ISSN 2042-2695

CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 1215

May 2013

Has Job Polarization Squeezed the Middle Class? Evidence from the Allocation of Talents

Michael J. Boehm

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
Over the last two decades, earnings in the United States increased at the top and at the bottom of the wage distribution but not in the middle - the intensely debated middle class squeeze. At the same time there was a substantial decline of employment in middle-skill production and clerical occupations - so-called job polarization. I study whether job polarization has caused the middle class squeeze. So far little evidence exists about this because the endogenous selection of skills into occupations prevents credible identification of polarization's effect on wages. I solve the selection-bias problem by studying the changes in returns to occupation-specific skills instead of the changes in occupational wages using data over the two cohorts of the National Longitudinal Study of Youth (NLSY). This data features multidimensional and pre-determined test scores, which predict occupational sorting and thus measure relative occupation-specific skills. My estimation equations are derived from the Roy (1951) model over two cross-sections with job polarization amounting to a shift in the occupation-specific skill prices. In line with polarization, I find that a one percentage point higher propensity to enter high- (low-) as opposed to middle-skill occupations is associated with a .29 (.70) percent increase in expected wages over time. I then compute a counterfactual wage distribution using my estimates of the shifts in occupation-specific skill prices and show that it matches the increase at the top of the wage distribution but fails to explain the increase at the bottom. Thus, despite the strong association of job polarization with changes in the returns to occupation-specific skills, there remains room for alternative (e.g. policy related) explanations about the increase in the lower part of the wage distribution.

Keywords: Job polarization; wage inequality; talent allocation; Roy model
JEL Classifications: J21, J23, J24, J31

This paper was produced as part of the Centre's Labour Markets Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
I am extremely grateful to my supervisor Jörn-Steffen Pischke. I very much thank Yona Rubinstein, Alan Manning, Luis Garicano, David Autor, David Dorn, Pedro Carneiro, John Van Reenen, Guy Michaels, Esteban Aucejo, Barbara Petrongolo, Claudia Steinwender, Johannes Boehm, Georg Graetz, Martin Watzinger, and Pinar Hosafci for comments. I also thank Joseph Altonji, Prashant Bharadwaj, and Fabian Lange for providing me with their data and code, and Steve McClaskie for support with the NLSY data.

Michael Boehm is an Occasional Research Assistant at the Centre for Economic Performance and a PhD candidate, Department of Economics, London School of Economics and Political Science.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Over the last two decades, wages of middle class workers in the United States have been squeezed, in that earnings in the middle of the wage distribution have stagnated or even fallen while earnings at the top and at the bottom have increased. This has coincided with a decrease of employment in middle-skill production and clerical occupations, and an increase of employment in low-skill services and high-skill professional and managerial occupations—so-called job polarization. Many economists believe that job polarization and the middle class squeeze are two sides of the same coin. In particular, they think that a negative demand shock for the middle-skill occupations has simultaneously reduced middle-skill employment and middle-class wages. If this is true, the middle class squeeze is a consequence of market forces, and it will be difficult to design policies that reverse the trend and help the middle class without hampering the efficiency of the economy.[1]

However, there is little evidence so far which establishes a direct link between job polarization and the middle class squeeze. On the one hand, a large body of research in labor economics and international trade has found a drop in the demand for jobs that can be replaced by computers or off-shored and shown that many of these jobs are in middle-skill occupations. On the other hand, there are plenty of hypotheses about other factors which could have contributed to the U-shaped change in wage inequality that characterizes the middle class squeeze—including increases in the minimum wage, de-unionization, and the deregulation of financial and related professions. If such policy-related or institutional factors have caused the downward pressure on the middle of the wage distribution, policy makers may be called to action in order to support the middle class.

The goal of my paper is thus to answer the question: does job polarization explain the middle class squeeze? I do this by studying how the wages of workers who would have chosen the high-, middle-, or low-skill occupations in the 1980s have changed over time. To be exact, since the same workers cannot be observed both before and

[1]The struggles of the middle class are a major issue in the public and political debate. For example, this editorial in the International Herald Tribune from August 30, 2012 takes the market-based view: "The economic reality is that, thanks to smart machines and global trade, the well-paying, middle-class jobs that were the backbone of Western democracies are vanishing. Neither Mitt Romney's smaller state nor Barack Obama's larger one will bring them back."(Freeland 2012)

after polarization has taken place, I study the returns to talents that are associated with choosing the particular occupations over time. The Roy (1951) model of self-selection into sectors guides my empirics: in the model, workers' skills in occupations are made up of observable as well as unobservable components and the returns to talents that I estimate only reflect the observable part. However, using the sorting of- and the returns to observables, I can estimate the shifts in occupation-specific prices per unit of skill, which also apply to the unobservables, and examine how much of the middle class squeeze they explain. In addition, I assess the role that heterogeneous gains from switching occupations may play for the change of the wage distribution.

So far, the fundamental problem in linking job polarization to the wage distribution has been that one could not estimate the effect of occupational demand on workers' wages. Job choices are naturally dependent on the price movements so that the skill selection into occupations changes endogenously. Hence, a comparison of wages in high-, middle-, and low-skill occupations over time would confound the relative demand shifts with a changing composition of workers' skills in each occupation.[2] The problem is exemplified by the fact that average wages in the middle-skill occupations have not declined compared to average wages in the low-skill occupations in several datasets and samples (for example, Goos and Manning 2007, and the data used here).

The point of departure for my analysis is the regression equation formulation of the Roy model as in Heckman and Sedlacek (1985):[3] every worker possesses a vector of talents which combine into skills in each occupation and which are only partly observed in the data. The log wage offered to workers in a given occupation is then the sum of an occupation-specific log skill price, which is the regression intercept, an observable component of skill, which is the regressor, and an unobservable component of skill, which is the orthogonal regression error. In this framework, the relative demand shocks of polarization amount to a shift in the occupation-specific skill prices.

[2] To quote the well-known survey paper by Acemoglu and Autor (2010, p78): "[...] because the allocation of workers to tasks is endogenous, the wages paid to a set of workers previously performing a given task can fall even as the wages paid to the workers now performing that task rise. [...] a regression of wages on tasks currently performed, or their change over time, would be difficult to interpret."

[3] The only difference is in labels: I call talents what Heckman and Sedlacek (1985) call skills and I call skills what they call tasks.

This has the effect that the relative returns to workers' talents change, but also that workers switch occupations depending on their observed and unobserved skills. The switching due to the unobserved skills causes selection bias in occupational wages, since we do not know whether rising wages in an occupation are due to a rise in the price per unit of skill in this occupation or due to a better selection of workers with respect to the unobserved component of skill.

My paper solves the selection bias problem, and it circumvents the structural estimation of the Roy model, by shifting the analysis from occupational wages to the returns to occupation-specific skills. I estimate the changes in returns to the observable component of occupation-specific skills with a two-stage procedure. First, using workers' talents I predict their propensities to enter the high-, middle-, and low-skill occupations in the period before polarization took place. Second, I estimate the changes in the returns to these propensities.[4]

In order to implement this procedure, I need two cross-sections of data with consistent measures of workers' talents that predict occupational sorting but are not influenced by occupational choice and thus not endogenous to polarization. Such data has only recently become available in the form of the National Longitudinal Survey of Youth (NLSY):[5] the NLSY cohorts of 1979 and 1997 contain detailed and multidimensional measures of talents which are hardly malleable and determined well before a worker's entry into the labor market. The measures include mathematical, verbal, and mechanical test scores as well as risky behaviors and parental education. In addition, the data are available for two representative cross-sections of 27 year olds in the end of the 1980s and the end of the 2000s, and therefore well-timed for studying polarization and the middle class squeeze.[6]

The estimation results on the returns to observable occupation-specific skills indicate a strong impact of polarization on wages. I find that a one percentage point

[4]Acemoglu and Autor recommend a similar procedure but lack the data to implement it. In their words: "[...] the approach here exploits the fact that task specialization in the cross section is informative about the comparative advantage of various skill groups, and it marries this source of information to a well-specified hypothesis about how the wages of skill groups that differ in their comparative advantage should respond [...]" (Acemoglu and Autor 2010, p78)

[5]Until recently the 1997 cohort of the NLSY was too young to warrant a reliable analysis of labor market outcomes.

[6]Moreover, the data from the NLSY79 and the NLSY97 were designed to be comparable to one another.

higher propensity to enter the high- as opposed to the middle-skill occupation is associated with a .29 percent increase in wages over time. A one percentage point higher propensity to enter the low as opposed to the middle-skill occupation is associated with a .70 percent increase in wages. Workers with a high propensity to enter the middle-skill occupations in the 1980s actually suffer an absolute decline in their expected real wages. This finding is robust to controlling for absolute skill measures such as educational attainment, which supports the idea that it is relative occupational skills rather than absolute skills whose returns have changed over time.

The effect identified in these estimations is a combination of the direct demand effect of polarization on talent returns as well as the potentially heterogeneous wage gains for workers of different talents from reallocating out of the middle-skill occupations. Moreover, at age 27, the workers in the NLSY97 are young enough to have chosen their occupations when most of polarization has already taken place. Thus, the effect on their wages is likely to be largely due to ex ante different talent endowments and not due to having acquired occupation-specific experience whose value has changed ex post. This indicates a long-lasting effect of polarization on relative wages that will not fade when the current generation of workers retires.[7]

The changing returns to propensities of entering high-, middle-, and low-skill occupations may in fact be driven by alternative factors which are correlated with occupations. I address this concern by exploiting the Roy model's prediction about specific talent returns under polarization: if only occupation-specific skill prices are shifting, the Roy model implies that the change in the return to each talent solely depends on how that talent is associated with occupational choice and how the association changes over time.

I use this prediction to estimate the change in relative occupation-specific skill prices and to test the null hypothesis that all changes in returns to talents—and equivalently all changes in returns to occupational propensities—were driven by polarization. In the data, I observe each talent's initial and final association with the three classes of occupations but not the adjustment path over time. I therefore linearly interpolate the adjustment path, which gives relative price estimates that are

[7]It thus implies the need for long-term policy responses, e.g. long-term changes in education or tax policy instead of income support for the current generation of workers.

close to the actual prices and at the same time robust to different distributions of unobserved skills. Since the NLSY provides more talents—three test scores plus the risky behaviors and other demographics—than the two unknown relative prices, I obtain over-identifying restrictions on talent returns from the model which I use to test the polarization hypothesis and to estimate the relative occupation-specific skill price changes.

The over-identifying restrictions test does not reject the polarization hypothesis in the data. Moreover, the relative skill price increase in the high compared to the middle-skill occupation is precisely estimated at 20 percent, while the relative skill price increase in the low-skill occupation is imprecisely estimated with a point estimate of 31 percent. The relative skill price estimates are crucial to assess the overall impact of polarization on the wage distribution. This is because the returns to observable talents or occupational propensities alone can only explain a small part of the change in the wage distribution—just as they can only explain a small part of the variation of wages in the cross-section. In contrast, the relative occupation-specific skill price estimates change the return to the observable as well as the unobservable components of skill in each occupation and thus allow me to assess the full effect of polarization on the wage distribution.

Therefore, I compute a counterfactual wage distribution which is due to the relative occupation-specific skill price effect of polarization and compare it to the actual distribution. I do this by assigning the estimated relative skill price changes to each worker in the NLSY79 according to his occupation. It turns out that the counterfactual distribution closely matches the increase of wages at the top of the actual distribution compared to the middle. However, it fails to match the increase of wages at the bottom of the actual distribution compared to the middle. The reason is that the wage rate estimates and the dispersion of wages within occupations is so high that also many middle-earners' wages are lifted by the price changes and that some low-skill occupation workers become middle-earners themselves.

Finally, if polarization is to be the main driver of the middle class squeeze, the remaining difference at the bottom between the counterfactual and the actual change in the wage distribution must be due to the heterogeneous effect of optimal occupational switching in response to polarization on different parts of the wage distribution—a reallocation effect. Since there are no clear predictions from the Roy model about this effect, I conduct rule-of-thumb experiments to assess whether the reallocation effect may in principle explain the remainder: I assign the lowest earning workers in the middle-skill occupation in the initial period gains that they could obtain from switching to the low-skill occupation due to polarization and examine the effect that this has on the change in the lower part of the wage distribution. Experiments with a substantial gain from switching can relatively well match the wage distribution in the bottom as well as average wages in occupations. However, the assumptions that I need to make for this are strong and they are not supported by the reallocation (effect) of observable skills, which I can measure in the data.

It thus seems that, despite its strong effect on relative wages, polarization can account well for only part of the evolution of the wage distribution over the past two decades. The results therefore suggest that market-based forces may not be responsible for all of the changes in the lower half of the wage distribution. This opens the door for policy-related and institutional factors—such as de-unionization and the minimum wage—that other studies have found to have an impact on earnings at the bottom of the wage distribution over this period (Machin and Van Reenen 2008, Autor, Manning, and Smith 2010, Firpo, Fortin, and Lemieux 2011).

The findings in this study are qualitatively similar when implementing alternative definitions of occupations or tasks in occupations that have been used in the literature on polarization. These include grouping occupations according to initial median wages or average education, splitting up the large middle-skill group into blue collar and white collar occupations, and employing continuous measures of routine and nonroutine (analytical and manual) task content in occupations.

The paper continues as follows. The remainder of this section discusses the relation to the existing literature. Section 2 demonstrates that job polarization and wage inequality in the NLSY are similar to what is found in the commonly used Current Population Survey (CPS), and it shows that workers sort themselves systematically into occupations according to the talent measures available in the NLSY. The Roy model and its empirical predictions are analyzed in section 3. Section 4 presents the empirical results on the returns to occupation-specific skills. Section 5 estimates the occupation-specific skill prices and tests the model, while section 6 assesses whether the resulting counterfactual wage distribution may match the actual. Section 7 concludes.

1.1 Related Literature

There are other studies that have tried to link job polarization to changes in the wage distribution. The most explicit effort is a recent paper by Firpo, Fortin, and Lemieux (2011) who use a Roy-style model to study the effect of shifts in the demand for tasks on occupational wages. They also carry out a decomposition to assess the effect of different factors such as occupational demands, skill supply, unionization, and minimum wages on the change in the wage distribution. Neither of these exercises control for the endogenous selection of workers with respect to unobservable skills. This limitation of Firpo, Fortin, and Lemieux (2011)'s analysis is noted by Acemoglu and Autor (2010), whose comparative advantage model predicts a changing self-selection of workers into occupations due to movements in wages rates across occupations or tasks. Exercises similar to Firpo, Fortin, and Lemieux (2011)'s that feature as part of broader papers may thus be regarded as mostly descriptive (e.g. Goos and Manning 2007, Autor, Katz, and Kearney 2008).

An alternative method to deal with endogenous selection is to employ panel data and worker fixed effects. Cortes (2012) uses data from the Panel Study of Income Dynamics to analyze the transition from middle- to high- and low-skill occupations due to polarization and its associated wage changes. Liu and Trefler (2011) similarly estimate the impact of trade in services with China and India on US workers using matched data from the Current Population Survey (see also Ebenstein, Harrison, McMillan, and Phillips 2011). Cortes finds a substantial impact of polarization on workers' wages while Liu and Trefler (2011) find a rather small impact of trade. A general difficulty with the panel data approach is the need to make an appropriate assumption about—or to control for—workers' counterfactual experience profiles of wages and occupations in the absence of polarization. Moreover, contrary to this paper and the one by Firpo, Fortin, and Lemieux (2011), these studies do not link their estimated earnings impacts of polarization to the change in the aggregate wage distribution.

The large literature on the causes of job polarization provides the hypothesis on occupational demands analyzed in my paper. During the last decade, many studies in labor economics and international trade have examined rapidly changing information and communication technology (ICT) and the off-shoring of goods and services production as causes of polarization. For example, papers that consider technological change include Autor, Levy, and Murnane (2003), Goos and Manning (2007), Michaels, Natraj, and Van Reenen (2010), Acemoglu and Autor (2010), and Autor and Dorn (2012). Papers that consider trade and offshoring include Blinder (2009), Becker, Ekholm, and Muendler (2009), Crinò (2010), Ottaviano, Peri, and Wright (2010), and Autor, Dorn, and Hanson (2012). Many of these studies find that it is largely occupations in the middle of the skill distribution that are affected by technology or trade.[8]

My approach to linking job polarization with changes in the wage distribution relies critically on Roy (1951)'s model of occupational choice and the development of his ideas by Gronau (1974) and Heckman (e.g. Heckman 1974, Heckman and Sedlacek 1985). In particular, the mathematical specification of how occupational skills are composed of observable and unobservable worker characteristics is identical to that of Heckman and Sedlacek (1985). Gould (2002) and Mulligan and Rubinstein (2008) are the first papers to explicitly link the Roy model to increases in wage inequality and skill-biased technological change (see also Yamaguchi 2012). Compared to these papers I study the Roy model in relation to job polarization and the U-shaped change of wage inequality.

Finally, there exists a large and diverse body of literature that analyzes hypotheses about drivers of wage inequality other than polarization. The most important of those is skill-biased technological change (SBTC), which is detached from demand for specific occupations (e.g. Bound and Johnson 1992, Autor, Katz, and Krueger 1998). Hypotheses complementing that of SBTC in the top of the wage distribution have emphasized firm size and organization as well as pay increases in financial services and other professions (e.g. Garicano and Rossi-Hansberg 2004, Gabaix and Landier 2008, Tervio 2008, Philippon and Reshef 2009).[9] In terms of the developments specific to

[8]Jaimovich and Siu (2012) find that job polarization and jobless recoveries after recessions are related to one another.

[9]For example, Garicano and Rossi-Hansberg (2004) show that improvements in communication technology lead to lower inequality at the bottom and higher inequality at the top of the wage distri-

the lower part of the wage distribution, changes in policy variables and labor market institutions such as minimum wages and unionization have been prominent in the discussion (see Machin and Van Reenen 2008, Autor, Manning, and Smith 2010, Firpo, Fortin, and Lemieux 2011).

Both a falling and a rising supply of skills have been analyzed as different sets of explanations for the change in inequality. Card and Lemieux (2001) and Goldin and Katz (2008) consider a slowdown in the rate of supply of college graduates, while Lemieux (2006, p461) argues that a large part of the changes in the wage distribution that we observe is due to "composition effects linked to the secular increase in experience and education". My study is most closely related to the papers that analyze the supply of, and returns to, ability test scores (e.g. Murnane, Willett, and Levy 1995, Blau and Kahn 2005, Altonji, Bharadwaj, and Lange 2008). After Altonji, Bharadwaj, and Lange (2012, ABL), this is also the first study to analyze labor market outcomes across the two cohorts of the NLSY. While ABL examine the effect of changes in overall skill supply on wage levels and inequality in the economy, my paper analyzes the effect of shifts in skill demand across occupations.

2 Data and Empirical Facts

This section establishes the stylized facts of job polarization and the u-shape change in wage inequality in my data. Median real wages for 27 year old males rise only very little, so the other characteristic of the middle class squeeze—stagnating incomes—is also present in my data. Moreover, the section shows how workers systematically sort themselves into the occupations affected by polarization depending on their talent endowments.

2.1 Job Polarization and the U-Curve of Wages

I use data from the National Longitudinal Survey of Youth (NLSY) cohorts of 1979 and 1997, which contain detailed information on individuals' fundamental talents that is not available in other datasets. Moreover, the two cohorts are specifically designed to to be comparable to one another. When possible, I compare my results

bution, and thus a squeezed middle, in a hierarchy model of endogenous firm size and organization.

to the more standard Current Population Survey Merged Outgoing Rotation Groups (CPS) over the same period.

The individuals in the NLSY surveys are born between 1956 and 1964 and between 1980 and 1984, respectively. I restrict my attention to 27 year olds, which is the oldest age that I have enough data in the NLSY97 for to analyze, and to males.[10] The sample selection and attrition weighting is done closely in line with a recent paper using both of the NLSY cohorts by Altonji, Bharadwaj, and Lange (2008). Labor supply by hours worked and hourly wages are defined as in Lemieux (2006). The details of the sample construction can be found in Appendix A.1. Table 1 accounts for how I end up with a sample of 3,054 and 1,207 individuals in the NLSY79 and the NLSY97, respectively.

For the overall (male) labor force, the wage distribution change from the end of the 1980s to the end of the 2000s is characterized by a U-shape, i.e. wages increase substantially at the top of the distribution and somewhat less at the bottom but hardly at all in the middle (the middle class squeeze). Moreover, there is job polarization in the sense that employment in the middle-skill occupations decreases and employment in the high-skill and low-skill occupations increases. For the details of these facts, see the survey paper by Acemoglu and Autor (2010).

I start with the stylized fact about the wage distribution in my data. Figure 1 graphs the empirical cumulative log wage distribution in the NLSY79 and the NLSY97 in the top two sub-figures and the change in wages by distribution quantile compared to the CPS in the bottom sub-figure. We see that the wage distribution levels and, more importantly, the changes in the NLSY and the CPS align well for both cohorts. This establishes the well-known U-shape in the wage distribution for the NLSY.[11]

The second important fact is job polarization. The literature has measured high-, middle-, and low-skill occupations in different ways and arrived at the same results. It has ranked them by initial median wages or average education (e.g. Autor, Katz, and

[10] At the time of writing, NLSY97 data was available up to 2009. The periods that I compare are thus 1983-1991 and 2007-2009.

[11] The increase at the top for 27 year olds is not as pronounced as previous papers have found for prime age males (e.g. Acemoglu and Autor 2010). This is not surprising, since the wage trajectory for high-skilled workers is steep around the age of 27 and thus the differences, and their changes, are likely to be larger at older ages.

Kearney 2006, Goos and Manning 2007). Alternatively, it has grouped managerial, professional, and technical occupations as high-skill; sales, office and administrative, production, and operator and laborer occupations as middle-skill; and protective, food, cleaning and personal service occupations as low-skill (e.g. Acemoglu and Autor 2010, Cortes 2012, Jaimovich and Siu 2012).

I use the latter approach of grouping occupations in figure 2 and in the paper more generally for two reasons: it is becoming a standard in the literature and it explicitly delineates occupations by the extent of abstract (high-skill), routine (middle-skill), and manual (low-skill) tasks that they require (see Acemoglu and Autor 2010, Jaimovich and Siu 2012). The upper two sub-figures graph the employment shares in the three occupation groups for the NLSY79 and NLSY97 and compared to the CPS. The share of employment in the middle-skill occupations is declining while the share of employment in the high- and the low-skill occupations is rising. This can be seen more clearly in the lower sub-figure, which plots the changes in employment shares. These facts establish job polarization in the NLSY, which is very similar to what can be found for 27 year olds in the CPS. The findings are the same if I use the alternative approaches of grouping occupations as low, middle, and high-skilled.

Before moving on, figure 3 shows average 1979 real wages in high-, middle-, and low-skill occupations and how they have changed over the two cohorts in the NLSY and, for comparison again, the CPS. Unsurprisingly, average wages in high-skill occupations are higher than in middle-skill occupations, which in turn are higher than average wages in low-skill occupations. The changes are more interesting. While wages in high-skill occupations have increased robustly in levels and compared to the other two occupations, wages in low-skill occupations have lost somewhat further ground against wages in middle-skill occupations in the NLSY and also slightly in the CPS.[12] One might find this surprising under the demand side explanation for job

[12]Note that the small differences between wages, occupational employment, and occupational wages in the NLSY and the CPS are unlikely to stem from systematic sample attrition or non-test-taking in the NLSY. This is because sample attrition or non-test-taking are much lower in the NLSY79 than the NLSY97, while the differences between CPS and NLSY are equally large for the two cohorts. Further, note again that the scope of the NLSY and the CPS are different. The CPS is supposed to be representative of the resident population in the survey year while the NLSY is supposed to be representative of those individuals in the survey year who were between 14 and 21 years old in 1979 and between 12 and 16 in 1997, respectively.

polarization, which should decrease employment and wages in the middle at the same time. Yet, just as the size of occupations, the composition of skills in occupations does not stay constant when relative demands change.[13] Appropriately adjusting for this effect is the main contribution of my paper.

2.2 Talent Sorting into Occupations

Workers do not choose to work in the high-, middle-, and low-skill occupations at random. This section uses choice regressions to establish and quantify systematic occupational sorting in the data.

2.2.1 Measures of Talent

The NLSY data provides a long array of characteristics of its respondents. Out of these, I focus on variables that are early determined, that are relevant for occupational choice and wages, that may approximate different dimensions of skill, and that can be compared over the two cohorts.[14]

Table 2 reports labor force averages of NLSY variables that fulfill the four criteria ("early skill determinants") and some demographic variables and contemporary skill determinants that are available in more standard datasets. In terms of the early skill determinants, I construct intuitive composite measures of mathematical, verbal, and mechanical talent by combining test scores on mathematics knowledge, paragraph comprehension and word knowledge, and mechanical comprehension and auto- and shop information, respectively. In addition, I report the AFQT score, which is commonly taken as a measure of general intelligence.[15]

[13] Also other studies find a further decrease in low-skill compared to middle-skill wages (Goos and Manning 2007). Autor and Dorn (2012) find that relative wages in clerical occupations rise while quantities fall.

[14] Thus, the popular non-cognitive skill measures of locus of control and self-esteem have to be left out of the analysis because they are not available in the NLSY97.

[15] All these measures are taken from the Armed Services Vocational Aptitude Battery of tests (ASVAB) which consists of ten components: arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, general science, numerical operations, coding speed, auto and shop information, mechanical comprehension, and electronics information. The breakup into mathematical, verbal, and mechanical talent is very similar to what a factor analysis of test scores suggests. AFQT is essentially the average of arithmetic reasoning, word knowledge, paragraph comprehension, and mathematics knowledge.

The advantages of the early skill determinants—and in particular the composite measures of mathematical, verbal, and mechanical talent—compared to the contemporary skill determinants—and in particular measures of education—for my study are threefold: First, the early skill determinants are largely exogenous to an individual's actual occupational choice as they are hardly malleable and determined before entry into the labor market. Second, the test scores are finer measures of individual differences in skill than education, which has a lot of bunching at points like high school graduate (12 years of education) or college graduate (16 years of education). This is a sizeable advantage when I want to use test scores to compare similarly skilled individuals over the two cohorts. And finally, the test scores provide proxies for multiple dimensions of individuals' skills. Thus, they can be used to determine comparative advantage as I show in the next subsection.

Before moving on, we see from table 2 that the level of AFQT, which is a measure of IQ, does not change in the male labor force over the two cohorts. In addition, table 3 reports that the cross-correlation of the composite test scores and AFQT remained virtually the same. This supports my identification assumption in the following that the tests measure similar dimensions of talent over the two cohorts and that "within test score groups" individuals can be considered on average the same across cohorts.

2.2.2 Sorting into Occupations

Figure 4 depicts average mathematical, verbal, and mechanical talent in the three occupation groups in both cohorts. We see that the levels of all three talents are much higher in the high-skill occupation than in the middle-skill occupation which, in turn, is higher than the low-skill occupation. Thus, there is a clear ordering of absolute advantage in occupations independent of the talent considered. This underlines the appropriates of the classification of high-, middle-, and low-skill occupations.

Yet, in the absence of restrictions to enter occupations, workers' choice should not be governed by their absolute but by their comparative advantage and thus depend on their relative skills (for details, compare Sattinger 1993). We see in figure 4 that average mathematical talent in the high-skill occupation is higher than average verbal or mechanical talent, while average mechanical talent is considerably higher in the middle-skill occupation than mathematical or verbal talent. Verbal talent is higher than mathematical and mechanical talent in the low-skill occupation.

This strongly suggests sorting according to comparative advantage as in the well-known Roy model—with workers who have high math talent choosing the high-skill occupation, workers who have relatively high mechanical talent choosing the middle-skill occupation, and workers who have relatively high verbal talent choosing the low-skill occupation. It is also intuitive, since high analytical skills are required to pursue a career in managerial, professional, or technical jobs while individuals who have relatively strong mechanical skills or a practical inclination may prefer to work in production or clerical jobs. Verbal skills may be relatively helpful to communicate in personal and protective service occupations. In this case, the uniform absolute ranking of occupations in the three talents should stem from the high cross-correlations between them as seen in table 3.

To test the idea of sorting according to comparative advantage I run multinomial choice regressions. Let $\{K_{it}\}$ be a set of indicator variables that take the value of 1 when individual i works in occupation $K \epsilon \{L, M, H\}$ and zero otherwise. The timing is such that $t = 0$ when the members of the NLSY79 are 27 years old and $t = 1$ when the members of the NLSY97 are 27 years old. For now, I model the conditional choice probabilities as multinomial logit (MNL):[16]

$$p(K_{it} = 1 | x_{it}, \pi_t) = \frac{exp(b_{K0t} + b_{K1t}x_{1it} + ... + b_{KJt}x_{Jit})}{\sum_{G=H,M,L} exp(b_{G0t} + b_{G1t}x_{1it} + ... + b_{GJt}x_{Jit})}. \quad (1)$$

Maximum likelihood estimation of equation (1) yields the coefficients of this model and it provides conditional probabilities ("propensities") to enter each occupation based on the observable talents. As I show in the next section, these propensities can

[16]This is a commonly made modeling decision because the MNL is convenient to work with. For example, the relative risk of choosing occupation K rather than the base category M becomes

$$log\left[\frac{p(K_{it} = 1)}{p(M_{it} = 1)}\right] = (b_{K0t} - b_{M0t}) + (b_{K1t} - b_{M1t})x_{1it} + ... + (b_{KJt} - b_{MJt})x_{Jit}.$$

Using a multinomial probit (MNP) model with uncorrelated disturbances across options instead of the MNL would have been a natural choice, too. Although more difficult to interpret, the MNP has the attraction of being motivated by a latent normal random vector. Empirically, there is often little difference between the predicted probabilities from probit and logit models (see Cameron and Trivedi 2005, p489ff) and in particular my results are robust to using the MNP. Both, the MNL and the MNP, invoke an Independence of Irrelevant Alternatives (IIA) assumption (i.e. uncorrelated errors) which is too restrictive if one wants to interpret the regression coefficients as structural parameters of an economic model.

be interpreted as individuals' predicted relative skills in an occupation as opposed to the other two occupations. However, note that the descriptive choice regressions do not in general identify any parameters of the economic model that I introduce then.

Table 4 reports the results from the multinomial choice regressions. These extract the marginal effect of another unit of each talent on occupational choice when the respective other talents are held constant. For ease of discussion, focus on the first column which gives the sorting into high- and low-skill occupations relative to the omitted middle-skill occupation in the NLSY79. Conditional on the other talents, a one unit higher math score is associated with an about 4.7 percent higher probability to enter the high-skill versus the middle- or the low-skill occupation. A one unit higher mechanical score is associated with a 1.4 and 2.3 percent lower probability to enter the high- and the low-skill occupation as opposed to the middle-skill occupation, respectively. On the other hand, a one unit higher verbal score decreases the probability to enter the middle- as opposed to the high- or the low-skill occupation by about two percent. Thus, the idea of sorting according to comparative advantage is strongly supported by these regressions—with workers who have (conditionally) high math skills moving into the high-skill occupation, workers with conditionally high mechanical and low verbal skills moving into the middle, and those workers with low math and mechanical skills moving into the low-skill occupations. Also, the results underscore the importance of measuring multiple dimensions of skill for linking occupational demand to workers' comparative advantage in my data. They are the same when looking at the NLSY97 in figure 4 and in column three of table 4.

Finally, the regressions in columns two and four of table 4 are run for creating the propensities to enter occupations based on observables that are used in the following. The test scores are split into terciles in order to also allow for a U-shape in the change in demand for skill levels. Moreover, normalized measures of illicit activities and engagement in precocious sex are added. The regressions omit parental education because is is not available for about a third of respondents. However, the results below are qualitatively robust to adding parental education, omitting the risky behavior measures, or using the regressions in columns one and three for creating propensities.

3 Theory and Econometric Methods

On the one hand, as explained in the introduction, the large body of research on job polarization indicates that the drop of employment in the middle-skill occupations is due to a decrease in demand. On the other hand, the empirical analysis of occupational choice shows that there is systematic sorting with respect to talents in the NLSY data. This naturally motivates a Roy model of occupational choice in order to analyze the effect of demand changes on the supply side.

In this model, a given worker i chooses the occupation that offers him the highest log wage:

$$w_{it} = max\{w_{Hit}, w_{Mit}, w_{Lit}\}, \tag{2}$$

where $\{H, M, L\}$ indexes the high-, middle-, and low-skill occupation, respectively. The timing is such that $t = 0$ when the members of the NLSY79 are 27 years old and $t = 1$ when the members of the NLSY97 are 27 years old. The w_{Kit}s with $K \epsilon \{H, M, L\}$ can more generally be utility levels in each occupation.

As seen above, the NLSY provides a multidimensional array of relevant talent proxies for each respondent. Thus, the log occupational wages can be written as a sum of log prices and quantities of occupation-specific skills in the following way:

$$w_{Kit} = \pi_{Kt} + s_{Kit} = \pi_{Kt} + \beta_{K0} + \beta_{K1}x_{1it} + ... + \beta_{KJ}x_{Jit} + u_{Kit}, \tag{3}$$

where π_{Kt} is the price per unit of skill in occupation K, s_{Kit} individual i's specific skill in occupation K, $x_{it} = [x_{1it}, ..., x_{jit}, ..., x_{Jit}]'$ are the observed talents, the β_{Kj}s are the corresponding linear projection coefficients, and u_{Kit} is an orthogonal regression error which represents the unobserved component of skill in occupation K. This linear factor formulation is adopted from Heckman and Sedlacek (1985).[17]

The demand side hypothesis about job polarization in terms of this model is therefore

$$\triangle(\pi_H - \pi_M) > 0 \text{ and } \triangle(\pi_L - \pi_M) > 0, \tag{4}$$

[17] Contrary to Roy (1951) or Heckman and Sedlacek (1985) I will not make a distributional assumption on the unobserved component of skill in the following. Moreover, the primary interest is not in the sectoral distribution of skills and wages, but in changes in returns to occupation-specific skills.

i.e. the relative occupation-specific skill price in the middle falls compared to the high- and the low-skill occupation. The polarization hypothesis examined in the following has two components: first, that equation (4) is true, and second, that it is the reason for the U-shape change in the wage distribution.

The assumption that it is the occupation-specific skill prices that are changing under polarization is crucial. This is in fact the same as in much of the existing literature on job polarization, which models the effect of shifting demand for tasks or occupations on labor supply via changing wage rates. For example, the driving force on the labor market in the original papers of Autor, Levy, and Murnane (2003) and Autor, Katz, and Kearney (2006) is a drop in the relative wage rate for the routine task due to computerization. In Acemoglu and Autor (2010), similar to this paper, the authors analyze how technological change and offshoring alter wages and worker sorting via the relative price of the tasks corresponding to low-, middle-, and high-skill occupations.[18]

In the following I analyze whether the simple assumption of shifting occupation-specific skill prices may get us all the way to explaining the change in the wage distribution over the last decades. Since the theoretical argument and explanation of empirical methods is rather involved and the general case requires complex notation, I use a maximally simplified version of the model for the rest of this section. The results can be extended to the general case for the empirical analysis.

3.1 A Simplified Model

In order to strip the model of equations (2)-(4) to its essence, assume there are only two occupations, middle M and nonmiddle N, with $\triangle(\pi_N - \pi_M) > 0$ according to the polarization hypothesis. Moreover, there is only one observable talent x_i with mean zero ($E(x_i) = 0$) and variance one ($Var(x_i) = 1$), and β_{K0} is zero. I indicate the difference between N and M sector variables by a tilde, i.e. $\tilde{\pi}_t \equiv \pi_{Nt} - \pi_{Mt}$, $\tilde{\beta} \equiv \beta_N - \beta_M$, and $\tilde{u}_i \equiv u_{Ni} - u_{Mi}$. I suppress the index t for x_i and u_{Ki} because the only variables that change in the model are the prices π_{Nt} and π_{Mt} and their

[18]Other papers that make essentially the same assumption include Cortes (2012) and Liu and Trefler (2011).

functions. Wages in occupations $K \epsilon \{N, M\}$ become:

$$w_{Kit} = \pi_{Kt} + s_{Ki} = \pi_{Kt} + \beta_K x_i + u_{Ki} \tag{5}$$

For intuition, we can think of x_i as math talent where a high value is associated with the non-middle occupation and higher wages in the initial period.

How do the workers who have a comparative advantage in the middle occupation fare over time? Since I do not observe the same individual workers in both points in time (the counter-factual), the prediction from the Roy model will have to be in terms of conditional moments with respect to observable talents. Let K_{it} be an indicator variable that takes the value of 1 when individual i works in occupation K and zero otherwise and consider his expected wage conditional on his observable x_i:

$$E(w_{it}|x_i) = E(w_{Mit}|x_i, N_{it} = 1) + p_N(x_i, \tilde{\pi}_t) \left[E(w_{Nit}|x_i, N_{it} = 1) - E(w_{Mit}|x_i, N_{it} = 0) \right],$$

where the notation

$$p_N(x_i, \tilde{\pi}_t) \equiv p(N_{it} = 1|x_i) = Pr(\tilde{u}_i > -(\tilde{\pi}_t + \tilde{\beta} x_i))$$

emphasizes the fact that the probability to enter occupation N is a function of the differences in price per unit of skill between the two occupations. All of the economics of the Roy model can be found in this equation because the probability $p_N(x_i, \tilde{\pi}_t)$ and the conditional wages

$$E(w_{Kit}|x_i, K_{it}) = \pi_{Kt} + \beta_K x_i + E(u_{Ki}|x_i, K_{it} = 1)$$

are determined by the worker's optimal choice given his skills and the prices that he faces. Note that $\tilde{\beta} x_i$ is the expected relative skill given x_i and, for a given $\tilde{\pi}_t$, $p_N(x_i, \tilde{\pi}_t)$ is a monotone function of it. The propensity to enter occupation N for worker i estimated from the data can thus be interpreted as a predictor of his relative skill in occupation N.

Under the price change of polarization $\triangle(\pi_N - \pi_M) > 0$, the change in the conditional expected wage from $t = 0$ to $t = 1$ can be approximated as a sum of three components:

$$\triangle E(w_{it}|x_i) = \triangle\pi_M + p_N(x_i,\tilde{\pi}_0)\triangle(\pi_N - \pi_M)+ \\ + \triangle p_N(x_i,\tilde{\pi}_t)\left[E(w_{Ni0}|x_i, N_{it}=1) - E(w_{Mi0}|x_i, N_{it}=0)\right]+ \\ + p_N(x_i,\tilde{\pi}_0)\triangle E(s_{Ni}|x_i, N_{it}=1) + p_M(x_i,\tilde{\pi}_0)\triangle E(s_{Mi}|x_i, N_{it}=0) \tag{6}$$

The first component is the direct price effect, the second the effect of moving out of occupation M (since workers react optimally to the relative price shifts $\triangle p_N(x_i,\tilde{\pi}_t) \geq 0$), and the third a composition effect of skills within occupations. I call the first component the price or wage rate effect and subsume the second and third components under the name reallocation effect. However, without an assumption on the distribution of the unobserved skill vector u_i, one cannot make a prediction on the relative size of these two effects for workers with different observable talents x_i.[19] One way to evaluate the average effect of polarization on workers of different observable talents would thus be to assume the normal distribution and structurally estimate the Roy model in the NLSY79 and the NLSY97 cross-section, respectively. However, without convincing exclusion restrictions or instrumental variables that affect only occupational choices but not wages, the identification of the parameter estimates would solely rely on the potentially incorrect functional form assumption for the skill distribution.

For this reason, I take a different approach in my paper by starting out from a clear prediction on relative wages for marginal shifts in the π_{Kt}s and then applying it beyond the margin. Consider the change in worker i's wages for a marginal shift in prices:

$$dw_{it} = \begin{cases} d\pi_N & \text{if } N_{it} = 1 \\ d\pi_M & \text{if } N_{it} = 0, \end{cases}$$

where d denotes a marginal change. Thus, due to the optimality of workers' occupa-

[19] Even with a distributional assumption, say normality, $E(w_{Kit}|x_{it}, y_{Kit} = 1)$ and its change remain hard to interpret economically as there is no simple expression for the expectation of the maximum of correlated normal random variables. Results on the truncated normal provided for example in Heckman and Sedlacek (1985) apply only to the bivariate case, so for my more general three-occupation case things get very complicated. Hsieh, Hurst, Jones, and Klenow (2012) use an extreme value distribution to solve the problem, but this comes at the cost of the very strong assumption that individuals' skills are uncorrelated across occupations.

tional choice and the envelope theorem, the effect on wages of a marginal change in π_{Kt}s is only the direct price effect

$$dE(w_{it}|x_i) = d\pi_M + p_N(x_i, \tilde{\pi}_t)d(\pi_N - \pi_M). \tag{7}$$

According to prediction (7), under the polarization hypothesis, workers who are ceteris paribus more likely to enter the nonmiddle occupation are expected to see their relative wages increase. For example, randomly picking two workers from the population, the worker with lower math talent (call him $\bar{m}$ with $x_{\bar{m}} = \bar{m}$) will *on expectation* have a lower wage increase under polarization than the worker with higher math talent (call him m with $x_m = m$) because $p_N(x_{\bar{m}}, \tilde{\pi}_t) < p_N(x_m, \tilde{\pi}_t)$ and $d(\pi_N - \pi_M) > 0$. The nice feature about this result on the margin is that it is solely in terms of variables that I can straightforwardly estimate from the information on wages, occupational choice and my observables, i.e. $E(w_{it}|x_i)$ and $p_N(x_i, \tilde{\pi}_t)$, and parameters that I have hypotheses about or that I want to estimate, i.e. $d(\pi_N - \pi_M) = d\tilde{\pi}_t$.

Prediction (7) also holds qualitatively beyond the margin, i.e. the expected overall wage gain from polarization rises with the initial probability to work in the nonmiddle occupation. Note that the change in worker i's expected wage is the sum over his marginal expected wage changes along the adjustment path from π_0 to π_1. Hence, we can integrate prediction (7) from $t = 0$ to $t = 1$ to obtain:

$$E(w_{i1}|x_i) - E(w_{i0}|x_i) = \triangle\pi_M + \int_{\tilde{\pi}_0}^{\tilde{\pi}_1} p_N(x_i, \tilde{\pi}_t)d\tilde{\pi}_t, \tag{8}$$

where the structure of $p_N(x_i, \tilde{\pi}_t) = Pr(\tilde{u}_i > -(\tilde{\pi}_t + \tilde{\beta}x_i))$ illustrates that on the adjustment path of prices, the ranking of $p_N(x_i, \tilde{\pi}_t)$ with respect to x_i remains unchanged. In terms of the example, if $p_N(x_{\bar{m}}, \tilde{\pi}_0) < p_N(x_m, \tilde{\pi}_0)$ then $p_N(x_{\bar{m}}, \tilde{\pi}_t) \leq p_N(x_m, \tilde{\pi}_t)$ for all $t\epsilon(0, 1]$. Therefore, we expect a higher increase in wages for worker m than for worker $\bar{m}$.[20]

In section 4 I estimate the change in wages associated with $p_N(x_i, \tilde{\pi}_0)$ between the NLSY79 and NLSY97. Because of prediction (7), I expect the return per unit of $p_N(x_i, \tilde{\pi}_0)$ to increase over time. Note, though, that this return change includes the

[20] Another way of deriving equation (8) is illustrative: Concentrate on a specific worker i first and note again that $\tilde{\pi}_t \equiv \pi_{Nt} - \pi_{Mt}$, $\triangle\tilde{\pi}_t > 0$, and N_{it} is an indicator for working in occupation N such that $w_{it} = w_{Mit} + N_{it}(w_{Nit} - w_{Mit})$. Defining the relative price that makes i indifferent as

direct price effect and the reallocation effect discussed in equation (6). In terms of the example, the expected wage increase for worker m versus worker $\bar{m}$ between $t = 0$ and $t = 1$ includes the initial difference in propensities $p_N(x_m, \tilde{\pi}_0) - p_N(x_{\bar{m}}, \tilde{\pi}_0)$ and the change in this difference along the adjustment path. The identification assumption in my data is that the distribution of unobservable skill components conditional on x_i is the same across the NLSY79 and the NLSY97, i.e. that a given value of math talent measures on average the same person in both cohorts. Section 4 explains the details.

3.2 Identifying the Change in the Occupation-Specific Skill Prices

The second and more difficult question is to identify the actual changes in relative prices $\triangle(\pi_{Nt} - \pi_{Mt}) = \triangle\tilde{\pi}_t$. One way or another I will have to make an additional assumption for this and I argue that my approach of choice is particularly attractive for several reasons.

The overall change in worker i's expected wage is the sum over his marginal expected wage changes along the adjustment path from π_0 to π_1 as shown in equation (8). In this equation, I want to estimate $\triangle\tilde{\pi}_t$ and possibly $\triangle\pi_M$. I know $E(w_{t1}|x_i)$ and $p_N(x_i, \tilde{\pi}_t)$ in points in time $t = 0$ and $t = 1$ in the sense that I can consistently estimate them from my primary data. I do not know, however, $p_N(x_i, \tilde{\pi}_t)$ within the interval $t\epsilon(0, 1)$ and I will need to make an assumption on it.

The estimation problem can be nicely illustrated in a graph. In figure 5, I want

$\tilde{\pi}_t^i \equiv -\tilde{s}_i = -(s_{Ni} - s_{Mi})$, we get:

$$\begin{aligned} w_{i1} - w_{i0} &= \triangle\pi_M + N_{i1}(w_{Ni1} - w_{Mi1}) - N_{i0}(w_{Ni0} - w_{Mi0}) \\ &= \triangle\pi_M + \begin{cases} \triangle\pi_N - \triangle\pi_M = \tilde{\pi}_1 - \tilde{\pi}_0 & \text{if } N_{i0} = 1, N_{i1} = 1 \\ \tilde{\pi}_1 + \tilde{s}_i = \tilde{\pi}_1 - \tilde{\pi}_t^i & \text{if } N_{i0} = 0, N_{i1} = 1 \\ 0 & \text{if } N_{i0} = 0, N_{i1} = 0 \end{cases} \\ &= \triangle\pi_M + \int_{\tilde{\pi}_0}^{\tilde{\pi}_1} N_{it} d\tilde{\pi}_t. \end{aligned}$$

Taking expectations w.r.t. $\tilde{u}_i$ conditional on x_i on the top left and bottom of this equation gives result (8). Hence, since within occupations the wage gain is constant, the overall gain for a specific worker depends solely on the "distance" of the adjustment that the worker is still in the middle $(\pi_N^i - \pi_{N0})$ and already in the nonmiddle $(\pi_{N1} - \pi_N^i)$ occupation. This principle is the same for expected wages and probabilities of being in the nonmiddle occupation.

to back out the distance on the x-axis between $\tilde{\pi}_1$ and $\tilde{\pi}_0$ while I know the starting and the end point (the thick dots A_1 and A_2) of the function (the arch) over which I need to integrate and the value of the integral (the shaded area). I thus need to make an assumption about the shape of the curve connecting A_1 and A_2. This curve has to be (weakly) monotonically increasing (as with higher $\tilde{\pi}_t$ the number of workers in occupation N will increase) but it can be concave as in the picture or convex.

The first assumption that comes to mind is to simply assume that it is a horizontal line through the point A_1, which implies no reallocation of workers due to the price change and thus to plug $p_N(x_i, \tilde{\pi}_t) = p_N(x_i, \tilde{\pi}_0)$ into (8). In the figure, the difference between $E(w_{i1}|x_i)$ and $E(w_{i0}|x_i)$ is then assumed to be only the rectangle a. This results in the marginal prediction (7) holding exactly for the discrete price change as well and the regression in section 4 on the propensity $p_N(x_i, \tilde{\pi}_0)$ directly identifying the price change. Of course, this is not a good assumption.

A more subtle version of it but essentially the same assumption is to recognize that workers reallocate away from the middle occupation but to impose that the extent of reallocation does not differ across observables x_i. In terms of the example it is to assume that the probability change for the high math worker m is the same as for the low math worker $\bar{m}$, i.e. $\triangle p_N(x_{\bar{m}}, \tilde{\pi}_t) = \triangle p_N(x_m, \tilde{\pi}_t)$. In this case, equation (8) becomes

$$E(w_{i1}|x_i) - E(w_{i0}|x_i) = \triangle\pi_M + const + p_N(x_i, \tilde{\pi}_0)\triangle\tilde{\pi}_t.^{21}$$

Again the regression on $p_N(x_i, \tilde{\pi}_0)$ in section 4 directly identifies the price change. This is also not a good assumption as it does not allow for a differential reallocation effect across worker groups, e.g. that the low math worker $\bar{m}$ may be able to reallocate out of the middle to a larger extent than the high math worker m because the latter is more likely in the nonmiddle to start with. In figure 5 this means that the arch connecting A_1 and A_2 is restricted to be the same no matter where we start off on the y-axis (even if we start off high, i.e. close to probability one).

[21] Suppose $\frac{\partial p_N(x_i, \tilde{\pi}_t)}{\partial \tilde{\pi}_t} = F'(\tilde{\pi}_t) \geq 0$. Then $p_N(x_i, \tilde{\pi}_t) = p_N(x_i, \tilde{\pi}_0) + F(\tilde{\pi}_t) - F(\tilde{\pi}_0)$ and

$$const = \int_{\tilde{\pi}_0}^{\tilde{\pi}_1} [F(\tilde{\pi}_t) - F(\tilde{\pi}_0)] d\tilde{\pi}_t.$$

A seemingly attractive alternative would be to assume that $\tilde{u}_i$ is normally distributed (for simplicity assume $\tilde{\sigma} = 1$), which modifies (8) to

$$E(w_{i1}|x_i) - E(w_{i0}|x_i) = \triangle\pi_M + \int_{\tilde{\pi}_0}^{\tilde{\pi}_1} \Phi(\tilde{\pi}_t + \tilde{\beta}x_i)d\tilde{\pi}_t.$$

For this to be helpful, I need to know the structural parameter $\tilde{\beta}$ from the model. I could in principle estimate it from a probit model or a Heckman two stage regression.[22] But then I am estimating the price change by relying on a distributional assumption in (8) and, in order to implement it, estimating the necessary parameter $\tilde{\beta}$ relying on the distributional assumption in the first stage. This appears to be no improvement to outright structurally estimating the Roy model with a normality assumption in both cross-sections and comparing the estimated $\tilde{\pi}_0$ and $\tilde{\pi}_1$.

I therefore instead decide for an approach which makes full use of the empirical evidence in $t = 0$ and $t = 1$. I linearly approximate

$$p_N(x_i, \tilde{\pi}_t) \approx p_N(x_i, \tilde{\pi}_0) + \frac{p_N(x_i, \tilde{\pi}_1) - p_N(x_i, \tilde{\pi}_0)}{\tilde{\pi}_1 - \tilde{\pi}_0}(\tilde{\pi}_t - \tilde{\pi}_0). \tag{9}$$

In figure 5, this amounts to approximating $p_N(x_i, \tilde{\pi}_t)$ as the y-coordinate for the point on the line $\overline{A_1A_2}$ that corresponds to $\tilde{\pi}_t$ and by approximating $E(w_{i1}|x_i) - E(w_{i0}|x_i)$ as the trapezoid $a + b$. If the shape of $p_N(x_i, \tilde{\pi}_t)$ in $\tilde{\pi}_t \epsilon (\tilde{\pi}_0, \tilde{\pi}_1)$ is not too convex or concave, the approximation should be reasonably close. Whether it is sufficiently accurate will be tested below.

Equation (8) now becomes

$$E(w_{i1}|x_i) - E(w_{i0}|x_i) = \triangle\pi_M + \frac{p_N(x_i, \tilde{\pi}_1) + p_N(x_i, \tilde{\pi}_0)}{2}\triangle(\pi_N - \pi_M). \tag{10}$$

This is one equation in two unknowns. However, as it holds for all x_i, I could for example identify $\triangle(\pi_N - \pi_M)$ and $\triangle\pi_M$ by imposing it for workers with high and low math talent m and $\bar{m}$, respectively.

A more attractive way to estimate $\triangle(\pi_N - \pi_M)$ is to multiply both sides of equation (10) by x_i and taking expectations. By the law of iterated expectations, this results

[22] In the case of three occupations, this would be multinomial probit with correlated errors or structural estimation of the three-sector Roy model.

in

$$cov(w_{i1}, x_i) - cov(w_{i0}, x_i) = \frac{cov(N_{i1}, x_i) + cov(N_{i0}, x_i))}{2} \triangle(\pi_N - \pi_M), \tag{11}$$

where $cov(w_{it}, x_i)$ is the coefficient from a linear wage regression of w_{it} on x_i and $cov(N_{it}, x_i)$ the coefficient from a linear allocation regression of occupational dummy N_{it} on x_i.

If I had just one talent as in this simple example, I could exactly solve equation (11). Yet, as I have J different talents in my empirical implementation, prediction (11) has to hold for each single one of them so that I get J different moment conditions

$$m_j(\triangle\tilde{\pi}_t) = cov(w_{i1}, x_i) - cov(w_{i0}, x_i) - \frac{cov(N_{i1}, x_i) + cov(N_{i0}, x_i)}{2} \triangle(\pi_N - \pi_M) = 0$$

from the model. I can stack those moment conditions in a column vector and apply the minimum distance estimator for $\triangle\tilde{\pi}_t$ which minimizes the quadratic form:

$$m(\triangle\tilde{\pi}_t)' W m(\triangle\tilde{\pi}_t), \tag{12}$$

where the asymptotically optimal W takes into account the variance-covariance matrix of the first-stage estimates of $cov(w_{it}, x_i)$ and $cov(N_{it}, x_i)$. The objective function (12) in optimum also provides a test statistic for the joint test of the polarization hypothesis and my linear approximation of the reallocation adjustment path. Section 5 details and implements this estimation and testing procedure in the more general case of three occupations in my data.

Overall, the procedure of estimating the relative price changes described here has two advantages over the standard approach of estimating the Roy model under normality. It should give relative price estimates that are close to the actual prices and at the same time be robust to different distributions of unobserved skills, and it is transparent and easy to implement.

4 Polarization's Effect on Observable Skills

How do the workers who have a comparative advantage in the high-, middle-, and low-skill occupation fare over time? This section analyzes the effect of polarization on the returns to propensities to enter occupations and on absolute skill measures.

4.1 Prediction

Predictions (7) and (8) generalize to the three-occupation case (for detailed derivation see Appendix B.1):

$$dE(w_{it}|x_{it}) = d\pi_{Mt} + p_H(x_{it}, \pi_t)d(\pi_{Ht} - \pi_{Mt}) + p_L(x_{it}, \pi_t)d(\pi_{Lt} - \pi_{Mt}), \tag{13}$$

and

$$\begin{aligned} E(w_{i1}|x_{i1}) - E(w_{i0}|x_{i0}) = \triangle\pi_M + \int_{\pi_{H0}-\pi_{M0}}^{\pi_{H1}-\pi_{M1}} p_H(x_{it}, \pi_t)d(\pi_{Ht} - \pi_{Mt}) + \\ + \int_{\pi_{L0}-\pi_{M0}}^{\pi_{L1}-\pi_{M1}} p_L(x_{it}, \pi_t)d(\pi_{Lt} - \pi_{Mt}). \end{aligned} \tag{14}$$

where $p_K(x_{it}, \pi_t)$ is the probability of working in occupation $K \epsilon \{H, M, L\}$ under the price vector π_t. Moreover, I now give a time subscript to the observable characteristics to indicate which dataset they are from.

Hence, under the polarization hypothesis (4), workers who are ceteris paribus more likely to enter the high- and the low-skill occupation are expected to see their relative wages increase. In order to evaluate this, I estimate ordinary least squares (OLS) regressions for pooled data of the form

$$\begin{aligned} w_{it} = \alpha_0 + \alpha_1 p_H(x_{it}, \pi_0) + \alpha_2 p_L(x_{it}, \pi_0) + \alpha_3 \times NLSY97 + \\ + \alpha_4 p_H(x_{it}, \pi_0) \times NLSY97 + \alpha_5 p_L(x_{it}, \pi_0) \times NLSY97 + \varepsilon_{it}, \end{aligned} \tag{15}$$

where NLSY97 is a dummy for whether a particular observation is from the NLSY97 and $p_H(x_{it}, \pi_0)$ and $p_L(x_{it}, \pi_0)$ are the probabilities to choose the high- and the low-skill occupation in the NLSY79, i.e. under the old prices. Hence, the approach is to hold groups of workers constant over time in terms of their predicted occupation-specific skills (the probabilities) and study their average wages over time. According to prediction (13), I expect the parameter estimates for α_4 and α_5 to be positive.[23] There are no predictions from the theory on α_1, α_1 and α_3, although one would think that a higher probability to enter the H and the L occupation is associated with higher and lower wages, respectively.

Of course, the occupational choice probabilities are not directly available in the data and they have to be estimated in a preceding step in the NLSY79. The parameter estimates are then used to predict $p_H(x_{it}, \pi_0)$ and $p_L(x_{it}, \pi_0)$ for each individual in the NLSY79 and the NLSY97. This makes the estimation of (15) a two-step procedure. In fact, I am using two-step estimation procedures throughout this paper since my empirical strategy exploits measuring comparative advantage in occupations with respect to observable talents and then relating this comparative advantage to changes in the returns to talents:

> "[...] the approach here exploits the fact that task specialization in the cross section is informative about the comparative advantage of various skill groups, and it marries this source of information to a well-specified hypothesis about how the wages of skill groups that differ in their comparative advantage should respond [...]"

These are the words of Acemoglu and Autor (2010, p78) who suggest the same procedure in their well-known survey paper but lack the data that I have to implement it satisfactorily.

In terms of the two-step procedure used here, two clarifications are in order. First, different functional form assumptions can be used to specify $p_K(x_{it}, \pi_0)$. A linear probability model, i.e. OLS regression, provides the best linear estimator for the probabilities but some predicted values from it will be above one and below zero, i.e. they are not probabilities themselves. Therefore, many researchers would prefer a multinomial logit or probit model. I report the results from the multinomial logit that I ran in table 4 in the following but my results do not change if I use the other

[23] In general, regression (15) provides the best linear predictor of

$$\triangle E\left(w_{it}|p_H(x_{it}, \pi_0), p_L(x_{it}, \pi_0)\right) = \alpha_3 + \alpha_4 p_H(x_{it}, \pi_0) + \alpha_5 p_L(x_{it}, \pi_0).$$

options to specify $p_K(x_{it}, \pi_0)$.

Second, the standard errors in the second stage regression (15) have to reflect the fact that $p_H(x_{it}, \pi_0)$ and $p_L(x_{it}, \pi_0)$ are estimates and thus possess sampling variation. Among others, Murphy and Topel (1985) provide a procedure to do this, which is however somewhat tedious.[24] Therefore I report bootstrapped standard errors instead, which are also asymptotically consistent.

Note that, although they identify the average relative wage changes for workers of different observables x_{it} due to polarization, the parameter estimates for α_4 and α_5 do not identify the structural relative price changes $\triangle(\pi_H - \pi_M)$ and $\triangle(\pi_L - \pi_M)$. This is because, as we have seen in equation (6), the conditional wage changes for different x_{it} consist of a combination of the direct price effect and a reallocation effect. As discussed at length in section 3.2, the latter may differ across worker groups, while interpreting α_4 and α_5 as the relative price changes would impose that it is the same across x_{it}.

4.2 Results

Table 5 reports the results from wage regressions a la (15) on the propensities to enter the high- and the low-skill occupation in the NLSY79 and the NLSY97. As expected, in column one we see that a higher propensity to enter the high-skill occupation compared to the omitted middle-skill occupation is associated with a significantly higher wage. The reverse is true for the propensity to enter the low-skill occupation.

The prediction from polarization in equation (13) is however about changes in returns to propensities over time, which are indicated in the table by (x NLSY97). We see that the coefficients change strongly and significantly in the expected direction. For the propensity to enter the high-skill occupation, the coefficient almost doubles (from .31 to .60) while the coefficient for entering the low-skill occupation rises by almost a third (from -1.65 to -.95). The level of the change in the low-skill coefficient is twice that of the high-skill coefficient, which may come as a surprise. However, note that it is also much less precisely estimated. Moreover, when scaling the size of

[24]Two stage least squares or joint estimation (in ML or GMM) of step one and two in a standard statistical package would be a convenient option to get the correct standard errors automatically. However, this is not feasible here as for the individuals in the NLSY97 the regressors are estimated in a different dataset.

the effect by the respective standard deviations of the propensities, the change in the effect of the propensity to enter the high-skill occupation is larger: a one standard deviation increase in the high- and low-skill propensities, respectively, is associated with a 11.3 percent higher and 5.2 percent lower wage in the NLSY97 compared to a 5.9 percent higher and 8.4 percent lower wage in the NLSY79.[25]

For illustration of the effect of different propensities to enter the three occupations, figure 6 plots the predictions from linear wage regressions on each propensity at a time together with their probability densities.[26] In the top left sub-figure we see the positive effect of having a higher propensity to enter the high-skill occupation in the NLSY79 indicated by the upward-sloping line. This effect increases further in the NLSY97 as the dashed line is even steeper. In the top right sub-figure, we see that there is a strong negative effect of the propensity to enter the low-skill occupation, which is however less severe in the NLSY97. Moreover, we see again that the range of propensities to enter the low-skill occupation is very limited in the data. Finally, for the propensity to enter the middle-skill occupation there is already a negative effect in the NLSY79 but this becomes substantially more negative in the NLSY97. For individuals with a very high propensity to enter the middle, which is quite frequent in the data, expected real wages even decline during the two decades between the NLSY79 and the NLSY97. This is indicated by the crossing of the two lines.

The identification of changes in returns to propensities in regression (15) is based on the assumption that for a given vector of talents x_{it} workers are in expectation the same in terms of their relative labor market productivities over the two cohorts. Tables 2 and 3 provided support for this assumption as they showed that the level and cross-correlation of observable early skill determinants is very similar in the NLSY79 and NLSY97. Consequently, unreported descriptive statistics show that the distribution of predicted propensities is very similar in the NLSY79 and NLSY97, i.e. that the distribution of relative occupational skills according to my observable measures has not changed over the two cohorts. Combined, these pieces of evidence

[25]For the NLSY79 multiply the coefficients on the propensities to enter the high- and low-skill occupations of 0.31% and -1.65% by the standard deviations of these propensities of 19.0 and 5.1. For the NLSY97 multiply the coefficients on the propensities to enter the high- and low-skill occupations of 0.60% and -0.95% by the standard deviations of these propensities of 18.8 and 5.5.

[26]The coefficients and standard errors from these wage regressions on each propensity separately are not reported in a table for saving space.

lend substantial support to my identification assumption.[27]

Given this identification assumption, the changes of the propensity coefficients provide the increase in average wages that is associated with relative advantage in the high- or the low-skill occupation compared to the middle. The workers in the NLSY97 entered the labor market only recently when the bulk of the occupational demand change had likely already taken place. Hence, polarization's effect on their relative wages mostly reflects returns changes to ex ante relative talent differences and not to skills that they acquired in a specific occupation. Identifying that there exists a substantial ex ante effect is relevant for policy makers as it implies that the relative earnings effects of polarization will not fade over time and that temporary policy responses are therefore not sufficient.

The result in column one of table 5 does not exclude the possible influence of other factors than polarization on wages of workers with comparative advantage in the high- or the low-skill occupation. In particular, skill-biased technological change that is independent of occupational demand constitutes an alternative hypothesis to polarization and may thus have an important effect on talent returns. According to this view, comparative advantage in occupations is not important because returns to skills change across the board. The SBTC amounts to $d\beta_{Kj} = d\beta_j$ in my framework and it is easily incorporated in prediction (13) in addition to polarization:[28]

$$
\begin{gathered}
dE(w_{it}|x_{it}, \pi_t) = d\pi_M + p_H(x_{it}, \pi_t)d(\pi_H - \pi_M) + p_L(x_{it}, \pi_t)d(\pi_L - \pi_M) + \\
+ d\beta_0 + d\beta_1 x_{1it} + ... + d\beta_J x_{Jit}
\end{gathered}
$$

When allowing for SBTC with all the talents included on top of polarization, the identification will have to rely on the functional form of $p_H(x_{it}, \pi_t)$ and $p_L(x_{it}, \pi_t)$, because the same variables that are used for estimating the propensities are directly entered into the wage regression. This may potentially lead to near multicollinearity of the explanatory variables in the regression and imprecise estimates. In additional regressions, I thus use education indicators as absolute skill measures.

The remaining columns of table 5 assess the potential importance of the SBTC

[27]The racial distribution does however change over the cohorts. Therefore, I control for race in all my analyses.

[28]Actually, SBTC may predict that the return to $p_L(x_{it}, \pi_t)$ falls instead of rises.

hypothesis versus polarization. Column two adds a dummy of whether the individual completed a four-year college or more to the regression. We see that the level of the coefficient on the propensity to enter the high-skill occupation drops all the way to zero but that the changes in both coefficients are remarkably stable. On the other hand, the level of return to college is large and highly significant while its change does not significantly increase once I control for the propensities. The result is similar if I control for four different degree dummies (high school dropout and graduate, some college, and at least four year college) in column three.[29] This indicates that Mincerian returns to education are important to explain wages in the cross-section, but that they have much less power than relative skills in occupations to explain the change in wages that took place over the twenty years from the NLSY79 to the NLSY97.

Finally, the regression reported in column four of the table adds the same specification of talents that I use to estimate the occupational propensities in the first place. The parameter estimates on the propensities remain in the right direction and become even stronger but they also become very imprecise and insignificant, which is due to the high degree of multicollinearity between the regressors in this specification. Therefore, the regression is not as informative as the preceding ones.

How much of the U-shape change in the wage distribution can the changing returns to observable skills explain? Figure 7 plots the actual and the predicted change in the wage distribution when the changing coefficient values from the regressions reported in columns one and four of table 5 are assigned to workers' wages in the NLSY79. As we can see, the propensities to enter occupations with their functional form restriction do not do a worse job in matching the wage distribution than a very flexible specification of the same talents that are included in estimating the propensities. However, both options do not explain are large share of the change in the wage distribution. This is not surprising since the observables also only explain a relatively small share of variation in wages in the cross-section. The remainder should thus be explained by changes in returns to unobservable occupational skills (u_{Kit} in the notation of the model).

[29]The coefficient estimates on the degree dummies and the talents included in column three and four of the table do not provide additional insight and are not reported in order to save space.

To sum up, I conclude that the results reported in this section indicate a substantial longterm decline in relative wages of workers with comparative advantage in the middle-skill occupation. Moreover, the driver of this decline is more likely to be relative demand changes for occupations as implied by the polarization hypothesis than increases in absolute returns to skills that are detached from comparative advantage (SBTC). Nonetheless, the analysis so far remains unsatisfactory in two dimensions: it does not formally exclude other drivers of skill returns than polarization and, because a substantial part of skill is unobserved, the changing returns to observable talents can naturally only hope to match part of the wage distribution. The next sections tackle these two shortcomings.

5 Estimating the Change in Occupation-Specific Skill Prices

The last section provided convincing evidence for polarization to have driven workers' skill returns over the past two decades. In this section, I formally test whether the polarization model can explain the whole variation in observable skill returns in the NLSY via a test of over-identifying restrictions. The procedure yields an estimate of the implied change in occupation-specific skill prices. In the next section, I use this estimate to explore how much of the U-shape change of wage inequality can be explained by relative price changes across occupations and the potential role of reallocation to explain the rest.

5.1 Methodology

A more detailed assessment of the effect of polarization looks at each talent in turn. I use the fact that I observe $x_{it} = [x_{1it}, ..., x_{Jit}]'$ and that individuals have comparative advantages in occupations varying with each x_{jit} in order to over-identifying restrictions from the polarization hypothesis. The intuition is that the return to a talent should change depending on which occupational choice it predicts and how that changes.

Linearly approximating the probabilities under the integral in prediction (14) as discussed in relation to figure 5 (see also Appendix B.1), and writing in terms of regression coefficients gives:

$$\triangle\gamma_j = \frac{\delta_{Hj0} + \delta_{Hj1}}{2}\triangle(\pi_H - \pi_M) + \frac{\delta_{Lj0} + \delta_{Lj1}}{2}\triangle(\pi_L - \pi_M), \tag{16}$$

where $\delta_{Kjt} = \frac{cov(K_{it}, x_{jit})}{var(x_{jit})}$ with $\delta_{Hjt} + \delta_{Mjt} + \delta_{Ljt} = 0$, K_{it} is an indicator for working in occupation K, and $\gamma_{jt} = \frac{cov(w_{it}, x_{jit})}{var(x_{jit})}$. These parameters can be recovered from OLS allocation

$$K_{it} = \delta_{K0t} + \delta_{K1t}x_{1it} + \delta_{K2t}x_{2it} + ... + \delta_{KJt}x_{Jit} + v_{Kit} \tag{17}$$

and wage regressions

$$w_{it} = \gamma_{0t} + \gamma_1 x_{1it} + \gamma_2 x_{2it} + ... + \gamma_J x_{Jit} + u_{it}.^{30} \tag{18}$$

Therefore, result (16) provides a simple to implement procedure to assess polarization's effect on the returns to detailed talents. I have data on individuals' talents, their choices of entering high, middle, or low-skill occupations, and their wages in the periods before ($t = 0$) and after ($t = 1$) polarization took place. First, I run four allocation regressions (17) for $K = H$ and $K = L$ in $t = 0$ and $t = 1$, which recover the partial correlations of the observed talents and occupational choices δ_{Kjt}. Second, I run two wage regressions (18) for $t = 0$ and $t = 1$, which recover the partial correlations of the observed talents and wages γ_{jt} in each period. Then, according to condition (16), the change of a talent's effect on the wage equals its effect in the allocation regressions times the change in relative prices.[31]

Condition (16) is in fact very intuitive. The return to a talent x_{jit} should change by the extent to which, conditional on the other talents, it increases the probability to work in occupations H and L, i.e. δ_{Hj0} and δ_{Lj0}, and the extent to which this association changes, i.e. (δ_{Hj1}-δ_{Hj0}) and (δ_{Lj1}-δ_{Lj0}).

In order to assess the validity of the polarization hypothesis in the data, one could

[30] To be exact, the allocation and wage regressions in fact recover the covariance of K_{it} and w_{it} with the residual of regressing x_{jit} on the other observable talents. This is what I use in the following.

[31] Note that the literature on SBTC has also run linear wage regressions on test scores (e.g. Murnane, Willett, and Levy (1995)). The difference here is that the drivers of returns changes are explicitly examined in the allocation regressions and that the results are interpreted within an explicit model of sorting and occupational demand.

thus simply check whether the returns changes to individual talents line up with what their allocation coefficients imply. However, a more encompassing test of the model recognizes that condition (16) has to hold for all J talents at the same time. Thus, as long as there are more talents than the two unknown model parameters $\triangle(\pi_H - \pi_M)$ and$\triangle(\pi_L - \pi_M)$, I can use the over-identifying restrictions implied in (16) to devise an overall test of the model.

The first step in such a test is to implement a minimum distance estimator for the implied relative wage rate changes. Define $\bar{\delta}_{Kj} \equiv \frac{\delta_{Kj0}+\delta_{Kj1}}{2}$, and stack $\triangle\gamma_j$ and $\bar{\delta}_{Kj}$ into $J \times 1$ vectors. Then, using the first stage estimates $\hat{\triangle\gamma}$ and $\hat{\bar{\delta}}_K$ and defining the $J \times 1$ vector $m(\triangle\pi) = \hat{\triangle\gamma} - \hat{\bar{\delta}}_H\triangle(\pi_H - \pi_M) - \hat{\bar{\delta}}_L\triangle(\pi_L - \pi_M)$, this estimator minimizes

$$Q(\triangle\pi) = m(\triangle\pi)'Wm(\triangle\pi) \tag{19}$$

with respect to $\triangle(\pi_H - \pi_M)$ and $\triangle(\pi_L - \pi_M)$. Depending on the weighting matrix W, the minimizing wage rate changes can be the Equally Weighted Minimum Distance (EWMD) estimator if $W = I$, the Optimal Minimum Distance (OMD) estimator if $W = [Var(m(\triangle\pi))]^{-1}$, and the Diagonally Weighted Minimum Distance (DWMD) estimator if $W = [diag(Var(m(\triangle\pi)))]^{-1}$. The EWMD can be implemented by a simple OLS regression of $\hat{\triangle\gamma}$ on $\hat{\delta}_{Ht}$ and $\hat{\delta}_{Lt}$, the OMD by a (feasible) GLS regression, and the DWMD by weighted least squares.

Just as GLS the OMD is asymptotically optimal and it yields consistent estimates of the relative price changes $\triangle(\pi_H - \pi_M)$ and $\triangle(\pi_L - \pi_M)$. Moreover, the objective function (19) in optimum can be shown to be asymptotically chi-squared distributed with $J - 2$ degrees of freedom:

$$Q(\hat{\triangle\pi}) = m(\hat{\triangle\pi})'[Var(m(\hat{\triangle\pi}))]^{-1}m(\hat{\triangle\pi}) \overset{a}{\sim} \chi^2(J-2)$$

This provides me with an overall test of the cross-equation restrictions implied by the model.

Finally, Altonji and Segal (1996) and Pischke (1995) present evidence for potential bias of the OMD in small samples and recommend using the EWMD and the DWMD in addition, respectively. I thus report results for these two estimators as well. For more details of how I implement the minimum distance estimation and test, please refer to Appendix C.1.

Given optimal worker reallocation, the implied absolute wage change in the middle-skill occupation π_M can be bounded: under the initial prices, the initial worker allocation has to (weakly) dominate the new allocation and vice versa under the new prices. A natural approach is to impose this for average wages. Thus, $\triangle\pi_M$ has to be such that

$$\triangle E(w_{it}) \geq \triangle\pi_M + p_H(\pi_0)\triangle(\pi_H - \pi_M) + p_L(\pi_0)\triangle(\pi_L - \pi_M)$$

since otherwise it would yield higher wages if workers had stayed in the old allocation and

$$\triangle E(w_{it}) \leq \triangle\pi_M + p_H(\pi_1)\triangle(\pi_H - \pi_M) + p_L(\pi_1)\triangle(\pi_L - \pi_M)$$

since otherwise it would have yielded higher average wages if workers had been in the new allocation from the outset. The sample statistics corresponding to $\triangle E(w_{it})$, $p_H(\pi_t)$, and $p_L(\pi_t)$ are the change in average wages and the fraction of workers in the high- and the low-skill occupations, respectively. I take the midpoint between the two bounds as my preferred point estimate for $\triangle\pi_M$.

Finally, by assigning the estimated price changes to the workers in the NLSY79 and comparing the resulting change in the counterfactual wage distribution to the actual one, I can assess what the contribution of changes in occupational prices is to the overall change in the wage distribution. I also assess what share of the remainder may be due to reallocation.[32]

5.2 Empirical Results

Table 6 reports the reduced form allocation and wage regressions according to equations (17) and (18). In the first two columns, we see that math talent is associated with the high-skill occupation, mechanical talent with the middle-skill occupation, and verbal talent with the high-skill occupation to a lesser degree than math. The illicit activities are associated with not working in the high-skill occupation.

This is quite similar to the results from the MNL sorting regressions in table 4.

[32]In fact, the model does allow for a change in the population supply of talents to play a part. As we saw in table 2, this is however minuscule in the data.

However, contrary to the MNL, the OLS coefficients for each occupation in table 6 are not interpreted with respect to an omitted base occupation but with respect to the other two occupations taken together. Moreover, note that the R-squared for the low-skill occupation allocation regressions is very low, i.e. little of the variation in low-skill occupation choice is explained by the data. This will affect the precision of my relative price change estimates for the low-skill occupation below.

The changes in returns to talents are reported in column three of table 6. The returns to the highest math tercile increase significantly, the returns to mechanical talents fall, and the returns to illicit activities fall as well. This is largely in line with prediction (16). Thus, most of the returns changes to talents are in the direction predicted by the model, apart from verbal talents whose returns decline. Yet, with exception of the top math tercile and illicit activities, the changes are not statistically significant by themselves.

Overall, thus, the results from table 6 are neither clearly in favor of- nor against the polarization hypothesis. The formal test of the restrictions implied by prediction (16) across all talents may therefore be quite informative. Table 7 reports the results from this test and the implied occupation-specific skill price change for the asymptotically optimal minimum distance estimator and the two alternatives suggested by Altonji and Segal (1996) and Pischke (1995). The EWMD, which amounts to OLS estimation, is also the first step of the feasible GLS procedure to implement the OMD.

In the OMD, the point estimates of $\triangle(\pi_H - \pi_M)$ and $\triangle(\pi_L - \pi_M)$ are of the expected sign and of substantial magnitude: the wage rates in the high- and the low- compared to the middle-skill occupation increase by 20.1 and 31.4 percent, respectively. The implied absolute wage rate in the middle-skill occupation itself decreases slightly at 2.4 percent. The p-value of the hypothesis test is at 10.7 percent and thus the model is not rejected at conventional significance levels. Furthermore, the estimates for $\triangle(\pi_H - \pi_M)$ are precise and do not change in the two alternative implementations of the minimum distance estimator. In contrast to that, at a standard error of 35.2, $\triangle(\pi_L - \pi_M)$ is imprecisely estimated and it actually drops to negative point estimates in the EWMD and the DWMD.

With this caveat in mind, I use the price estimates from the OMD to evaluate what share of the overall change in the wage distribution is due to the occupation-specific skill prices in the next section.

6 Matching the Change in the Wage Distribution

In this last section, I assess whether the polarization model can in principle account for the change in the overall wage distribution.

First, I use the price estimates from the OMD to evaluate what share of the change in inequality is due to the occupation-specific skill prices. I obtain the skill price effect by assigning the price changes to the workers in the initial period. According to the model, the remaining differences between the actual and the counterfactual wage distribution should then be due to the reallocation effect. I conduct this exercise in the NLSY and in the CPS data from section 2.1. To use the CPS is now possible again because assigning the estimated skill prices only requires knowledge of workers' occupations and not their talents anymore.

Figure 8 displays the effect of the occupation-specific skill prices. We can see that in both datasets the increase in wages at the top of the distribution is quite well explained by the estimated price changes alone. The increase at the bottom is however hardly explained at all, despite the high point estimate of $\triangle(\pi_L - \pi_M) = 31.4\%$. This appears somewhat as a puzzle, since I would have expected that at least part of the increase in the bottom of the wage distribution should be due to higher relative prices in the low-skill occupation.

There are two interrelated reasons for the lack of an increase in the bottom of the counterfactual wage distribution compared to the middle. First, the dispersion of earnings within occupation groups is large, such that the respective occupational wage distributions overlap substantially and that an increase in the price per unit of skill in the low-skill occupation lifts wages of some middle-earners as well. Second, the estimated price changes are large enough such that an "overtaking effect" becomes empirically relevant, whereby some low-wage earners in the low-skill occupations become middle-wage earners and vice versa for some middle-wage earners in middle-skill occupations.[33] Together, these two factors prevent a strong increase of relative wages in the bottom of the counterfactual wage distribution despite the high point

[33] The corresponding statistics are not reported for the sake of brevity.

estimate for the relative price changes.

The results about the difference between actual and counterfactual wages are similar when I use the alternative definitions of occupation groups that have been used in the literature. These include grouping occupations according to initial median wages or average education, splitting up the large middle-skill group into blue collar and white collar occupations, and employing continuous measures of routine and nonroutine (analytical and manual) task content in occupations. As above, all these groupings share the feature that the wage dispersion within them is substantial.[34] However, in the case of tasks, one should note that measurement is far from perfect. This is because tasks that workers carry out are assigned on the three-digit occupation level (for details see the survey paper by Acemoglu and Autor 2010), which may capture only a relatively small share of the overall variation in workers' actual tasks. Therefore, job groupings or task measures that correspond more closely to the tasks that technology and trade have replaced may help to better match the wage distribution, since the dispersion of wages conditional on them may also be lower.[35]

In addition to the change in the overall wage distribution, figure 9 depicts the change in average wages in high-, middle-, and low-skill occupations for the NLSY and CPS. The counterfactual wage increase in the low-skill occupation is much higher than the actual in both datasets, while the increase in middle- and high-skill occupations is lower. Again, this is similar when I use alternative occupational groupings. Overall, hence, it seems that the estimated relative price changes across occupations alone cannot match the empirical facts about wages in the data.

What remains as an explanation, according to the model, is therefore the effect of reallocation on different parts of the wage distribution. In the data, there is a net outflow from the middle- to the low-skill and to the high-skill occupation of three and 3.5 percent of the overall workforce, respectively. I assume that the lowest earners in the middle who make up three percent of the workforce switch into the low-skill

[34] Again, the results on alternative occupational groupings are not reported in detail in order to save space but available from the author upon request.

[35] The occupation groups and task measures that are used here explain only around five to ten percent of the variation in wages in the cross-section. Hence, if it were available for my application, individual-level data on tasks as employed by Autor and Handel (2012) or by Spitz-Oener (2006) for Germany might improve the precision of measurement and the variation in wages that it captures substantially.

occupation and assign them a fifteen percent wage increase, i.e. about half of the maximum wage increase that they could possibly obtain (31.4% − 2.4%).[36] Figure 10 plots the resulting counterfactual wage distribution which fits the actual quite well, especially in the CPS. Moreover, figure 11 displays the corresponding changes of average wages in occupations, which are now also closer to the actual than without reallocation.[37]

Qualitatively, the reallocation effect at the bottom seems plausible. It not only matches better the unconditional wage distribution, but in addition brings occupational wages in the actual and the counterfactual closer together. Moreover, the low-earners in the middle-skill occupations may really have a strong incentive to switch jobs once the relative demand shock hits and it is also conceivable that they could to so gainfully: for example, given probably not too different skill requirements, someone who would have been a low-earning worker in a factory in the 1980s may instead relatively easily become a janitor today.

While qualitatively plausible, the assumptions made about reallocation in order to match the wage distribution in figure 10 are quite strong. Firstly, the concentrated switching of low-earners in the middle-skill occupation requires that the population distribution of skills in the low-skill occupation be very condensed so that the low-earners are the first to find it profitable to "switch down". This is hard to reconcile with the fact that the empirical wage distributions of the low- and the middle-skill occupation overlap substantially in both cross-sections. Secondly, the gains from switching that I need to assume seem high.

Moreover, the assumptions are not strictly testable. This is so because I do not know individual workers' unobserved skills in the occupations that they have not chosen and thus I cannot estimate their overall gains from reallocation. The only assessment I can make is about the gains from reallocation for the observable components of skill. It turns out that according to observable skills there is no

[36] An additional one percent of low earners is assumed to move to the high-skill occupation with the same wage gain.

[37] In fact, the fit may be even better than in figures 10 and 11 if the remaining difference between actual and counterfactual is due to small-sample variation for 27 year olds. For example, I have tried out assigning the same relative price estimates and making similar assumptions about reallocation to the larger group of 25-29 year olds in the CPS. This matches the actual changes almost perfectly. The same is the case if I do the exercise for prime age males aged 25-55.

clear evidence in favor of the idea that the low earners have the highest gains from reallocation. To see this crudely, compare figure 4 again: contrary to what one would expect in the case of strong switching of low-earners out of the middle-skill occupation, the average talent measures in the middle-skill occupation do not improve visibly and they do not deteriorate in the low-skill occupation. Moreover, unreported regressions of the gains from reallocation for observables on workers' wages in the 1980s yield no clear relationship. Finally, I obtain essentially the same results about reallocation when I use the alternative definitions of occupation groups or task measures discussed above.

Therefore, I conclude that, as it is currently implemented in the literature, polarization seems to explain much but not all of the changes in the wage distribution that have occurred over the last decades. Within the polarization story, the most promising avenues for matching the whole wage distribution are to provide evidence for a large reallocation effect at the bottom and to search for more precise empirical measurements of the jobs or the tasks for which demand has declined.[38] However, simply having some more occupation groups or tasks alone will not help much unless the increase in the variation in wages that these finer groups explain is large.

7 Conclusion

This article is the first to study the effect of job polarization on the wage distribution accounting for the endogenous sorting of skills. I do this by employing newly available data from the National Longitudinal Survey of Youth (NLSY) which provides detailed, multidimensional, and pre-determined measures of workers' talents (i.e. test scores) in order to hold different kinds of workers fixed and analyze the returns to occupation-specific skills over time. The estimation equations are derived from a

[38]The strong role for reallocation, if it was substantiated in further research, would be conceptually and economically important. First, workers in fact gain from switching down into an on average lower-paying occupation because they find a better match there. This is a conceptually important point that only models of relative—rather than absolute—advantage can make. It would thus emphasize the fact that there is no one mapping from occupations to the wage distribution. Second, contrary to some existing studies which find strong wage losses from workers switching down in panel data (e.g. Cortes 2012, Ebenstein, Harrison, McMillan, and Phillips 2011, Liu and Trefler 2011), it would suggest that switching may be an important channel to cushion the negative impact of polarization on the lowest earners.

Roy model over two cross-sections with job polarization amounting to a change in the occupation-specific skill prices. In this case, I show that predictions about wage changes depend exclusively on relative occupation-specific skills, which can be measured via the allocation of talents.

My results indicate that a one percentage point higher propensity to work in high- (low-) as opposed to the middle-skill occupations in the base period is associated with a .29 (.70) percent increase in wages over time, and therefore workers with comparative advantage in the middle-skill occupations lose out substantially over time. Furthermore, the effect of job polarization on workers' wages does well to match the changes at the top of the wage distribution but appears unable to wholly explain the changes at the bottom. Thus, occupational demand seems to have been the driving force of a substantial part but not all of the changes in the wage distribution over the past two decades.

These findings suggest that the dismal trend in middle class wages over the last couple of decades may not be fully explained by the changes in technology and globalization that coincided with it. In particular, (relative) incomes in the bottom and the middle of the distribution could have been affected by policy variables and labor market institutions such as the minimum wage and de-unionization. Thus, policies that encourage union formation or other measures that increase workers' bargaining power may be effective in raising middle class wages.

In future research it will be important to examine whether the result that the wage distribution at the bottom cannot be fully explained by demand shocks is robust in other datasets that may become available and for a more precise measurement of the jobs and tasks that may have declined. In addition, similar analyses for European countries, with their different labor market institutions, would help to disentangle the effect of policy instruments on the change in the wage distribution and their interaction with the undoubtedly existent demand shocks.

Finally, the methods developed in this paper can be applied more generally to study the effect of other important demand shocks on the labor market. For example, there are ongoing debates about a long-term increase in the demand for talent in the financial- and related sectors, and about the effect of the Great Recession on the wage distribution. These debates may be vitally informed by the "allocation of talents" perspective.

References

Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

Acemoglu, D., and D. Autor (2010): "Skills, Tasks and Technologies: Implications for Employment and Earnings," Working Paper 16082, National Bureau of Economic Research.

Altonji, J. G., P. Bharadwaj, and F. Lange (2008): "Changes in the Characteristics of American Youth: Implications for Adult Outcomes," Working Paper 13883, National Bureau of Economic Research.

Altonji, J. G., and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, 14(3), 353–66.

Aughinbaugh, A., and R. M. Gardecki (2007): "Attrition in the National Longitudinal Survey of Youth 1997," Mimeo.

Autor, D. H., and D. Dorn (2012): "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market," MIT Working Papers 15150, MIT.

Autor, D. H., D. Dorn, and G. H. Hanson (2012): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," Working Paper 18054, National Bureau of Economic Research.

Autor, D. H., and M. J. Handel (2012): "Putting Tasks to the Test: Human Capital, Job Tasks and Wages," Journal of Labor Economics, forthcoming 15116.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006): "The Polarization of the U.S. Labor Market," *American Economic Review*, 96(2), 189–194.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed The Labor Market?," *The Quarterly Journal of Economics*, 113(4), 1169–1213.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118(4), 1279–1333.

Autor, D. H., A. Manning, and C. L. Smith (2010): "The Contribution of the Minimum Wage to U.S. Wage Inequality over Three Decades: A Reassessment," Working Paper 16533, National Bureau of Economic Research.

Becker, S. O., K. Ekholm, and M.-A. Muendler (2009): "Offshoring and the Onshore Composition of Tasks and Skills," CEPR Discussion Papers 7391, C.E.P.R. Discussion Papers.

Blau, F. D., and L. M. Kahn (2005): "Do Cognitive Test Scores Explain Higher U.S Wage Inequality?," *The Review of Economics and Statistics*, 87(1), pp. 184–193.

Blinder, A. S. (2009): "How Many US Jobs Might be Offshorable?," *World Economics*, 10(2), 41–78.

Bound, J., and G. Johnson (1992): "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *The American Economic Review*, 82(3), pp. 371–392.

Cameron, A., and P. Trivedi (2005): *Microeconometrics: Methods And Applications*. Cambridge University Press.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116(2), 705–746.

Cortes, G. M. (2012): "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data," *Working Paper*.

Crinò, R. (2010): "Service Offshoring and White-Collar Employment," *The Review of Economic Studies*, 77(2), 595–632.

Ebenstein, A., A. Harrison, M. McMillan, and S. Phillips (2011): "Estimating the impact of trade and offshoring on American workers using the current population surveys," Policy Research Working Paper Series 5750, The World Bank.

Firpo, S., N. M. Fortin, and T. Lemieux (2011): "Occupational Tasks and Changes in the Wage Structure," IZA Discussion Papers 5542, Institute for the Study of Labor (IZA).

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998): "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *Journal of Human Resources*, 33(2), 251–299.

Freeland, C. (2012): "A U.S. Campaign Fixated on Another Era," Editorial from August 30, International Herald Tribune, Global Edition of the New York Times.

Gabaix, X., and A. Landier (2008): "Why has CEO Pay Increased So Much?," *The Quarterly Journal of Economics*, 123(1), 49–100.

Garicano, L., and E. Rossi-Hansberg (2004): "Inequality and the Organization of Knowledge," *The American Economic Review*, 94(2), pp. 197–202.

Goldin, C. D., and L. F. Katz (2008): *The race between education and technology / Claudia Goldin, Lawrence F. Katz*. Belknap Press of Harvard University Press, Cambridge, Mass. :.

Goos, M., and A. Manning (2007): "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, 89(1), 118–133.

Gould, E. D. (2002): "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States," *Journal of Labor Economics*, 20(1), 105–147.

Gronau, R. (1974): "Wage Comparisons-A Selectivity Bias," *Journal of Political Economy*, 82(6), 1119–43.

Heckman, J. (1974): "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42(4), pp. 679–694.

Heckman, J. J., and G. Sedlacek (1985): "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93(6), pp. 1077–1125.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2012): "The Allocation of Talent and U.S. Economic Growth," Working paper, Working Paper.

Jaimovich, N., and H. E. Siu (2012): "The Trend is the Cycle: Job Polarization and Jobless Recoveries," Working Paper 18334, National Bureau of Economic Research.

Lemieux, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 96(3), 461–498.

Liu, R., and D. Trefler (2011): "A Sorted Tale of Globalization: White Collar Jobs and the Rise of Service Offshoring," Working Paper 17559, National Bureau of Economic Research.

Machin, S., and J. Van Reenen (2008): "wage inequality, changes in," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke.

MaCurdy, T., T. Mroz, and R. M. Gritz (1998): "An Evaluation of the National Longitudinal Survey on Youth," *Journal of Human Resources*, 33(2), 345–436.

Michaels, G., A. Natraj, and J. Van Reenen (2010): "Has ICT Polarized Skill Demand? Evidence from Eleven Countries over 25 years," Working Paper 16138, National Bureau of Economic Research.

Mulligan, C. B., and Y. Rubinstein (2008): "Selection, Investment, and Women's Relative Wages Over Time," *The Quarterly Journal of Economics*, 123(3), 1061–1110.

Murnane, R. J., J. B. Willett, and F. Levy (1995): "The Growing Importance of Cognitive Skills in Wage Determination," *The Review of Economics and Statistics*, 77(2), pp. 251–266.

Murphy, K. M., and R. H. Topel (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 3(4), pp. 370–379.

Ottaviano, G. I., G. Peri, and G. C. Wright (2010): "Immigration, Offshoring and American Jobs," Working Paper 16439, National Bureau of Economic Research.

Philippon, T., and A. Reshef (2009): "Wages and Human Capital in the U.S. Financial Industry: 1909-2006," Working Paper 14644, National Bureau of Economic Research.

Pischke, J.-S. (1995): "Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study," *Journal of Business & Economic Statistics*, 13(3), 305–14.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

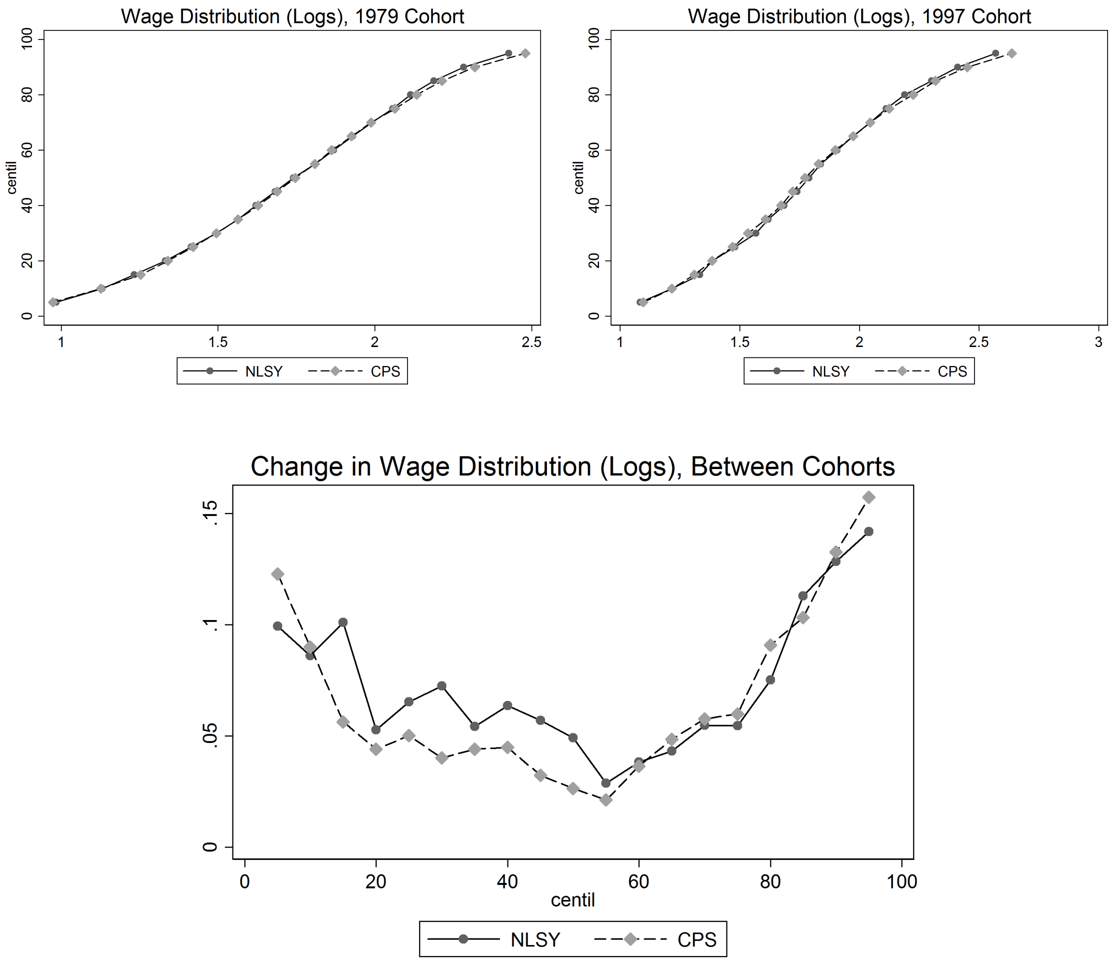
Sattinger, M. (1993): "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31(2), pp. 831–880.

Spitz-Oener, A. (2006): "Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure," *Journal of Labor Economics*, 24(2), 235–270.

Tervio, M. (2008): "The Difference That CEOs Make: An Assignment Model Approach," *American Economic Review*, 98(3), 642–68.

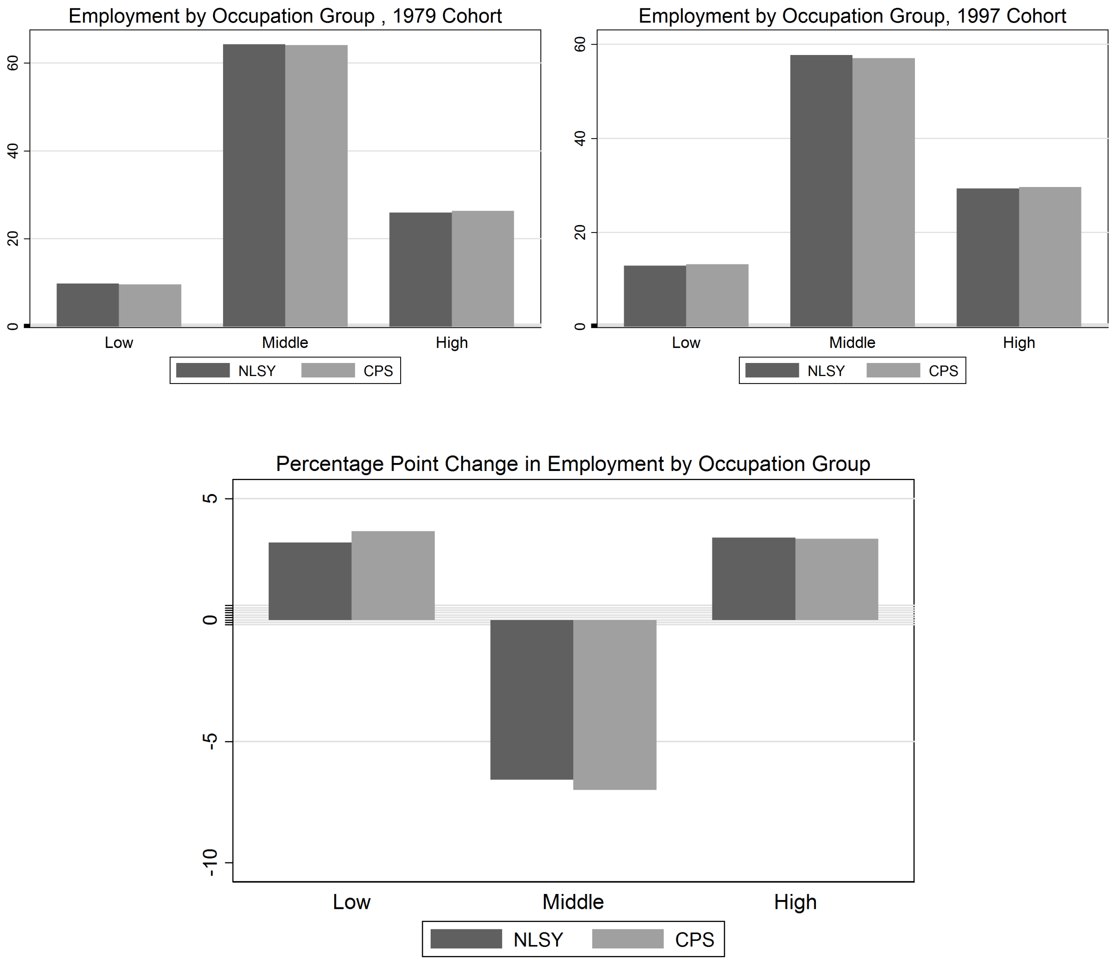
Yamaguchi, S. (2012): "Changes in Returns to Task-Specific Skills and Gender Wage Gap," *SSRN Working Paper*.

Figure 1: The Distribution of Log Wages and its Change



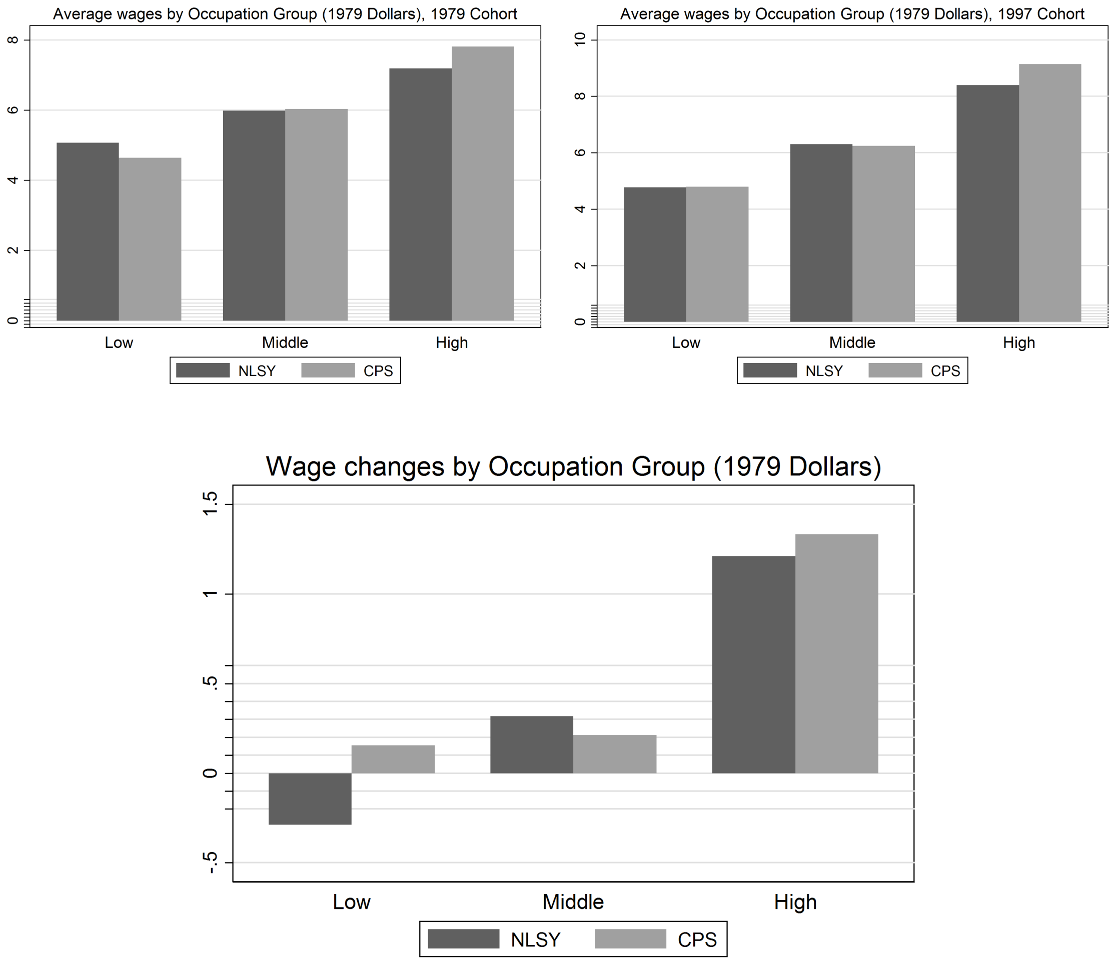
Notes: The subfigure on the top left depicts the empirical cumulative distribution of log real wages for 27 year olds in the NLSY79 cohort and for the comparable years and age group in the CPS. The subfigure on the top right does the same for the NLSY97. The subfigure at the bottom compares the changes in log real wages along the quantiles of the wage distribution over the two cohorts.

Figure 2: Employment Shares by Broad Occupation Group and their Changes



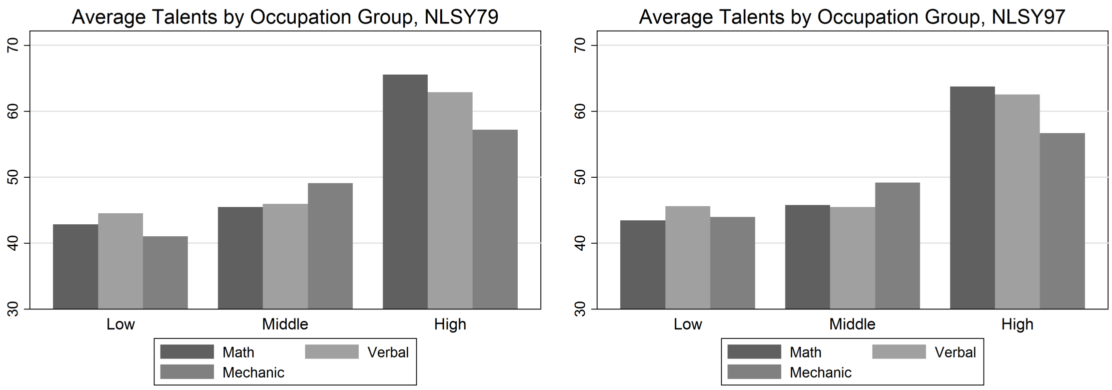
Notes: The subfigure on the top left depicts the employment shares of low-, middle-, and high-skilled occupations for the NLSY79 cohort and the comparable years and age group in the CPS. The subfigure on the top right does the same for the NLSY97. The subfigure at the bottom depicts the percentage point change in employment in the three occupation groups and again the CPS in comparison. The high skill occupation group contains managerial, professional services, and technical occupations. The middle skill occupation group contains sales, office / administrative, production, and operator and laborer occupations. The low skill occupation group contains protective, food, cleaning and personal service occupations.

Figure 3: Real (1979) Wages by Broad Occupation Group and their Changes



Notes: The subfigure on the top left depicts the average real wages of low-, middle-, and high-skilled occupations for the NLSY79 cohort and the comparable years and age group in the CPS. The subfigure on the top right does the same for the NLSY97. The subfigure at the bottom depicts the change in real wages in the three occupation groups and again the CPS in comparison. The high skill occupation group contains managerial, professional services, and technical occupations. The middle skill occupation group contains sales, office / administrative, production, and operator and laborer occupations. The low skill occupation group contains protective, food, cleaning and personal service occupations.

Figure 4: Average Talents in Occupation Groups, NLSY 1979 and 1997



Notes: The figures display the average math, verbal, and mechanical test scores in the three occupation groups for the NLSY79 and the NLSY97.

Figure 5: The Estimation Problem

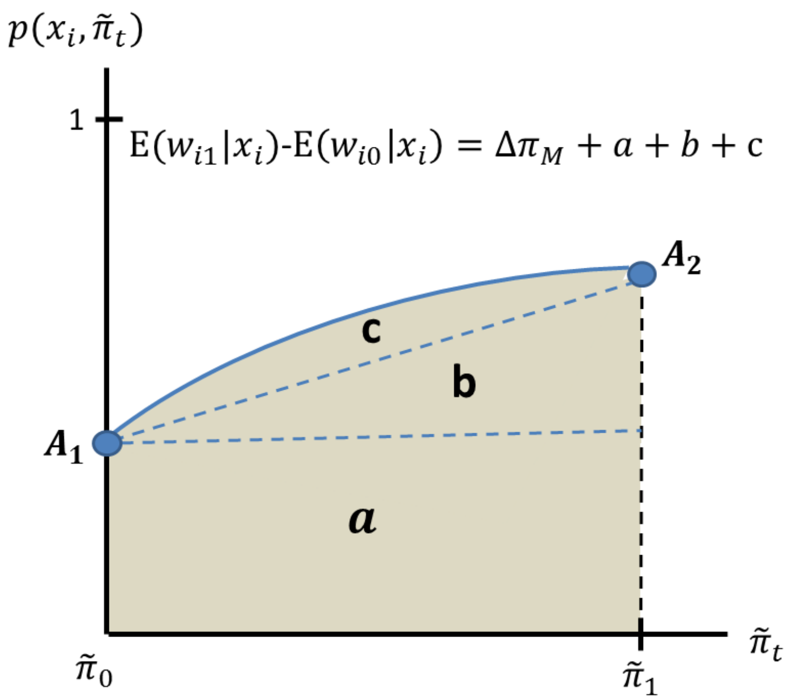
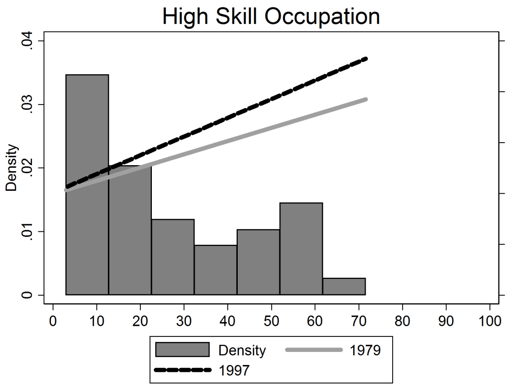
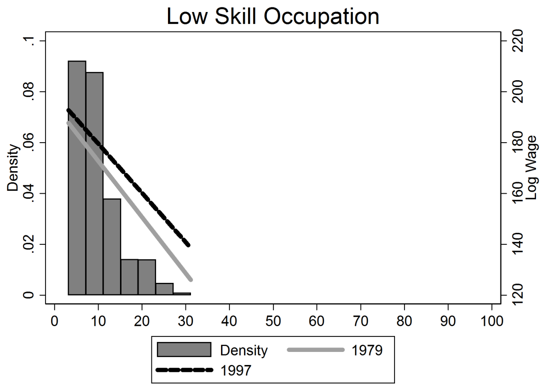


Figure 6: Predicted Relative Skill Returns and their Changes

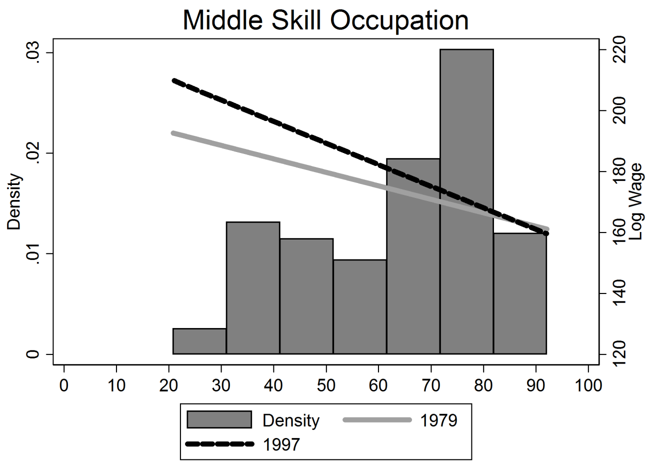
(a) Propensity High Occupation



(b) Propensity Low Occupation



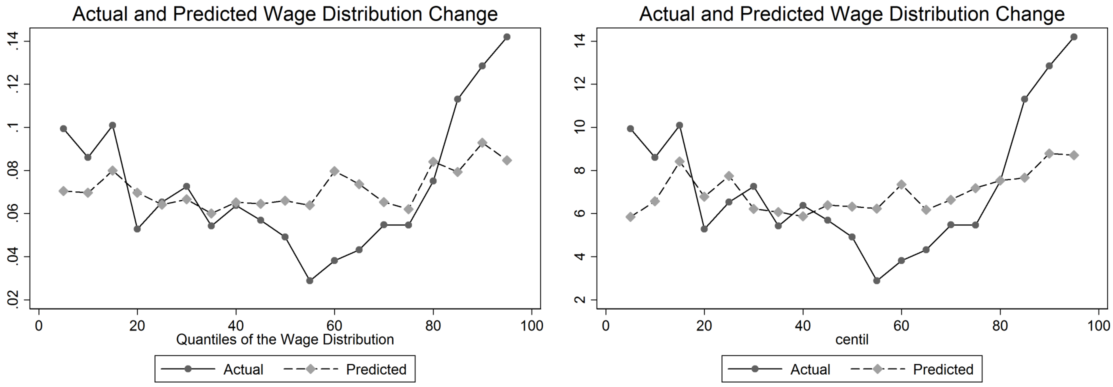
(c) Propensity Middle Occupation



Notes: The figures plot the returns to propensities of entering the respective occupation in the NLSY79 and the NLSY97 together with the empirical density of these propensities in the NLSY79. The returns are estimated in regressions of log wages on a constant and the respective propensity together with an interaction term for the NLSY97.

Figure 7: Actual and Predicted Wage Distribution Change

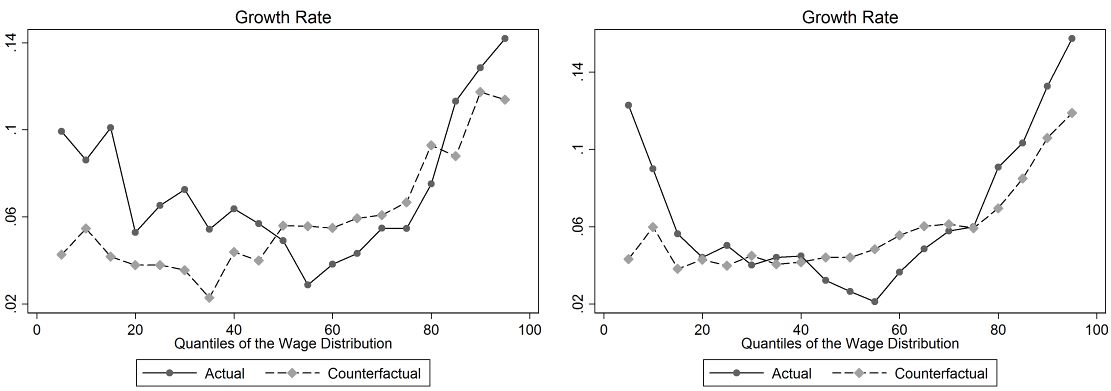
(a) Returns to Propensities (b) Returns to All Talents



Notes: The figures plot the actual and the predicted change in the wage distribution when workers in the NLSY79 are assigned the change in the returns to their observable characteristics between the two cohorts estimated in columns one and four of table 5.

Figure 8: Actual and Counterfactual Wage Distribution Change, NLSY and CPS

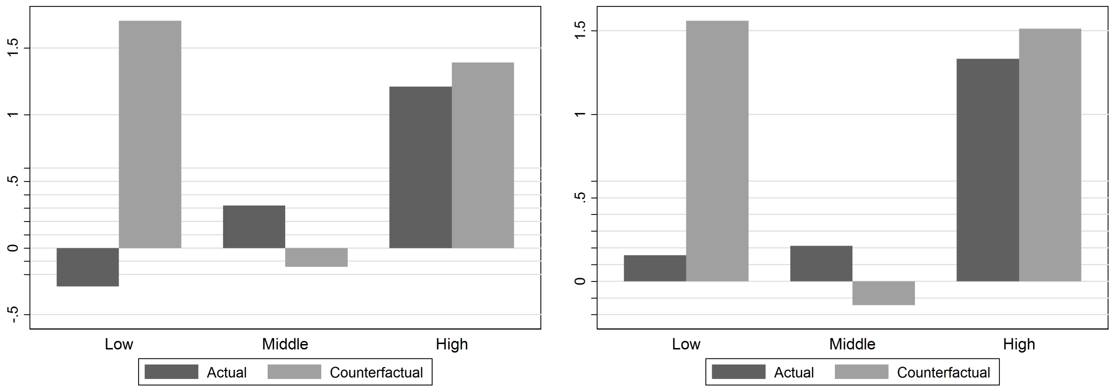
(a) NLSY (b) CPS



Notes: The figure plots the actual and the counterfactual change in the wage distribution when workers in the initial period are assigned the estimated price changes in their occupations from the optimal minimum distance estimator in table 7.

Figure 9: Actual and Counterfactual Occupational Wage Changes, NLSY and CPS

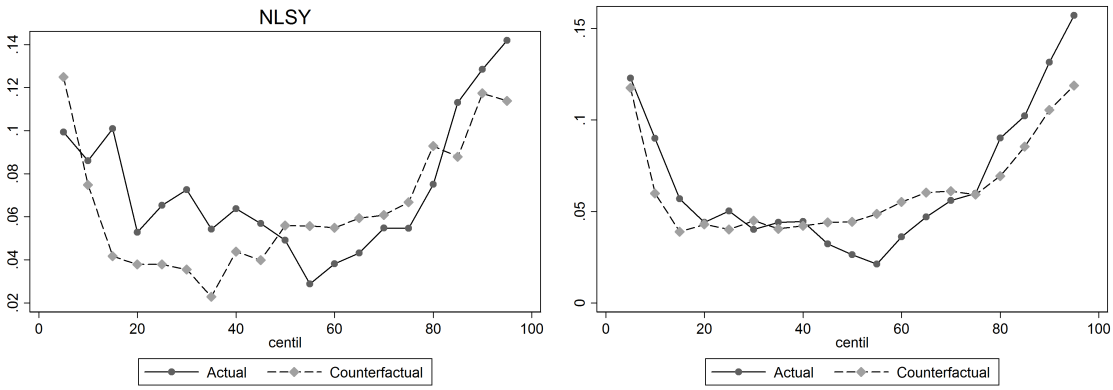
(a) NLSY (b) CPS



Notes: The figures plot the actual and the counterfactual change in occupational wages in the NLSY and CPS when workers in the initial period are assigned the estimated price changes in their occupations from the optimal minimum distance estimator in table 7.

Figure 10: Actual and Counterfactual Wage Distribution Change with Reallocation, NLSY and CPS

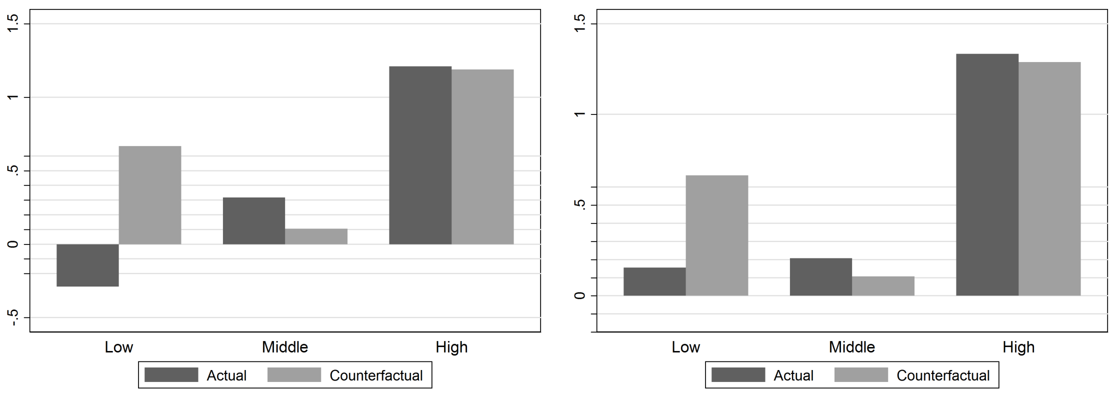
(a) NLSY (b) CPS



Notes: The figures plot the actual and the counterfactual change in occupational wages in the NLSY and CPS when workers in the initial period are assigned the estimated price changes in their occupations plus a reallocation effect: the lowest-earning three percent are assumed to move out of the middle- to the low-skill occupation with a 15 percent relative wage increase and the next low-earning one percent is assumed to move to the high-skill occupation with the same relative wage gain.

Figure 11: Actual and Counterfactual Occupational Wage Change with Reallocation, NLSY and CPS

(a) NLSY (b) CPS



Notes: The figures plot the actual and the counterfactual change in occupational wages in the NLSY and CPS when workers in the initial period are assigned the estimated price changes in their occupations plus a reallocation effect: the lowest-earning three percent are assumed to move out of the middle- to the low-skill occupation with a 15 percent relative wage increase and the next low-earning one percent is assumed to move to the high-skill occupation with the same relative wage gain.

Table 1: From the full NLSY to the analysis sample

	NLSY79 (Birthyears 1956-1964)	**NLSY97 (Birthyears 1980-1984)**
Reason for exclusion		
Total males	6,403	4,599
Excluded oversampled white and older arrivers in US than age 16	4,585	4,599
Birthyear > 1982	4,585	2,754
Type of attrition		
Ought to be present with ASVAB at age 27	4,585	2,754
No ASVAB excluded	4,299	2,081
%	94	76
Not present at age 27 excluded	3,939	1,737
%	86	63
Conditioned on working		
Excluded who report no or farm occupation, self-employed, and those with no wage income	3,054	1,207

Note: The table reports how I get from the full NLSY 1979 and 1997 to my analysis sample and where observations are lost or need to be dropped.

Table 2: Labor Supply with Respect to Average Demographics, Early, and Contemporary Skill Determinants

	NLSY79	**NLSY97**
Nbr of observations	3051	1210
Percentage of observations	71.60	28.40
Demographics		
Age	27.00	27.00
White	0.80	0.72
Black	0.13	0.14
Hispanic	0.06	0.14
Early skill determinants		
AFQT	167.31	167.65
Low AFQT Tercile	0.34	0.33
Middle AFQT Tercile	0.33	0.34
High AFQT Tercile	0.33	0.32
Math Score (NCE)	50.45	50.73
Verbal Score (NCE)	50.26	50.49
Mechanical Score (NCE)	50.41	50.69
Illicit Activities (NCE, Measured 1980)	49.98	50.01
Precocious Sex (NCE, Measured 1983)	49.91	50.24
Mother's Education (Years)	11.86	13.11
Father's Education (Years)	10.83	13.09
Contemporary skill determinants		
High School Dropout (HSD)	0.12	0.07
High School Graduate (HSG)	0.43	0.58
Some College (SC)	0.20	0.06
College Graduate (CG)	0.19	0.24
Advanced Degree (AD)	0.06	0.04
North East	0.22	0.17
North Central	0.29	0.25
South	0.32	0.35
West	0.17	0.21

Note: The table shows average demographics and skill proxies in the NLSY79 and NLSY97 for all individuals weighted by hours worked. NCE indicates variables in the population (including non-workers) are standardized to "normal curve equivalents" with mean 50 and standard deviation 21.06. This is done when absolute values of these variables cannot confidently compared over the two cohorts.

Table 3: Pairwise Correlations between Composite Test Scores

	NLSY79 AFQT	Math	Verbal	NLSY97 AFQT	Math	Verbal
AFQT (NCE)	1			1		
Math Score (NCE)	0.82	1		0.83	1	
Verbal Score (NCE)	0.93	0.71	1	0.92	0.75	1
Mechanical Score (NCE)	0.63	0.53	0.61	0.63	0.54	0.63
Nbr Observations	2936			1210		

Note: The table shows the pairwise correlations between composite test scores after standardizing to normal curve equivalents with mean 50 and standard deviation 21.06.

Table 4: Sorting into Occupation Groups, Multinomial Logit Regressions

	(1) NLSY79	(2) NLSY79	(3) NLSY97	(4) NLSY97
High				
Constant	-4.024***	-1.710***	-3.176***	-1.384***
Black	0.235	0.159	-0.152	-0.106
Hispanic	0.03	-0.031	-0.472*	-0.456*
Math (NCE)	0.047***		0.034***	
Verbal (NCE)	0.023***		0.032***	
Mechanic (NCE)	-0.014***		-0.019***	
Middle Math Tercile		1.144***		0.441*
High Math Tercile		2.315***		1.426***
Middle Verbal Tercile		0.207		0.670**
High Verbal Tercile		0.750***		1.445***
Middle Mechanic Tercile		-0.269		-0.258
High Mechanic Tercile		-0.552***		-0.618**
Illicit Activities (NCE)		-0.009***		-0.003
Precocious Sex (NCE)		-0.004		-0.006
Low				
Constant	-1.689***	-1.608***	-1.339***	-2.053***
Black	0.636***	0.762***	0.473*	0.658**
Hispanic	0.201	0.243	-0.216	-0.114
Math (NCE)	-0.002		-0.009	
Verbal (NCE)	0.018***		0.021**	
Mechanic (NCE)	-0.023***		-0.017**	
Middle Math Tercile		-0.381**		-0.07
High Math Tercile		0.128		-0.395
Middle Verbal Tercile		0.342		0.27
High Verbal Tercile		0.471*		0.790**
Middle Mechanic Tercile		-0.319		-0.281
High Mechanic Tercile		-0.908***		-0.608*
Illicit Activities (NCE)		-0.002		0.013*
Precocious Sex (NCE)		-0.003		-0.003
Pseudo R-Squared	0.132	0.123	0.114	0.112
N	2936	2936	1210	1210

Note: Each columns presents the results from a multinomial logit regression of occupational choice on demographics and talent proxies. The omitted group is the middle occupation. The first column uses only linear test scores in the NLSY79. The second column, which is the specification to estimate occupational propensities in the following, uses terciles of test scores and adds measures of risky behavior. The last two columns repeat these estimations for the NLSY97. In order to save space, standard errors are not reported but statistical significance is indicated: * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Returns to Occupational Propensities over the Two Cohorts

	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage
Constant	181.15***	185.17***	176.66***	183.21***
	(3.10)	(3.11)	(3.76)	(21.61)
Const x NLSY97	-7.90	-10.27	-12.59	-43.37
	(6.74)	(6.57)	(8.16)	(41.62)
Prop High Occup	0.31***	0.03	-0.06	0.13
	(0.07)	(0.08)	(0.08)	(0.57)
Prop H Occ x NLSY97	0.29***	0.25**	0.30**	1.41
	(0.11)	(0.13)	(0.13)	(1.03)
Prop Low Occup	-1.65***	-1.80***	-1.75***	-2.19**
	(0.17)	(0.17)	(0.17)	(0.97)
Prop L Occ x NLSY97	0.70*	0.86**	0.91**	2.26
	(0.39)	(0.38)	(0.38)	(1.92)
College		19.23***		
		(2.92)		
Coll x NLSY97		4.04		
		(5.20)		
Observations	4154	4149	4149	4154
R^2	0.09	0.11	0.12	0.10
Degree dummies	No	No	Yes	No
Talents directly	No	No	No	Yes

Note: The table reports OLS wage regressions of 100 times the deflated log wage on propensities to enter occupation groups (predicted relative occupation-specific skills) and the change in the coefficient between the NLSY79 and the NLSY97. The propensities are from the NLSY79 only and they are from multinomial logit regressions of occupational choice including mathematical, verbal, and mechanical talent terciles, illicit activities, precocious sex and dummies for respondents' race. The specifications in columns two to four add dummies for college degree, detailed education (HS drop out, HS graduate, Some college, College and above), and the talents that were used in the estimation of the propensities directly. "x NLSY97" stands for the interaction between the variable and an NLSY97 dummy, i.e. the change in the coefficient between the NLSY79 and the NLSY97. Standard errors are from bootstrapping the first (estimating the propensities) and second stage regressions together 500 times and they are reported below the coefficients. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Talent Allocation and Returns Changes

	High Skill Occup	**Low Skill Occup**	**Log Wage x NLSY97**
Constant	18.71***	11.87***	14.37**
	(6.21)	(5.51)	(2.08)
Black	-0.762	9.292***	0.647
	(-0.43)	(4.77)	(0.14)
Hispanic	-2.632	1.708	-1.586
	(-1.34)	(1.12)	(-0.36)
Middle Math Tercile	10.83***	-4.466***	-2.615
	(6.05)	(-2.85)	(-0.56)
High Math Tercile	34.90***	-5.997***	10.44*
	(13.48)	(-3.10)	(1.68)
Middle Mechanic Tercile	-2.505	-2.332	-5.767
	(-1.19)	(-1.52)	(-1.20)
High Mechanic Tercile	-7.043***	-4.827***	-1.740
	(-2.81)	(-2.97)	(-0.30)
Middle Verbal Tercile	3.505*	2.429	-0.282
	(1.84)	(1.48)	(-0.06)
High Verbal Tercile	15.34***	2.805	-4.535
	(5.67)	(1.45)	(-0.65)
Illicit Activities (NCE)	-0.129***	0.0388	-0.183*
	(-3.33)	(1.36)	(-1.89)
Precocious Sex (NCE)	-0.0612	-0.0120	0.0527
	(-1.64)	(-0.41)	(0.62)
R-squared	0.182	0.0281	0.0933
N	4146	4146	4146

Note: The first two columns present the coefficients from OLS allocation regressions of working in the low and high skill occupation with pooled NLSY79 and NLSY97 data. The third colum presents the change in the parameters between the two cohorts in an OLS wage regression. Coefficients represent 100 times the average partial increase in the probability of entering the occupation group and the log wage, respectively, for an additional unit of the regressor. "x NLSY97" stands for the interaction between the variable and an NLSY97 dummy, i.e. the change in the coefficient between the NLSY79 and the NLSY97. T-statistics below the coefficients. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Implied Wage Rate Changes and Cross-Equation Restriction Test

	Estim. $\triangle(\pi_H - \pi_M)$ in % (s.e.)	**Estim. $\triangle(\pi_L - \pi_M)$ in % (s.e.)**	**Implied $\triangle\pi_M$ in %**	**Test Statistic (p-value in %)**
OMD / Full GLS	20.1	31.4	-2.4	13.1
	(9.7)	(35.2)		(10.7)
EWMD / OLS	19.4	-4.4	1.7	13.2
	(10.8)	(32.0)		(10.5)
DWMD / WLS	22.0	-7.5	1.3	11.2
	(9.7)	(35.1)		(19.1)

Note: The table presents estimated relative wage rate changes in the high and the low skill occupation compared to the middle skill occupation, a point estimate for the absolute wage rate change in the middle, and the cross-equation restriction test of the polarization hypothesis. The characteristics used in the underlying allocation and wage regressions are my prefered specification, i.e. mathematical, verbal, and practical talent terciles, illicit activities, precocious sex, and dummies for respondents' race. There are 8 degrees of freedom for the test (10 coefficients minus 2 parameters estimated on them). Implied prices and the test statistics are reported for the Optimal Minimum Distance (Full feasable GLS) estimation and as alternatives for the Equally Weighted Minimum Distance (OLS regression of change in wage regression coefficients on allocation regression coefficients), and Diagonally Weighted Minimum Distance (WLS).

Appendices

A Data

A.1 Detailed Sample Construction

I use data from the National Longitudinal Survey of Youth (NLSY) cohort of 1979 and 1997. The individuals in these surveys are born between 1956 and 1964 and between 1980 and 1984, respectively. As is necessary for this paper, the NLSY studies provide detailed information about individuals' background, education, and labor market outcomes. Moreover, the two cohorts are specifically designed to to be comparable to one another.

Consistent with many papers on the NLSY and in the literature on polarization and wage inequality, I restrict my attention to males (e.g. Firpo, Fortin, and Lemieux 2011, Cortes 2012). There are several reasons for doing this. Firstly, polarization seems to have had the most dire effect on males (Acemoglu and Autor 2010). Secondly, female hours worked and thus the type of selection of females into the labor market (see Mulligan and Rubinstein 2008) changed substantially over the two NLSYs. In addition, females made strides in educational attainment, their wages rose across the whole distribution, and attitudes towards them and discrimination against them in the labor market seem to have changed drastically. Thus, there are diverse changes in (the structure of) female labor supply and demand that are likely to work aside from the forces of polarization. Restricting the analysis to males provides a cleaner comparison of workers across the two decades between the NLSY79 and the NLSY97.

I evaluate individuals' labor market outcomes at age 27. This is because, on the one hand, at younger ages the polarization facts that the paper sets out to analyze are not very pronounced in the NLSY as well as CPS data, which I use for comparison. On the other hand, at older ages than 27, I would loose too many observations from the NLSY97 as, at the time of writing, data is only available until the survey year of 2009. With the age 27 restriction, I already have to drop about two fifth of the NLSY97 sample (birth years 1983 and 1984 are dropped).

Table 1 summarizes how the sample restrictions, attrition, and labor market participation for males reduce my sample size from 6,403 to 3,054 and from 4,599 to 1,207 males in the NLSY79 and the NLSY97, respectively. I restrict the sample to individuals who participated in the Armed Services Vocational Aptitude Battery of tests (ASVAB) in the first survey year. This restriction is necessary because ASVAB will provide me with measures of different dimensions of talent for each individual that are comparable over the two cohorts. Moreover, I argue that the subtests from ASVAB are proxies of individuals' fundamental talents that do not react as elastically to changes in market returns as late skill determinants, such as education, which have been used in existing studies.

The participation in ASVAB is substantially lower in the NLSY97 than the NLSY79 where almost everyone participated. Moreover, sample attrition at age 27 is higher in the NLSY97 than the NLSY79 and overall only 63 percent of the NLSY79 participated in ASVAB and are also present at age 27. This problem is well known for the NLSY (e.g. Altonji, Bharadwaj, and Lange 2008, Aughinbaugh and Gardecki 2007). More generally, attrition rates in several panel surveys in the United States increased substantially during the 1990s (see also Fitzgerald, Gottschalk, and Moffitt 1998, MaCurdy, Mroz, and Gritz 1998). The attrition and non-test-participation rates in my data closely line up with those reported in the study by Altonji, Bharadwaj, and Lange (henceforth ABL). The only difference is that ABL consider outcomes at the younger age of 22 and thus have slightly lower attrition rates.

In their paper, ABL note that the higher attrition rate in the NLSY97 may be partly due to NLSY97 respondents being first interviewed at ages 12-16 versus ages 14-21 for the NLSY79 and thus had more time to attrit. ABL further extensively examine the potential non-randomness of attrition and non-test-participation and its likely impact in biasing important labor market outcomes. Aughinbaugh and Gardecki (2007) do a similar exercise but focus on social and educational outcomes. Both studies find evidence that attrition is not random with respect to youths' outcomes and their backgrounds. However, Aughinbaugh and Gardecki (2007) conclude that attrition from the NLSY97 does not appear to affect inference when estimating the three outcomes at age 20 that they are considering and ABL decide that the differences between non-attriters and the whole sample are not forbidding.

Moreover, ABL carefully select the samples of NLSY79 and NLSY97 to make them comparable to one another and compute weights that adjust for attrition and non-test-participation on observable characteristics. I closely follow their procedures for constructing my own sample. Thus, for even more information on the sample construction and statistics on the effects of attrition, please refer to ABL in addition to the description provided here.[39] First, I follow ABL in excluding from the NLSY79 immigrants who arrived in the United States after age 16. This is done because the scope of the NLSY97 (age 12-16) also doesn't include older than age 16 arrivals. Second, I exclude the economically disadvantaged whites and military supplemental samples from the NLSY79 because they were discontinued early on in the survey and thus don't provide labor market outcomes at age 27 (or for ABL's purposes). Table 1 reports that 1,818 observations are dropped by making these restrictions to the sample. For each individual I retain the observation that is closest to 27 years and 6 months of age and then measure labor market and final educational outcomes from this observation.

ABL use a probit model to adjust the NLSY79 and NLSY97 base year sample weights to account for attrition and non-test-participation according to several observable characteristics, such as parental education, parental presence at age 14, indicators by birth-year, urban and SMSA residence status, indicator variables for race and gender, and an interviewer coded variable describing the attitude of the respondent during the interview. I also employ a probit model to adjust weights for attrition and non-test-participation and use the same specification and variables as ABL apart from leaving out parental presence at age 14. Alternatively, a fully stratified set of indicators for birthyear, year, sex, and race, as employed by the Bureau of Labor Statistics for weighting, yields very similar results.[40] As ABL do in their paper, I proceed from this point with the assumption that, after attrition weighting, my two NLSY samples are representative of the population of young Americans that they are supposed to cover. These samples have the size of 3,939 and 1,737 individuals in the NLSY79 and the NLSY97, respectively.

I follow Lemieux (2006), who uses CPS May Outgoing Rotation Group data, in

[39] I am extremely grateful to Prashant Bharadwaj for providing me with their data and do-files.

[40] I thank Steve McClaskie and Jay Zagorsky for providing me with the official attrition-adjusted sample weighting program for the NLSY.

how I compute wages and in defining the sample of working individuals (henceforth labor supply). First, I use hourly wages reported for the current main job instead of imputing hourly wages from last year's income and total hours worked. Lemieux (2006) convincingly argues that the current main job measure is substantially more accurate because it better measures the wages of workers paid by the hour. Moreover, the reporting of weeks and hours per year worked in the NLSY seems somewhat inconsistent over the two cohorts. I normalize all wages to 1979 real values by adjusting with the PCE deflator provided by the St.Louis Federal Reserve Bank.[41] While Lemieux (2006) removes outliers with 1979 real wages below $1 and above $100, I remove the high wages from $40 onward because my NLSY wage data is very inaccurate for values above this threshold.

Finally, in order to condition on the sample of working individuals, I keep all individuals who report not to be self-employed, and who are employed in a non-farm, non-fishing and non-forestry occupation according to the Census 1990 three-digit occupation classification. This leaves me with an analysis sample of 3,054 and 1,207 males in the NLSY79 and NLSY97, respectively (compare table 1 again). I weight all of those individuals by the number of hours that they work per week on top of the sample weights that are adjusted for test-participation and attrition. Lemieux (2006) argues that weighting by weekly hours can be viewed as a reasonable compromise between concentrating on full-time workers only and looking at all workers including part-time workers. An additional advantage from this is that I am not losing any more observations from a full-time work restriction.

B Theory

B.1 Generalization of Results to the Three-Occupation Case

In the following I derive predictions (7) and (8) for the three-occupation case. For ease of exposition, wages in occupations (3) are reproduced here:

$$w_{Kit} = \pi_{Kt} + \beta_{K0} + \beta_{K1}x_{1it} + ... + \beta_{KJ}x_{Jit} + u_{Kit}.$$

[41]Source: "Personal Consumption Expenditures: Chain-type Price Index (PCECTPI)", accessed 2012-8-14, http://research.stlouisfed.org/fred2/series/PCECTPI

Note from equation (2) and the wages in occupations that:

$$w_{it} = \begin{cases} w_{Hit} = \pi_{Ht} + \beta_{H0} + \beta_{H1}x_{1it} + ... + \beta_{HJ}x_{Jit} + u_{Hit} & \text{if } H_{it} = 1 \\ w_{Mit} = \pi_{Mt} + \beta_{M0} + \beta_{M1}x_{1it} + ... + \beta_{MJ}x_{Jit} + u_{Mit} & \text{if } M_{it} = 1 \\ w_{Lit} = \pi_{Lt} + \beta_{L0} + \beta_{L1}x_{1it} + ... + \beta_{LJ}x_{Jit} + u_{Lit} & \text{if } L_{it} = 1 \end{cases}$$

When occupational wage rates change, by the envelope theorem, the marginal change in worker i's wage becomes

$$dw_{it} = \begin{cases} d\pi_H & \text{if } H_{it} = 1 \\ d\pi_M & \text{if } M_{it} = 1 \\ d\pi_L & \text{if } L_{it} = 1. \end{cases}$$

Thanks to its linearity, the change in the expectation can be written as

$$E(dw_{it}|x_{it}, \pi_t) = p_H(x_{it}, \pi_t)d\pi_H + p_M(x_{it}, \pi_t)d\pi_M + p_L(x_{it}, \pi_t)d\pi_L,$$

where $p_K(x_{it}, \pi_t)$ is the propensity for an individual of talent vector x_{it} to enter occupation K under prices π_t. Exploiting that the three probabilities sum to one gives prediction (13):

$$dE(w_{it}|x_{it}, \pi_t) = d\pi_{Mt} + p_H(x_{it}, \tilde{\pi}_{HMt}, \tilde{\pi}_{LMt})d\tilde{\pi}_{HMt} + p_L(x_{it}, \tilde{\pi}_{HMt}, \tilde{\pi}_{LMt})d\tilde{\pi}_{LMt},$$

where $\tilde{\pi}_{KMt} \equiv \pi_{Kt} - \pi_{Mt}$ for $K\epsilon\{H, L\}$,

$$\begin{aligned} p_H(x_{it}, \tilde{\pi}_{HMt}, \tilde{\pi}_{LMt}) = Pr[u_{Hi} - u_{Mi} > -(\pi_{Ht} - \pi_{Mt} + (\beta_H - \beta_M)'x_{it}), \\ u_{Hi} - u_{Li} > -(\pi_{Ht} - \pi_{Lt} + (\beta_H - \beta_L)'x_{it})], \end{aligned}$$

and similarly for $p_L(x_{it}, \tilde{\pi}_{HMt}, \tilde{\pi}_{LMt})$.

For convenience, omit the dependence on x_{it} from now on. Holding constant $\tilde{\pi}_{HMt}$ and $\tilde{\pi}_{LMt}$ at $t = 0$ and integrating equation (13) with respect to π_{Mt} we get

$$E(w_i|\pi_{M1}, \tilde{\pi}_{HM0}, \tilde{\pi}_{LM0}) - E(w_i|\pi_{M0}, \tilde{\pi}_{HM0}, \tilde{\pi}_{LM0}) = \triangle\pi_M.$$

Similarly,

$$E(w_i|\pi_{M1}, \tilde{\pi}_{HM1}, \tilde{\pi}_{LM0}) - E(w_i|\pi_{M1}, \tilde{\pi}_{HM0}, \tilde{\pi}_{LM0}) = \int_{\tilde{\pi}_{HM0}}^{\tilde{\pi}_{HM1}} p_H(\tilde{\pi}_{HMt}, \tilde{\pi}_{LM0}) d\tilde{\pi}_{HMt}$$
$$E(w_i|\pi_{M1}, \tilde{\pi}_{HM1}, \tilde{\pi}_{LM1}) - E(w_i|\pi_{M1}, \tilde{\pi}_{HM1}, \tilde{\pi}_{LM0}) = \int_{\tilde{\pi}_{LM0}}^{\tilde{\pi}_{LM1}} p_L(\tilde{\pi}_{HM1}, \tilde{\pi}_{LMt}) d\tilde{\pi}_{LMt}.$$

Summing these three expressions gives equation (14):

$$E(w_i|\pi_1) - E(w_i|\pi_0) = \triangle\pi_M + \int_{\tilde{\pi}_{HM0}}^{\tilde{\pi}_{HM1}} p_H(\tilde{\pi}_{HMt}, \tilde{\pi}_{LM0}) d\tilde{\pi}_{HMt} + \int_{\tilde{\pi}_{LM0}}^{\tilde{\pi}_{LM1}} p_L(\tilde{\pi}_{HM1}, \tilde{\pi}_{LMt}) d\tilde{\pi}_{LMt}$$

Finally, since we do not know the choice probabilities on the adjustment path, these have to be approximated analogously to equation (9) and figure 5

$$p_H(\tilde{\pi}_{HMt}, \tilde{\pi}_{LM0}) \approx p_H(\tilde{\pi}_{HM0}, \tilde{\pi}_{LM0}) + \frac{p_H(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM1}) - p_H(\tilde{\pi}_{HM0}, \tilde{\pi}_{LM0})}{\tilde{\pi}_{HM1} - \tilde{\pi}_{HM0}} (\tilde{\pi}_{HMt} - \tilde{\pi}_{HM0})$$
$$p_L(\tilde{\pi}_{HM1}, \tilde{\pi}_{LMt}) \approx p_L(\tilde{\pi}_{HM0}, \tilde{\pi}_{LM0}) + \frac{p_L(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM1}) - p_L(\tilde{\pi}_{HM0}, \tilde{\pi}_{LM0})}{\tilde{\pi}_{LM1} - \tilde{\pi}_{LM0}} (\tilde{\pi}_{LMt} - \tilde{\pi}_{LM0}),$$

which gives equation (16). Note that one might prefer using $p_H(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM0})$ instead of $p_H(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM1})$ in the first approximation and $p_L(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM0})$ instead of $p_L(\tilde{\pi}_{HM0}, \tilde{\pi}_{LM0})$ in the second, which are not observable in the data. Yet, $p_H(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM0}) > p_H(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM1})$ while $p_L(\tilde{\pi}_{HM1}, \tilde{\pi}_{LM0}) < p_L(\tilde{\pi}_{HM0}, \tilde{\pi}_{LM0})$, so this additional approximation error should not be too large.

C Econometrics

C.1 Details of the Minimum Distance Estimation and Test

The methods applied in the following can be found in the statistical appendix of Abowd and Card (1989) or chapter 6.7 of Cameron and Trivedi (2005). I explain them step by step.

First, I run seemingly unrelated wage and allocation regressions on the individual level in both points in time to obtain estimates $\hat{\delta}_{Ht}$, $\hat{\delta}_{Lt}$, and $\hat{\gamma}_t$ as well as an estimate of their joint covariance matrix. Second, I combine the wage rates $\triangle\tilde{\pi}_H \equiv \triangle(\pi_H - \pi_M)$, $\triangle\tilde{\pi}_L \equiv \triangle(\pi_L - \pi_M)$, and $\triangle\tilde{\pi} = [\triangle\tilde{\pi}_H, \triangle\tilde{\pi}_L]$. I also define the $J \times 1$ vectors $\hat{\triangle\gamma}$,

$\hat{\hat{\delta}}_K \equiv \frac{\hat{\delta}_{K0}+\hat{\delta}_{K1}}{2}$ for $K \epsilon \{H, L\}$, and $m(\triangle\tilde{\pi}) = \hat{\triangle\gamma} - \hat{\hat{\delta}}_H \triangle\tilde{\pi}_H - \hat{\hat{\delta}}_L \triangle\tilde{\pi}_L$.

The minimum distance estimator minimizes the quadratic form

$$Q(\triangle\tilde{\pi}) = m(\triangle\tilde{\pi})' W m(\triangle\tilde{\pi}),$$

with W being a $J \times J$ weighting matrix. Under some regularity conditions we can apply a central limit theorem to the OLS estimates $\hat{\delta}_{Ht}$, $\hat{\delta}_{Lt}$, and $\hat{\gamma}_t$ as well as to $m(\triangle\tilde{\pi})$:

$$\sqrt{N} m(\triangle\tilde{\pi}) \overset{a}{\sim} \mathcal{N}\left(Em(\triangle\tilde{\pi}), NVar(m(\triangle\tilde{\pi}))\right)$$

Under the polarization hypothesis, $Em(\triangle\tilde{\pi}) = 0$ and the variance of $m(\triangle\tilde{\pi})$ can be derived up to the parameter vector $\triangle\tilde{\pi}$ from the covariance matrix of the reduced form estimates:

$$\begin{aligned} \hat{Var}(m(\triangle\tilde{\pi})) = {} & \hat{Var}(\hat{\triangle\gamma}) + \triangle\tilde{\pi}_H^2 \hat{Var}(\hat{\hat{\delta}}_H) + \triangle\tilde{\pi}_L^2 \hat{Var}(\hat{\hat{\delta}}_L) + 2\triangle\tilde{\pi}_H \triangle\tilde{\pi}_L \hat{Cov}(\hat{\hat{\delta}}_H, \hat{\hat{\delta}}_L) - \\ & - 2\triangle\tilde{\pi}_H \hat{Cov}(\hat{\triangle\gamma}, \hat{\hat{\delta}}_H) - 2\triangle\tilde{\pi}_L \hat{Cov}(\hat{\triangle\gamma}, \hat{\hat{\delta}}_L) \end{aligned}$$

Since $\triangle\tilde{\pi}$ is unknown in $\hat{Var}(m(\triangle\tilde{\pi}))$, I run two step feasible GLS with the first stage being OLS using $W = I$ and plugging the resulting $\hat{\triangle\tilde{\pi}}_{OLS}$ into the weighting matrix $W = \hat{Var}(m(\triangle\tilde{\pi}))$ for the second step. The minimized value of the objective function can be shown to be chi-squared distributed asymptotically

$$m(\hat{\triangle\tilde{\pi}}_{FGLS})' [\hat{Var}(m(\hat{\triangle\tilde{\pi}}_{FGLS}))]^{-1} m(\hat{\triangle\tilde{\pi}}_{FGLS}) \overset{a}{\sim} \chi^2(J-2),$$

which provides the specification test.

Since there are concerns about small sample bias of $\hat{\triangle\tilde{\pi}}_{FGLS}$ (in particular Altonji and Segal 1996, Pischke 1995), results for $\hat{\triangle\tilde{\pi}}_{OLS}$ and $\hat{\triangle\tilde{\pi}}_{WLS}$ are also reported. In this case, the test statistic for the model test has to be adjusted (see Abowd and Card 1989).

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1214	Nattavudh Powdthavee Warn N. Lekfuangfu Mark Wooden	The Marginal Income Effect of Education on Happiness: Estimating the Direct and Indirect Effects of Compulsory Schooling on Well-Being in Australia
1213	Richard Layard	Mental Health: The New Frontier for Labour Economics
1212	Francesco Caselli Massimo Morelli Dominic Rohner	The Geography of Inter-State Resource Wars
1211	Stephen Hansen Michael McMahon	Estimating Bayesian Decision Problems with Heterogeneous Priors
1210	Christopher A. Pissarides	Unemployment in the Great Recession
1209	Kevin D. Sheedy	Debt and Incomplete Financial Markets: A Case for Nominal GDP Targeting
1208	Jordi Blanes i Vidal Marc Möller	Decision-Making and Implementation in Teams
1207	Michael J. Boehm	Concentration versus Re-Matching? Evidence About the Locational Effects of Commuting Costs
1206	Antonella Nocco Gianmarco I. P. Ottaviano Matteo Salto	Monopolistic Competition and Optimum Product Selection: Why and How Heterogeneity Matters
1205	Alberto Galasso Mark Schankerman	Patents and Cumulative Innovation: Causal Evidence from the Courts
1204	L Rachel Ngai Barbara Petrongolo	Gender Gaps and the Rise of the Service Economy
1203	Luis Garicano Luis Rayo	Relational Knowledge Transfers
1202	Abel Brodeur	Smoking, Income and Subjective Well-Being: Evidence from Smoking Bans

1201	Peter Boone Ila Fazzio Kameshwari Jandhyala Chitra Jayanty Gangadhar Jayanty Simon Johnson Vimala Ramachandrin Filipa Silva Zhaoguo Zhan	The Surprisingly Dire Situation of Children's Education in Rural West Africa: Results from the CREO Study in Guinea-Bissau
1200	Marc J. Melitz Stephen J. Redding	Firm Heterogeneity and Aggregate Welfare
1199	Giuseppe Berlingieri	Outsourcing and the Rise in Services
1198	Sushil Wadhwani	The Great Stagnation: What Can Policymakers Do?
1197	Antoine Dechezleprêtre	Fast-Tracking 'Green' Patent Applications: An Empirical Analysis
1196	Abel Brodeur Sarah Flèche	Where the Streets Have a Name: Income Comparisons in the US
1195	Nicholas Bloom Max Floetotto Nir Jaimovich Itay Saporta-Eksten Stephen Terry	Really Uncertain Business Cycles
1194	Nicholas Bloom James Liang John Roberts Zhichun Jenny Ying	Does Working from Home Work? Evidence from a Chinese Experiment
1193	Dietmar Harhoff Elisabeth Mueller John Van Reenen	What are the Channels for Technology Sourcing? Panel Data Evidence from German Companies
1192	Alex Bryson John Forth Minghai Zhou	CEO Incentive Contracts in China: Why Does City Location Matter?
1191	Marco Bertoni Giorgio Brunello Lorenzo Rocco	When the Cat is Near, the Mice Won't Play: The Effect of External Examiners in Italian Schools
1190	Paul Dolan Grace Lordan	Moving Up and Sliding Down: An Empirical Assessment of the Effect of Social Mobility on Subjective Wellbeing

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7404 0612
Email info@cep.lse.ac.uk Web site http://cep.lse.ac.uk