics (1993) controls. However, as can be seen in Figure B.1b, there is almost no change in the partial R-squared with respect to $\mathbf{x}$, suggesting that either the control variables are not a significant source of omitted variable bias or that the biases they induce

[54]We use five race/ethnicity indicators where they used four race indicators—we distinguish Hispanics from non-Hispanic whites, blacks, Asians, and other.

[55]The Bureau of Labor Statistics (1993) specification also includes indicators for whether the individual is in a central city or balance of a SMSA/CBSA or in a rural area, which we omit.

balance out on average. This suggests that the impact of including these control variables on measured labor-quality growth is likely to be quite limited. This is confirmed in Figure B.1c, which plots the resulting labor-quality indices. The indices with the Bureau of Labor Statistics (1993) controls are virtually indistinguishable from their uncontrolled counterparts, while the Aaronson & Sullivan (2001) controls appear to exert a modest negative drag on labor quality, on the order of a couple hundredths of a percentage point per year. These results suggest to us that control variables are not of first-order importance in measuring or forecasting labor-quality growth.

B.2 Additional results for CPS-ORG and CPS-ASEC

The majority of the results presented in the main text were produced using data from the ACS, but it is possible to conduct the same exercises using both the CPS-ORG and CPS-ASEC. In this section we evaluate the robustness of key results from the main text in these alternative data sources. All of the qualitative results hold up, with some minor differences in magnitude.

Figures B.2a and B.2b plot the adjusted R-squared ($\bar{R}^2$) against the 80th percentile standard error ($\tilde{\sigma}_{80}$) of the same specifications considered in Section 3 and Figure 2 for the CPS-ORG and CPS-ASEC, respectively. As we note in the main text, the large sample size of the ACS is relatively favorable to stratum-based specifications: with the CPS datasets, which are more than an order of magnitude smaller, the standard errors are an order of magnitude higher. In fact, the CPS-ASEC is small enough that for some of the more granular specifications more than 20 percent of the observations are in single-observation cells with infinite standard errors, leaving $\tilde{\sigma}_{80}$ undefined.[56] However, the tradeoff between fit and precision is still clearly visible, the age and education or age, education and occupation specifications strike a reasonable balance between fit and precision, and the *baseline* and *baseline+occupation* Mincer specifications dominate the stratum-based specifications. In short, the results are entirely consistent with our findings from the ACS.

[56]We substitute the highest observed percentile standard error, which is the source of the vertical lines in the upper right region of the figure.

Figures B.3a and B.3b plot the 2002-2013 labor-quality indices presented in Section 4 and Figure 3 for the CPS-ORG and CPS-ASEC, respectively. Once again the results are quite similar to those found in the ACS. The overfit specification (which includes all six variables considered in Section 3) and the underfit specification (which includes all variables except age and education) both show very little labor-quality growth over the 2000s. In the case of the CPS-ASEC the overfit and underfit specifications are quite noisy, with implausible jumps and changes in direction.

All of the age and education specifications (with or without occupation) and the *baseline* and *baseline+occupation* Mincer specifications, by contrast, are clustered together and quite similar to the ACS results in Figure 3a, although the CPS-ORG specifications show about 0.5 percent less cumulative labor-quality growth by 2013. The CPS-ORG results are also slightly more closely clustered than those for the other two datasets. This may be because hourly wages are measured directly in the CPS-ORG, whereas in the CPS-ASEC and ACS hourly wages are noisily derived from annual earnings divided by the product of usual weekly hours and weeks worked per year.

One notable difference is that ACS indices show an unexpected decline in labor quality between 2005 and 2006, while the CPS-based indices do not. This appears to be an data artifact induced by the tripling of the ACS sample size in 2006. A similar jump occurs when we calculate labor-quality growth between 2000 and 2001 in the ACS (not reported), and there also appears to be a slight tick in the 2012-2013 period for the ASEC, which saw a sample size change in 2013. Why changing sample size can induce these sharp adjustments in labor quality is somewhat unclear and bears more careful investigation.

Figures B.4a and B.4b plot the 2002-2013 counterfactual labor-quality indices presented in Section 4 and Figure 4 for the CPS-ORG and CPS-ASEC, respectively. The results are qualitatively the same in the CPS datasets as in the ACS, with changes in average hours worked contributing relatively little to labor-quality growth while changes in population demographics and demographic-specific EPOP rates both contributing significantly. However, there are two quantitative differences.

First, average hours appear to matter less in the CPS datasets. This is likely due to the fact that the ACS uses a categorical measure weeks worked after 2008, which induces additional noise in the measurement of average hours relative to the other two datasets. This is consistent with the fact that hours only make a significant difference after 2008. The relative unimportance of hours further strengthens our conviction that projecting average hours is not critical to a labor quality forecast and that attempts to do so are likely to introduce as much forecast error as they address.

Second, whereas for the ACS the evolution of EPOP rates induced more labor-quality growth than changing demographics (compare the yellow line to the red), in the two CPS datasets the contributions of employment and demographics are almost equal. Additionally, the contribution of EPOP rates, reflected in the yellow lines, is more obviously cyclical for the two CPS datasets — it is virtually flat before and after the Great Recession, with a substantial step increase during the Great Recession. The ACS, by contrast, shows significant labor-quality growth from EPOP rates even before the Great Recession, with the Great Recession simply accelerating the trend.

These observations have important implications for which of the scenarios presented in Section 5 one finds most compelling. If one believes the CPS datasets more accurately reflect the role of the employment margin in driving labor-quality growth, then the two plateau scenarios appear most compelling: they suggest that the U.S. experienced an unusual upskilling of employment during the Great Recession that will either persist or unwind, while offering little evidence of a pre-Great Recession upskilling trend in employment. If, on the other hand, one believes that the ACS data more accurately reflects the contribution of employment composition to labor quality then there appears to have been a significant pre-Great Recession structural trend, suggesting that the labor-quality growth from employment composition is unlikely to fully unwind and may even continue to drive a significant portion of labor-quality growth going forward.

Table 1: Different levels of granularity of classification of variables

	Variable	Number of classifications	Groups per classification			
			(I)	(II)	(III)	(IV)
1.	Gender	1	2	-	-	-
2.	Age	2	9	13	-	-
3.	Education	4	4	5	7	16
4.	Race/ethnicity	4	2	3	5	8
5.	Industry	2	12	50	-	-
6.	Occupation	3	10	22	51	-

Note: Total number of possible stratum specifications (including omission of one or more variables) is 1,799. Most granular definition includes 8,486,400 strata.

Table 2: labor-quality growth projections: 2015-2022

		EPOP ratio		
		Trend	Level	
		Extrapolated '02-'07 trends in EPOP (*I*)	Revert to pre-crisis EPOPs (*II*)	Persist at 2013 EPOPs (*III*)
Education	1. Extrapolate 2007-2013 trends in education	0.25/1.21	-0.20/0.86	0.15/0.51
	2. Revert to pre-crisis educational attainment	0.19/1.20	-0.27/0.88	0.09/0.50
	3. Persist at 2013 educational attainment	0.25/1.22	-0.21/0.88	0.14/0.51

Note: Reported are average annual log-growth rates for 2015-2022.
Row scenarios: (1) Uptick in educational attainment since Great Recession reflects a permanent acceleration in educational attainment of those 30 and younger. (2) Uptick since 2008 is fully cyclical and educational attainment reverts to its 2008 level. (3) Uptick in educational attainment since Great Recession reflects a step increase educational attainment of those 30 and younger.
Column scenarios: (*I*) Pre-Great-Recession ('02-'07) trends in EPOP continue. (*II*) EPOP ratios in 2025 have reverted to their 2007 levels. (*III*) EPOP ratios will remain at their 2013 levels.
More details about these scenarios can be found in section A.3.

Table 3: labor-quality growth projections: 2022-2025

		EPOP ratio		
		Trend	Level	
		Extrapolated '02-'07 trends in EPOP (*I*)	Revert to pre-crisis EPOPs (*II*)	Persist at 2013 EPOPs (*III*)
Education	1. Extrapolate 2007-2013 trends in education	0.47/0.60	0.28/0.37	0.27/0.41
	2. Revert to pre-crisis educational attainment	0.30/0.59	0.09/0.42	0.09/0.42
	3. Persist at 2013 educational attainment	0.32/0.60	0.12/0.41	0.12/0.42

Note: Reported are average annual log-growth rates for 2022-2025.
Row scenarios: (1) Uptick in educational attainment since Great Recession reflects a permanent acceleration in educational attainment of those 30 and younger. (2) Uptick since 2008 is fully cyclical and educational attainment reverts to its 2008 level. (3) Uptick in educational attainment since Great Recession reflects a step increase educational attainment of those 30 and younger.
Column scenarios: (*I*) Pre-Great-Recession ('02-'07) trends in EPOP continue. (*II*) EPOP ratios in 2025 have reverted to their 2007 levels. (*III*) EPOP ratios will remain at their 2013 levels.
More details about these scenarios can be found in section A.3.

Figure 1: Trade off between fraction of wage variation captured and precision of imputed wages.

(a) All 1799 specifications (b) All variables and specifications excluding age and education

(c) Age, education and industry and/or occupation (d) Age, education, and gender and/or race

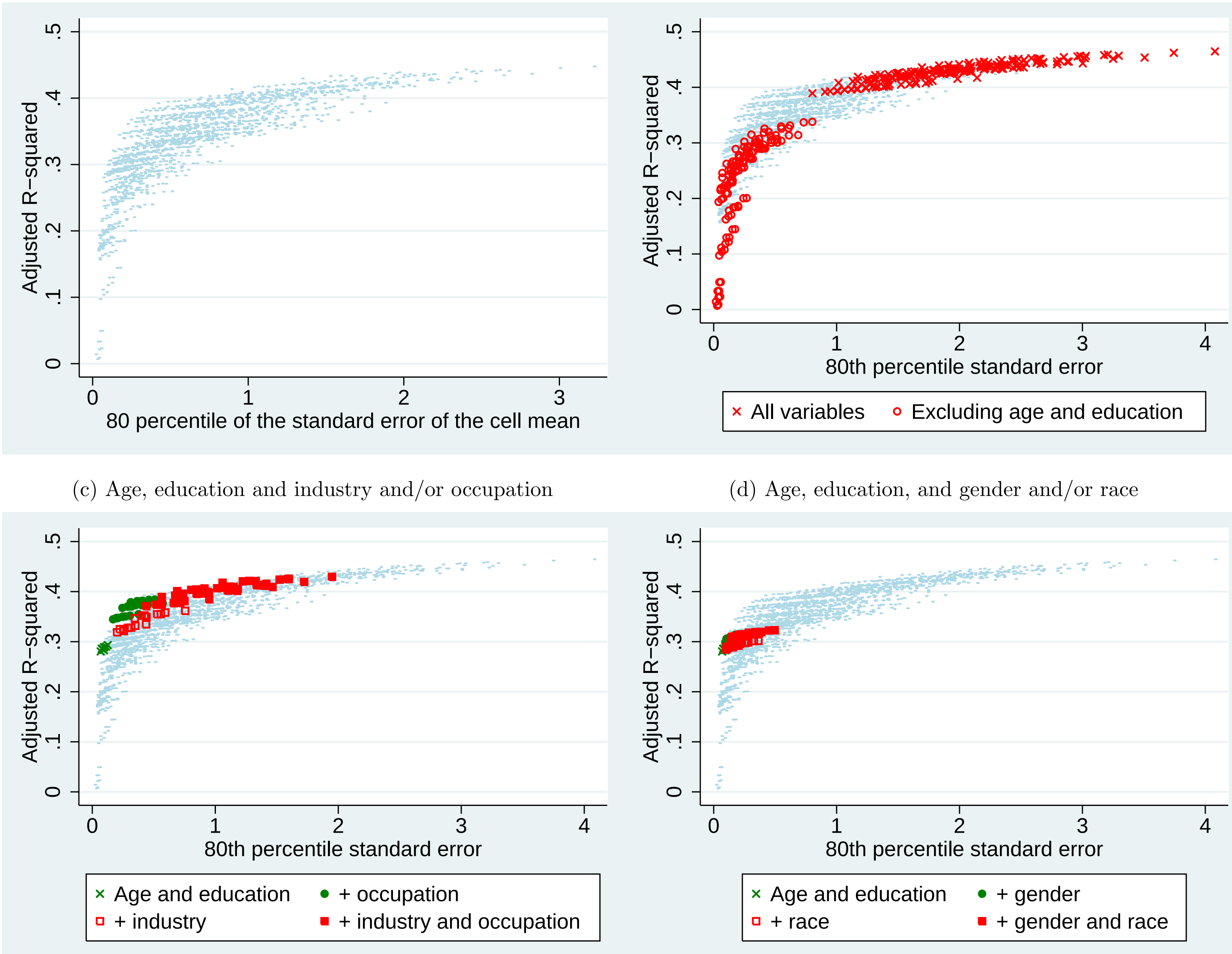


Figure 2: Regression-based fit and precision compared to stratum-based specifications.

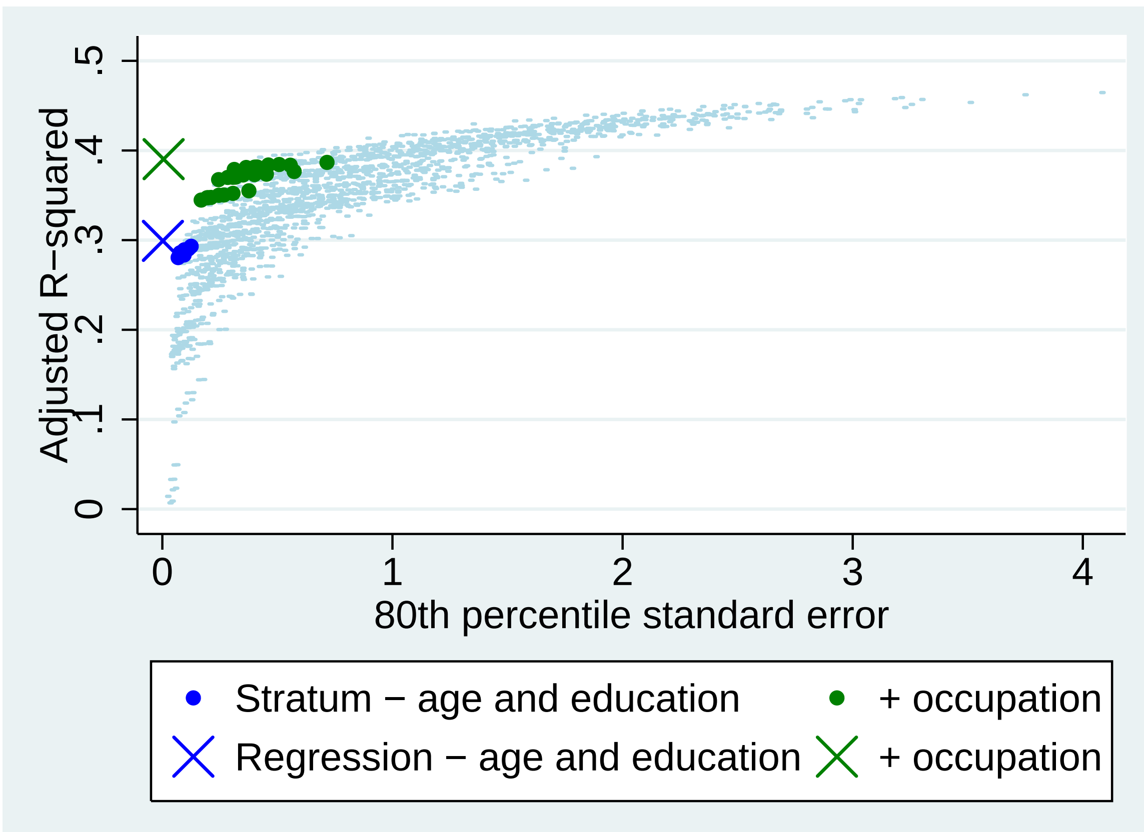
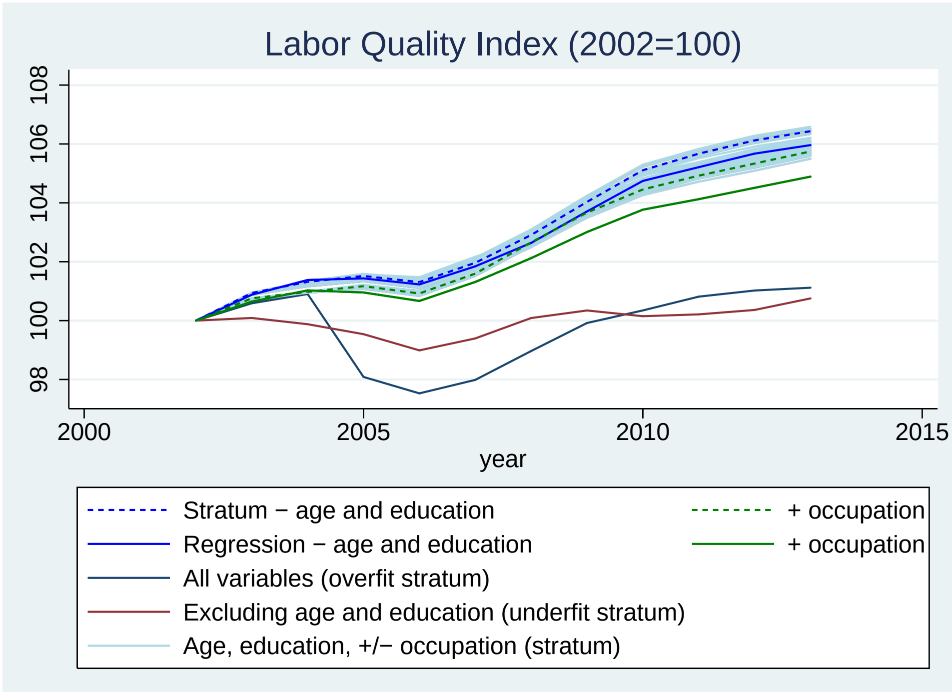


Figure 3: Comparison of results across specifications and datasets: 2002-2013.

(a) Different specifications using ACS data



(b) ACS, CPS-ASEC, and CPS-ORG

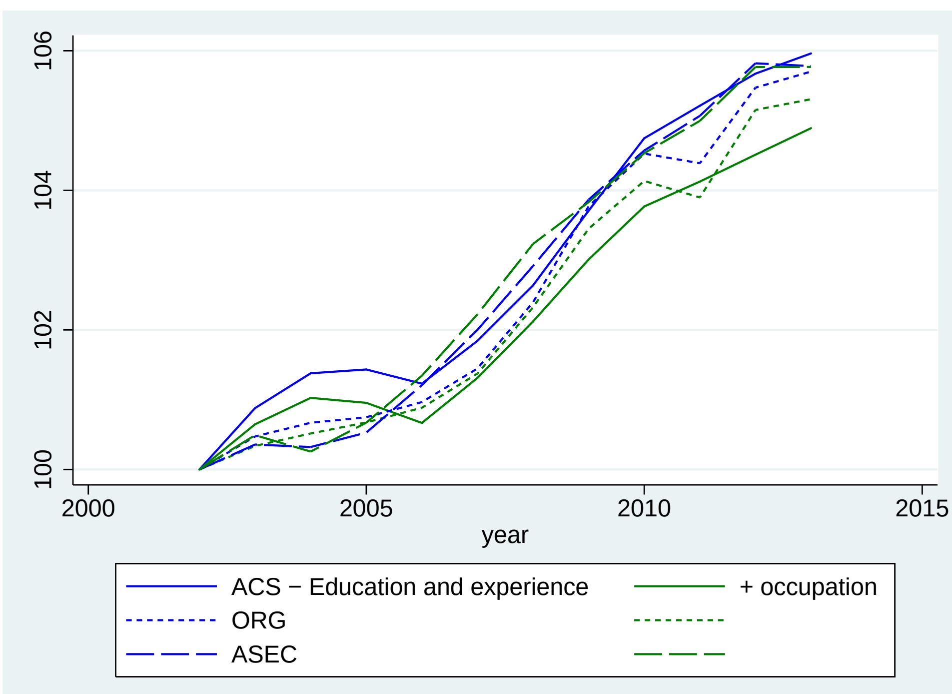


Figure 4: Counterfactual indices for 2002 baseyear hours, employment, and population.

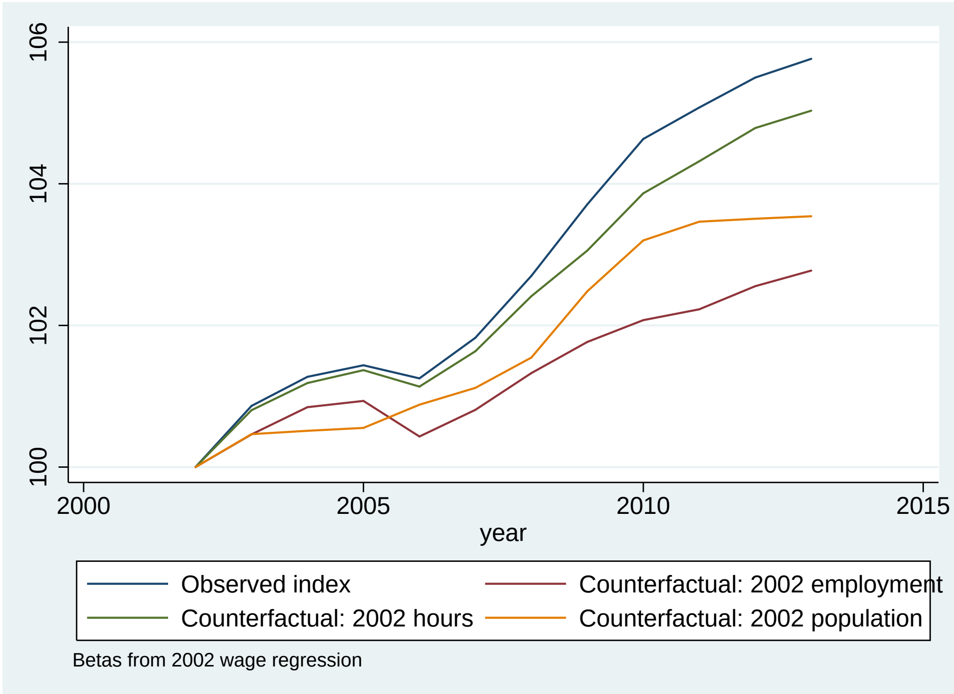
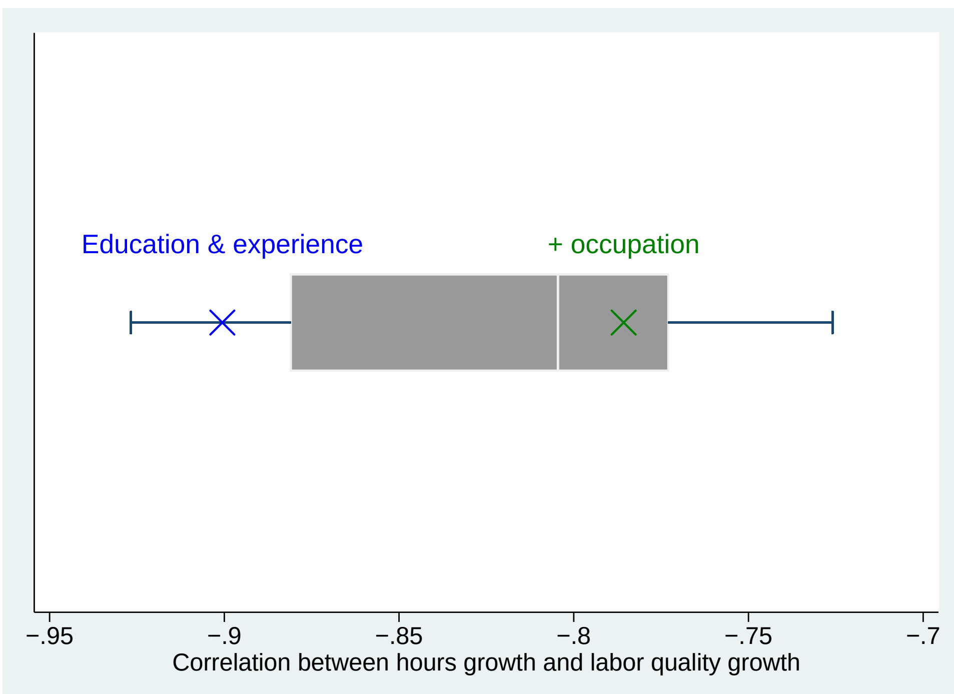
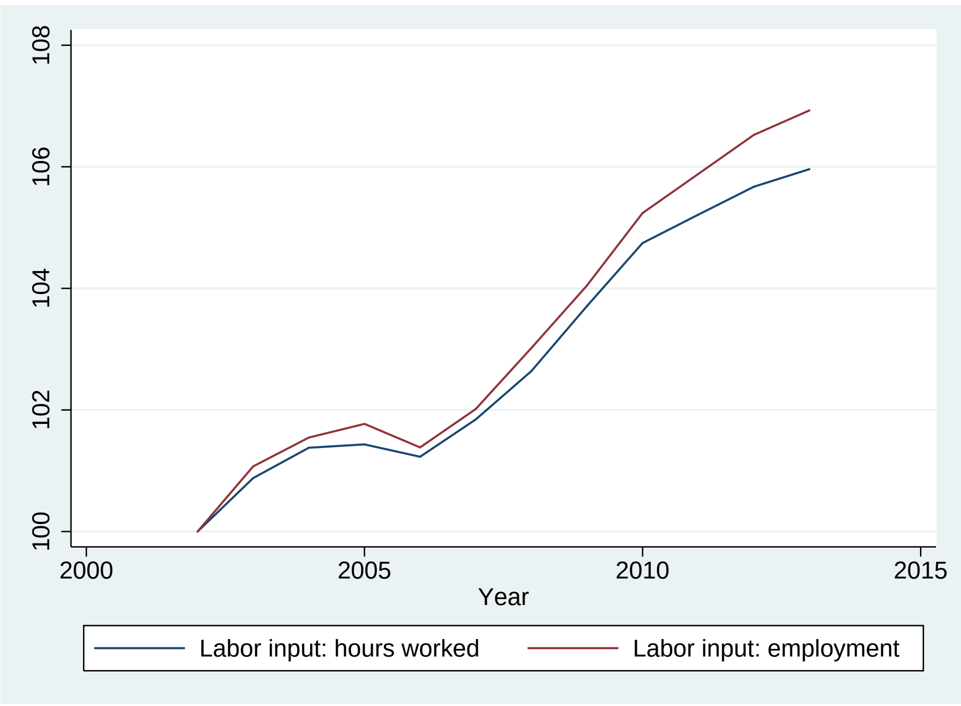


Figure 5: Correlation between labor-quality growth and hours growth for key indices.



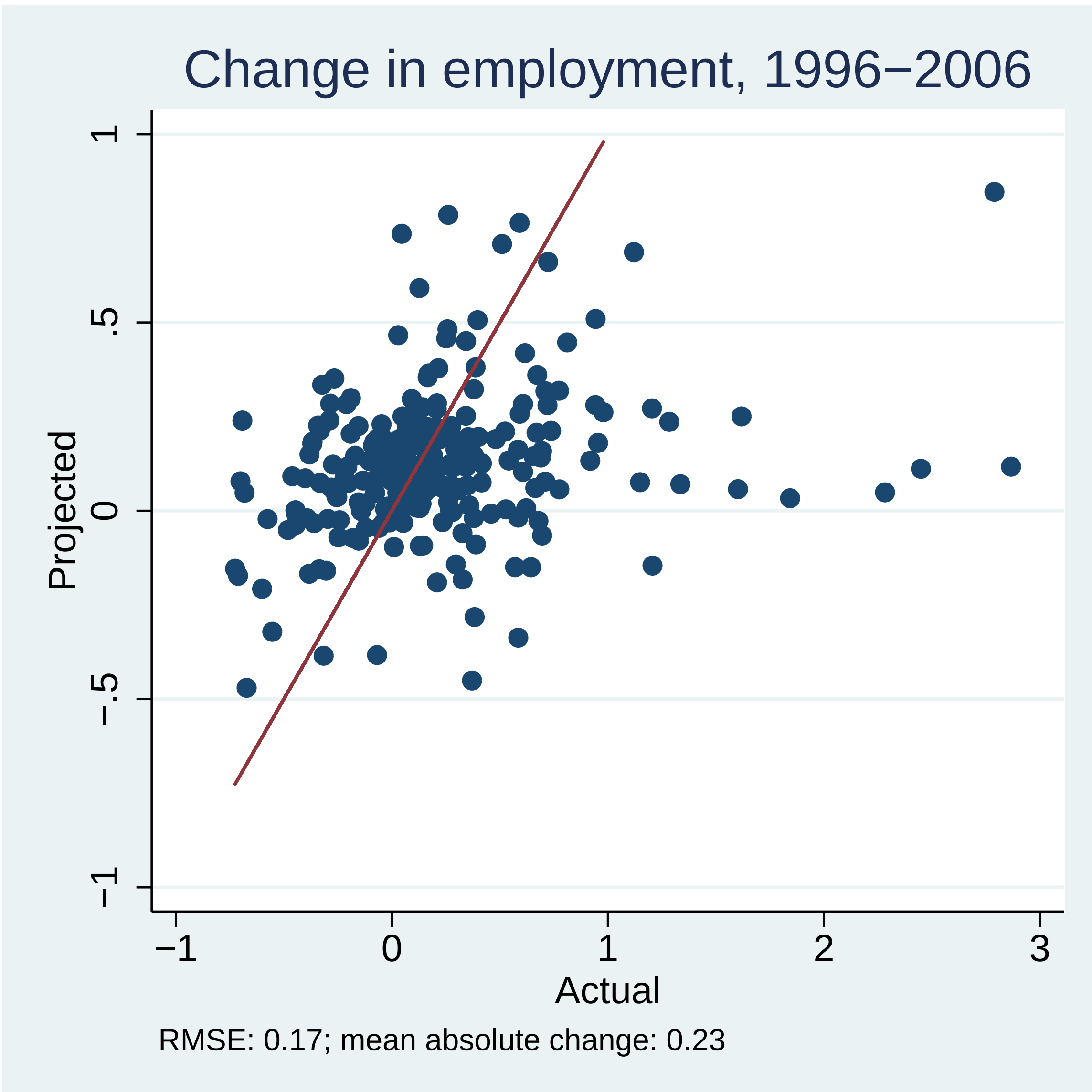
Note: The plotted correlations are from the age and education and the age, education, and occupation specifications by both the stratum- and regression-based methods. That is, they are all of the specifications plotted in Figure 3a, except for the overfit and underfit stratum specifications. The blue X identifies our *baseline* specification and the green X identifies our *baseline-plus-occupation* specification.

Figure 6: Hours- versus employment-based historical labor-quality indices.



Note: Employment-based indices ignore variation in average hours worked, $\eta_{i,t}$, across worker types.

Figure 7: Actual and projected employment growth by occupation (6-digit SOC).

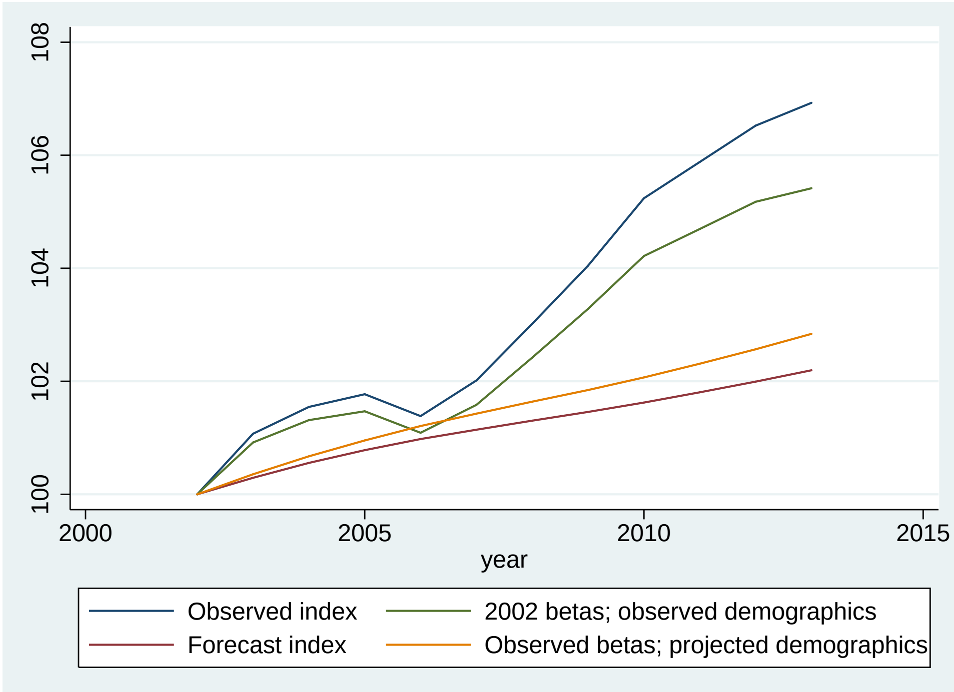


Note: plotted are actual and projected average annual percent changes in employment by 6-digit SOC occupation.

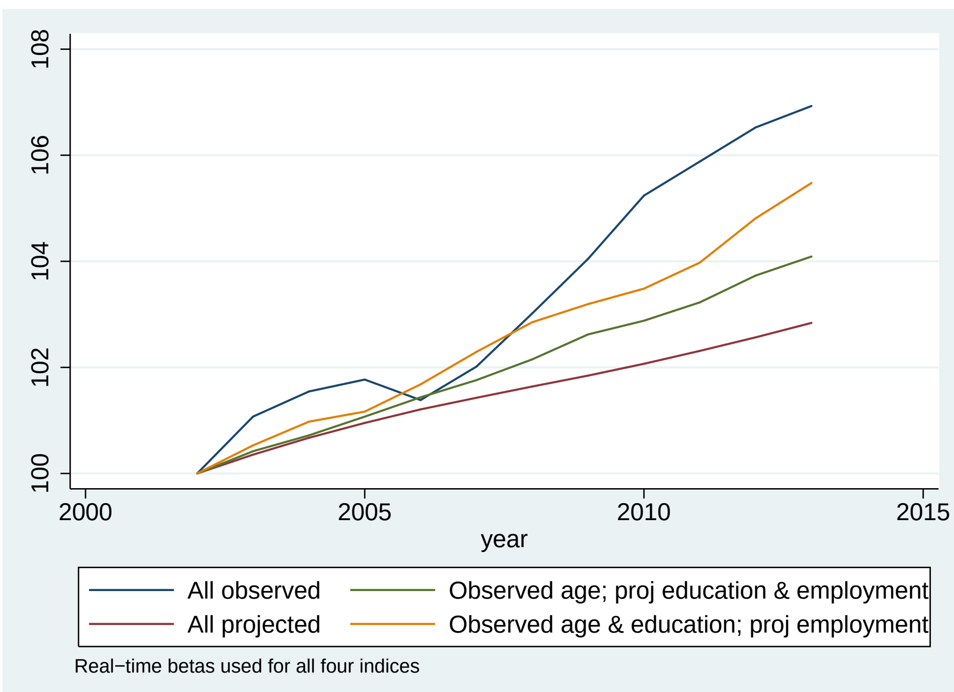
Source: Wyatt (2010).

Figure 8: Decomposition of forecast errors from 2002-2013.

(a) Projected hours distribution of $\mathbf{x}_i$ and projected $\boldsymbol{\beta}_t$



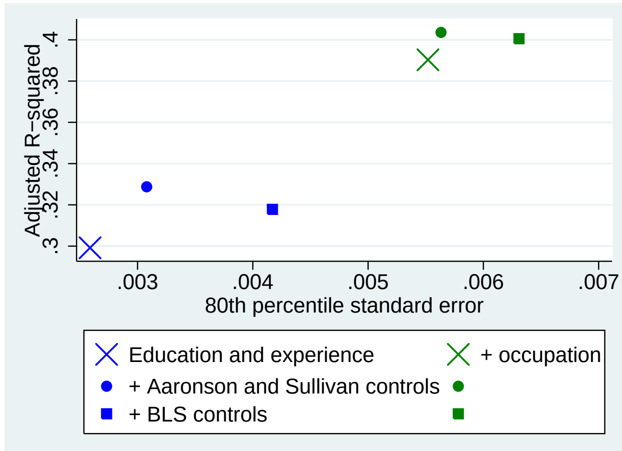
(b) Projected EPOP ratios and education rates for observed $\boldsymbol{\beta}$s



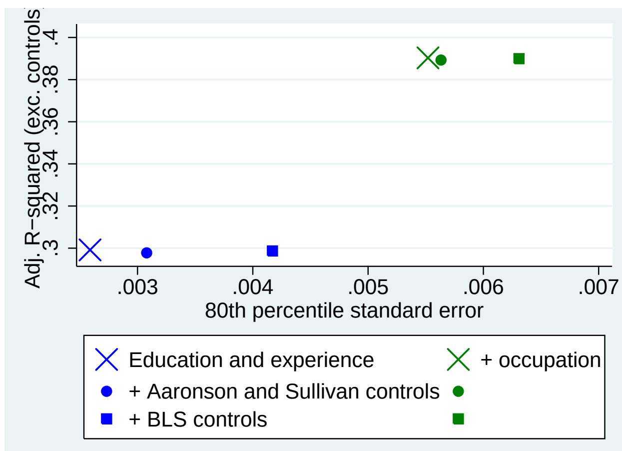
Note: EPOP and educational attainment models used to construct historical forecasts based on 1992-1997 CPS-ORG data. *Forecast index* in panel (a) is based on extrapolated cohort effects holding $\boldsymbol{\beta}$s constant at their 2002 values. *Observed betas* in panel (a) is the same as *All projected* in panel (b).

Figure B.1: Impact of including controls in Mincer specifications.

(a) Fit of both core and control variables



(b) Fit of the core variables only



(c) labor-quality indices, with and without controls

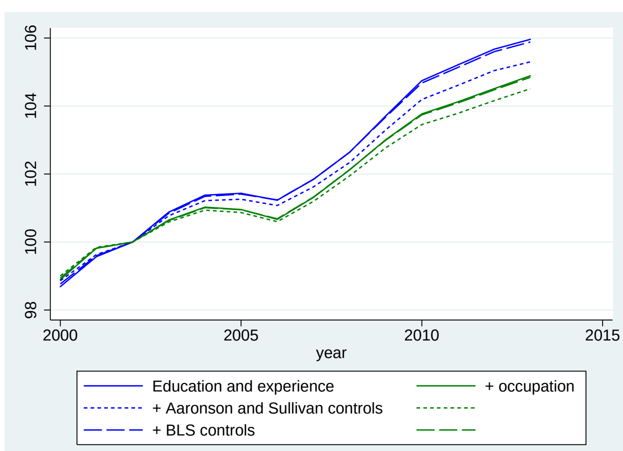
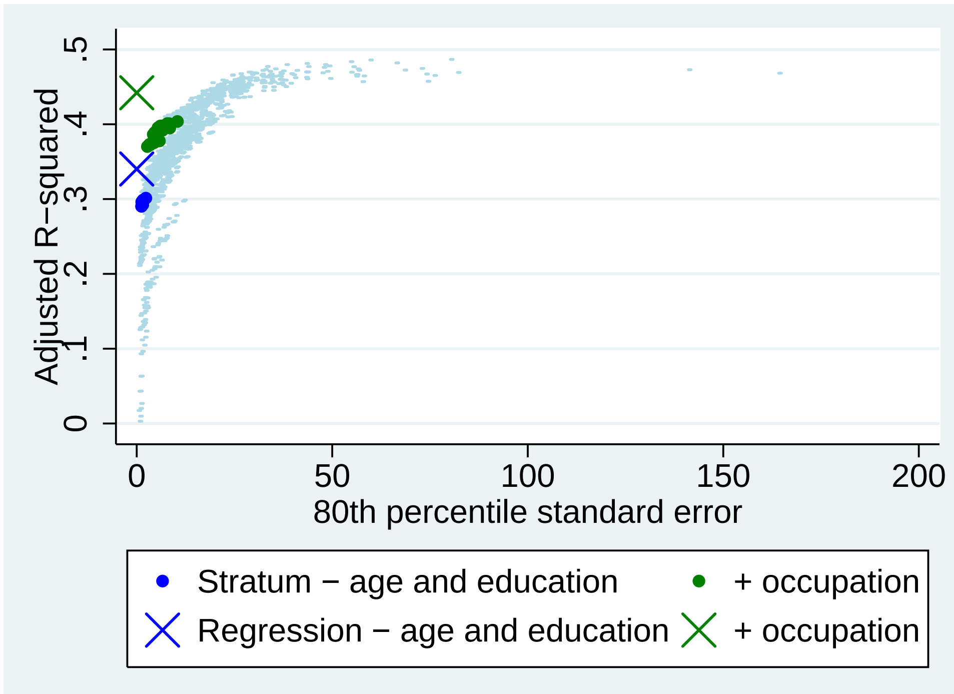


Figure B.2: Regression-based fit and precision compared to stratum-based specifications—CPS datasets.

(a) CPS-ORG



(b) CPS-ASEC

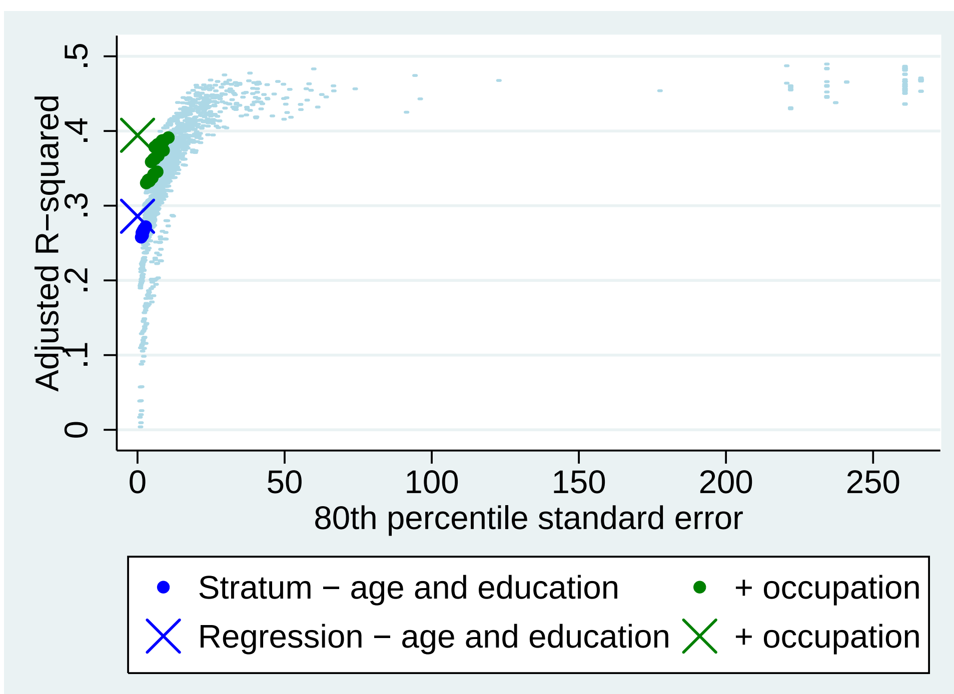
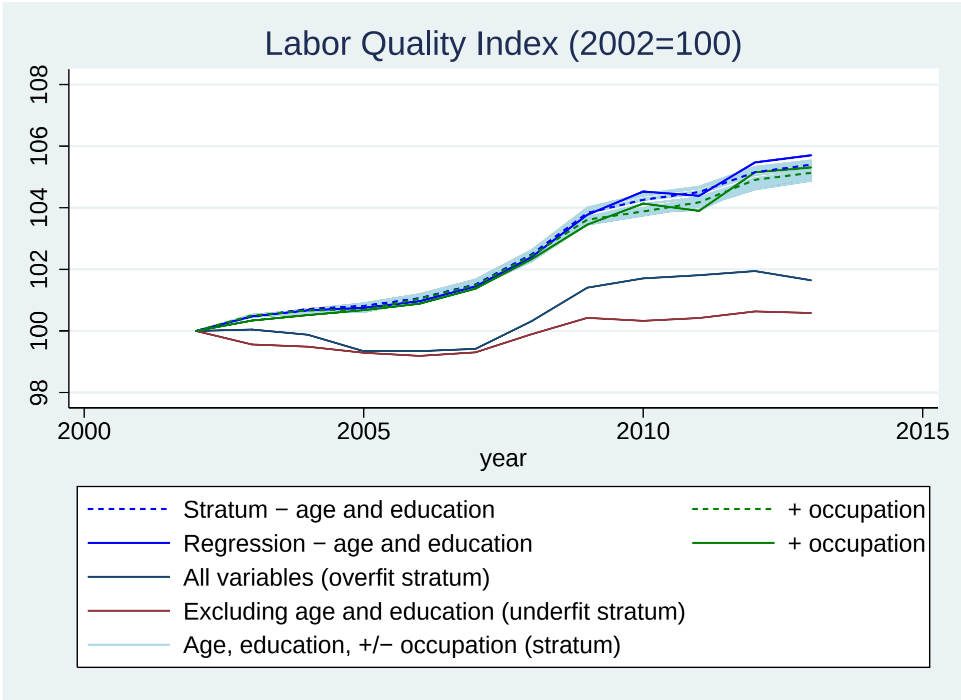


Figure B.3: Comparison of results across specifications: 2002-2013—CPS datasets.

(a) CPS-ORG



(b) CPS-ASEC

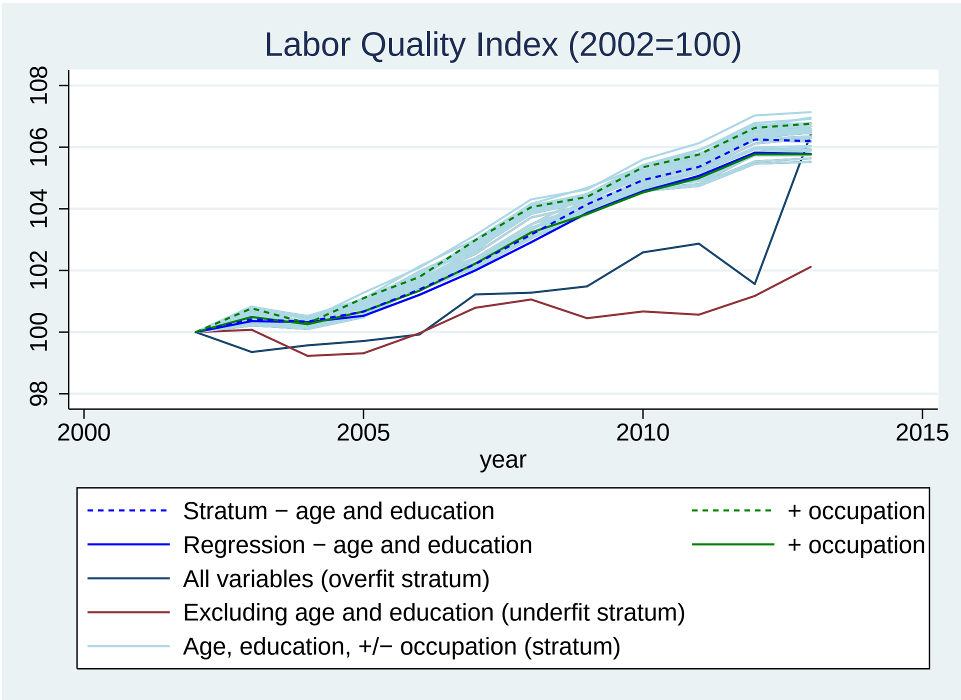
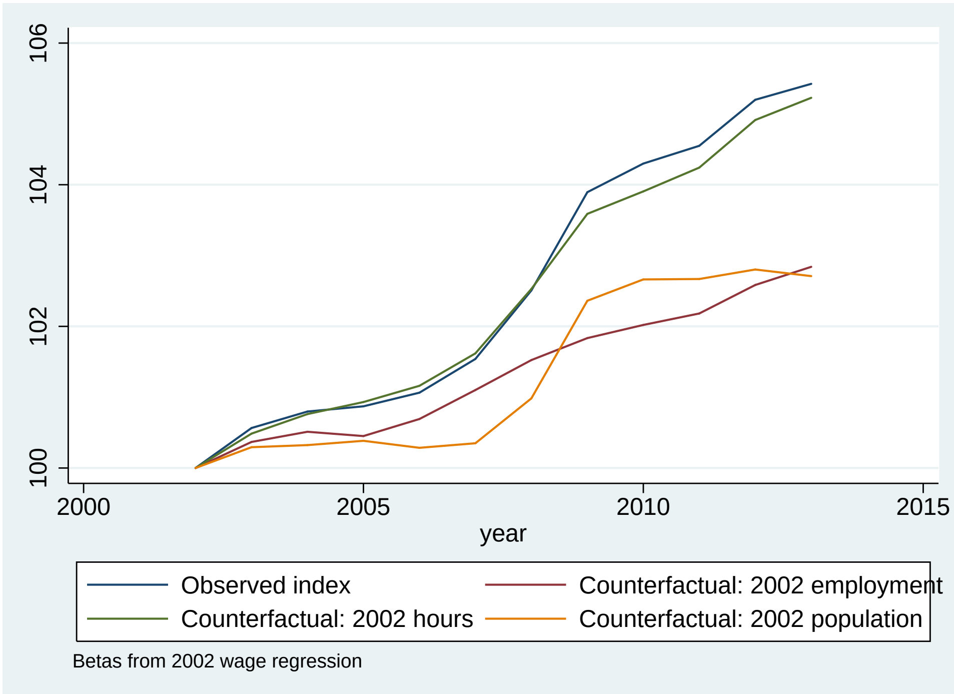


Figure B.4: Counterfactual indices for 2002 baseyear hours, employment, and population—CPS datasets.

(a) CPS-ORG



(b) CPS-ASEC