Job Amenities and Earnings Inequality

Alex Bell*

May 7, 2025

Abstract

A recent *Handbook of Labor Economics* chapter concludes that "[o]mitted measures of job and firm productivity are a primary source of bias in compensating differential estimates" (Mas, 2025). This paper shows that such bias can be resolved using a simple amenity pricing framework that leverages an imprecise proxy—or "anti-instrument"—for workers' offer sets. The method identifies amenity prices while holding offer sets fixed and recovers intuitively signed estimates in widely used cross-sectional datasets. The approach relies only on a single vertical index of job quality and a conditional independence assumption standard in prior work; it requires no assumptions about functional forms, distributions, or correlations between ability and preferences. Relative to prior observational methods, the estimates suggest a substantial role for amenity substitution in explaining the gender pay gap—accounting for roughly two-thirds of the observed gap—and little role for amenities in explaining inequalities by race or parent background. These findings offer new insight into the structure of labor market inequality and the large scope for equity-focused policy.

Keywords: Job Amenities, Gender Pay Gap, Labor Market Inequality, Compensating Differentials

Acknowledgments: This work has benefited from insightful advice from Isaiah Andrews, Steve Billings, Sophie Calder-Wang, Gary Chamberlain, Raj Chetty, Kirsten Clinton, David Cutler, Tanaya Devi, Natalia Emanuel, John Friedman, Roland Fryer, Edward Glaeser, Claudia Goldin, Nathaniel Hendren, Chinhui Juhn, Patrick Kline, Sandy Jencks, Corinne Low, Lawrence Katz, Philip Marx, Elizabeth Mishkin, Corina Mommaerts, Christian Moser, Alex Olssen, Amanda Pallais, Ariel Pakes, Jonathan Roth, Yulya Truskinovsky, and Shusheng Zhong. The paper has also benefited from feedback of seminar participants at Georgia Tech, Keio University, Osaka University, University of Bath, University of Essex, University of Nottingham, University of South Alabama, and Upjohn Institute for Employment Research. Finally, I am deeply grateful for the insights and encouragement of Sophie Calder-Wang, Stephen Billings, and Shusheng Zhong, who have become my coauthors on a companion paper advancing similar ideas in the housing context. This research has been supported by Harvard's Multidisciplinary Program in Inequality and Social Policy, the EO Foundation, and the Washington Center for Equitable Growth. The views expressed herein are my own and do not necessarily reflect the views of these organizations.

*Department of Economics, Georgia State University. Email: ambell@gsu.edu

1 Introduction

Workers differ both in terms of total compensation and how they split their compensation into wages and amenities. Determinants of total compensation may include productivity, discrimination by employers, search frictions, and rents. But income differences also arise when workers trade off between wages and other desirable attributes. This paper puts forward a new methodology for estimating the trade-offs that workers face between income and observable non-monetary characteristics of their jobs. These price estimates allow me to discern whether workers lie on the same or different frontiers of total compensation. I decompose earned income gaps by gender, race, and parent background into a component that is due to sitting on different wage-amenity frontiers (differences in total compensation) versus differences along the wage-amenity frontiers (differences in valuations).

The idea that job amenities command a wage premium, or compensating wage differential, dates back at least to Smith (1776). Two centuries, later Brown (1980) best summarized the existing empirical literature as holding "some clear support for the theory but an uncomfortable number of exceptions." More recently, quasi-experimental evidence has revived interest in the theory. A persistent barrier to studying the full impact of compensating differentials on income inequality is the inherent difficulty of estimating prices for a wide set of job amenities in a sample representative of the labor market.

The problem stems from the researcher's inability to isolate workers who had similar offer sets to each other but chose to split their compensation differently into wage and amenities. A classic illustration is that of physical safety on the job. Supposing workers tend to prefer jobs with higher levels of physical safety, one would expect workers in safer jobs to be paid less. Figure 1 shows the opposite is true in the NLSY. In general, safer jobs pay more, not less. The goal, however, is to look at how income and safety vary among workers who share the same level of productivity, and therefore could have obtained similar types of jobs. Panel B shows that although controlling for observable productivity-relevant characteristics weakens the relationship, in this case it does not even change it to the expected sign. Such a result might be expected if the productivity channels that determine a worker's compensation are not fully observed. In the absence of a full set of controls, income and safety are both determined by differences in (unobserved) productivity and preferences. The regression of wage on safety conflates compensating differentials with the returns to skill, much in the same way that a regression of quantities on prices conflates supply and demand. Controlling for a few observed ability measures has done as much to solve this simultaneity problem as controlling for a few shifters of supply would in a regression of price on quantity.[1]

The core contribution of this paper is to demonstrate the data used to construct Figure 1 is more than sufficient to remove the simultaneity bias previously seen as inherent to observational analysis of compensating differentials. Similar to instrumental variables, the anti-instrument approach can be viewed as treating an observed measure of ability as a shifter of the latent offer set.[2] It differs from IV estimation in that the variable that needs to be shifted – the offer set – is not observed.[3]

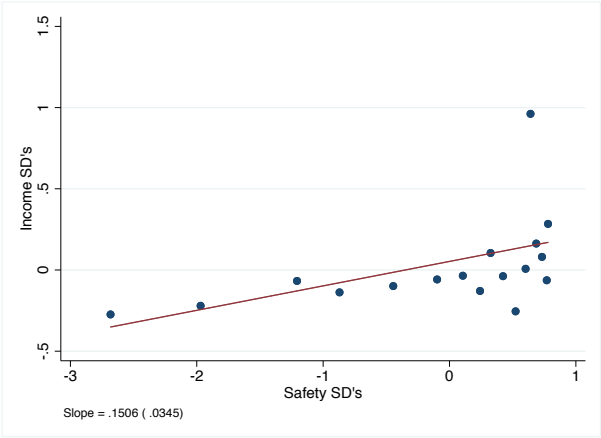
[1] Another analogy can be drawn to discrete choice analysis, in which choices are regressed on product characteristics (typically using logistic regression). That framework assumes that binary choices are informative of which options are better. My framework is similar, but uses an observed measure of worker productivity rather than choice. The typical estimating equation for compensating differentials in representative datasets is analogous to a regression of one product characteristic on other characteristics controlling for which product was chosen, which is generally not consistent as described by McFadden (1973).

[2] Other factors determining the offer set may include unobserved ability, but also factors typically viewed as "stochastic" such as search frictions, connections, or discrimination.

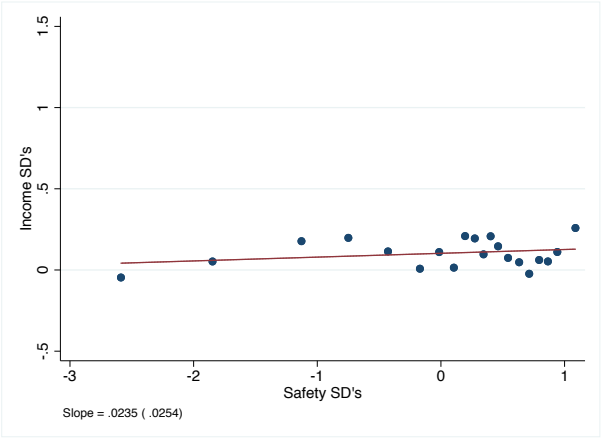
[3] My approach shares this feature in common with Olley and Pakes (1996), who describe how observed investment can identify the unobserved productivity term in production functions, as well as Combes et al. (2021) in the context

Figure 1: The Classic Problem

(a) Income by Safety



(b) With Ability Controls



Notes: Panel A is a binned scatterplot of income on ten bins of safety. The red line is the best-fit line on the unbinned data. Its slope and standard error are noted on the bottom left of the graph. Panel B includes controls that attempt to proxy for worker ability: years of education, age, and AFQT. Data are NLSY respondents in 2012 and standard errors are clustered at the occupation level; additional information on data is contained in Section 5.1.

Rather than "proxy," some researchers also refer to this as the "anti-instrument" approach (Bell et al., 2025a), where the prefix "anti-" denotes opposite,[4] referencing the fact that the opposite variable (i.e., the residual) is being instrumented from what we would traditionally think of as IV. Throughout the paper, I use the term "proxy" and "anti-instrument" interchangeably to describe the same method. This approach also shares some similarity with the selection-on-observables literature (Altonji et al., 2005; Oster, 2019; Diegert et al., 2023), which is described in more detail in Section 1.2.

The OLS approach to tracing out wage-amenity variation generated by preferences requires the researcher to control for the full set of variables that determine workers' abilities. In contrast, my approach requires only one such variable. Both approaches, it should be noted, are sensitive to a normalization on the part of the researcher as to which variable (or variables) belong as proxies for the offer set, rather than correlates of preferences. My approach allows me to explore the robustness of results when the offer set is proxied by any one of various types of cognitive, social, and physical characteristics of workers. Still, like other hedonic approaches, my approach is only suitable when the researcher believes they observe the key aspects of jobs that determine utility.[5]

of latent quality in the housing production function.

[4](e.g., anti-protons spin in the opposite direction as protons.)

[5]Empirically, this paper's main findings on demographic pay gaps seem unlikely to be driven by problems of unobserved amenities. For instance, I find that conditional on total wage and amenity compensation, women tend to earn better amenities and lower income than men. This finding implies amenities play a substantial role in the current gender pay gap, which seems difficult to explain by the problem of observing too few amenities. In stark contrast, I find more limited roles for the same set of amenities to explain gaps by race and by parent background. Unlike the gender gap, these income inequalities are shown to be primarily absolute gaps in total compensation rather than a function of how workers split their compensation into wages and amenities. Although the possibility still remains that additional racial and intergenerational income inequality is generated by amenities other than those I observe, my findings are consistent with a literature that has focused on the role of the absolute factors of skills and discrimination in causing these gaps. My findings on the gender gap are also consistent with a modern literature

The appeal of the approach is not just theoretical—the results relative to existing methods could not be starker. While intuitively "wrong-signed" amenity prices have plagued the literature, I have encountered that issue rarely if at all since adopting the framework.[6] In contrast to the challenges highlighted by contemporary work like those surveyed by Mas (2025), this paper reports that workers must give up pay to acquire jobs that: are safer, in the nonprofit industry, have more regular schedules, allow them to sit, afford them decision-making freedom, allow them to structure their tasks, involve repetition, or have less time pressure.

While these price estimates garner interest in their own right, they also hold potential to provide a new lens through which to view labor market *income* inequalities. The credible price estimates yield a new research question: How do costly amenities distort labor income inequality? By applying the price estimates to demographic gaps by gender, race, and parental income, I find most notably that the majority of the gender pay gap can be understood through a lens of costly amenity substitution. Put differently, the product of amenity differences by gender times the current amenity prices constitutes about two-thirds of the overall gender income gap. The implication of this analysis is that is that policies to reduce the costliness of several labor market amenities are particularly promising avenues to combat gender inequities in the labor market. This paper provides the first large-scale evidence from a nationally representative dataset about the salient role of costly amenities in gender inequality.

The identification strategy rests on two core assumptions. First, I assume the quality of jobs (or, analogously, the desirability of workers) can be summarized by a single vertical index of total compensation, such that wage and amenity lie on a common frontier. Second, I assume that a proxy for that vertical index—such as AFQT—is conditionally independent of preferences over amenities given this compensation level. These assumptions are sufficient to identify the amenity price, even in cross-sectional data, and are notably weaker than those required by typical observational approaches.

The outline for the rest of the paper is as follows. In the remainder of this section, I provide an intuitive example of the estimation framework and relate my work to previous research on compensating differentials as well as subsequent research citing earlier unpublished versions of this paper. In Section 2, I provide my rendition of the usual structural model of compensating differentials; I discuss the identification assumption in the context of that structural model and in the reduced-form. Section 3 proves formally that the Rosen frontier is non-parametrically identified by observational data, and Section 4 discusses estimation in the linear model, including potential sources of bias. Section 5 explores amenity price estimates in the NLSY linked with O*NET occupation data. Section 6 relates these price estimates to differences in quantities of amenities by gender, race, and parent background. Finally, Section 7 concludes.

1.1 Intuitive Example

Figure 2 illustrates the difference between OLS and my proposed strategy using hypothetical data constructed to mirror the important facets of real labor market data. Suppose the researcher observes only three coarse bins of workers' ability, represented by colors. The correspondence between observed ability and the actual offer set is quite noisy; some observably low-skilled workers enjoy very high wage and amenity. Panel A depicts the standard regression of wage on amenity with indicators for observed ability (as an average of regressions within observed ability levels).

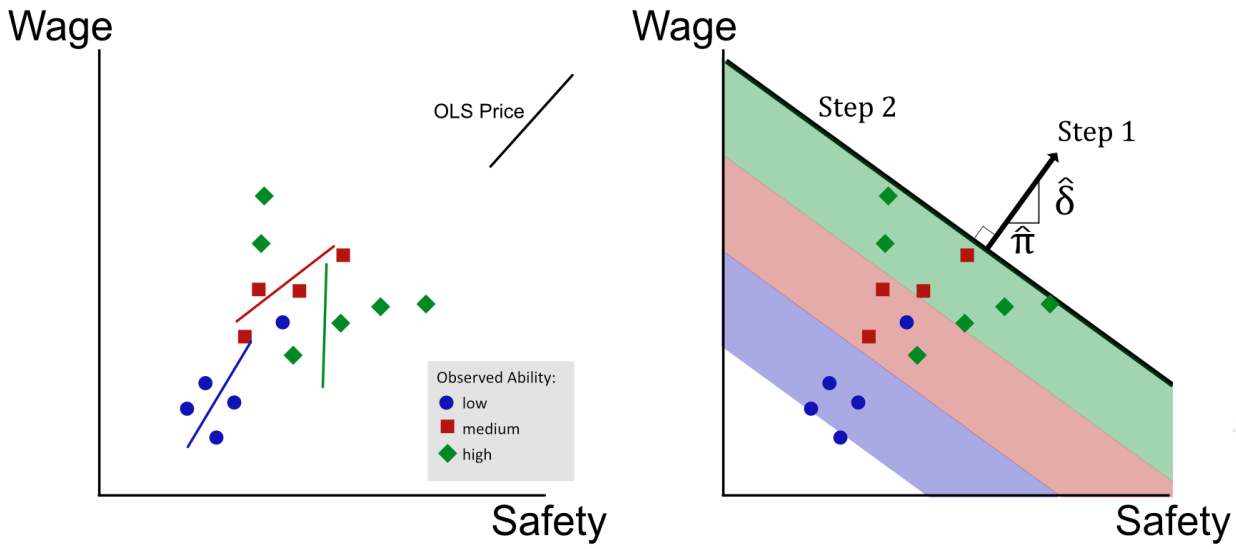
pointing to the importance of how work is remunerated.

[6] In theory, missing amenities should still have the potential to cause bias, potentially as severe as flipping the signs of estimates. The fact that this happens so infrequently if at all in practice is likely due to the tendency for good amenities to come bundled together in the labor market.

Figure 2: Graphical Intuition

(a) OLS

(b) Proposed Estimator



Notes: This figure describes two estimation strategies on hypothetical data in which higher-ability workers get more wage and amenity, such that a trade-off exists among workers of a given ability. This vertical segmentation is imperfectly observed by the researcher in only 3 rough bins, which I call "observed ability." Panel A describes a regression of wage on amenity with fixed effects for 3 levels of observed ability. The fixed effects estimator here is depicted as the within-ability estimator of wage on amenity, and the final average over all ability levels is given by the black line. The upward-sloping OLS estimate mirrors the finding of Figure 1. Panel B describes the anti-IV estimator of observed ability on wage and amenity, described in more detail in the text. The slope of the regression line labeled Step 1, which tells us the direction in which observed ability increases, is $\frac{\hat{\delta}}{\hat{\pi}}$. Of greater interest, the line orthogonal to that regression line, labeled Step 2, corresponds to wage-amenity variation when skill is held constant. The formula for the slope of that line is the negative inverse of the regression line, $-\frac{\hat{\pi}}{\hat{\delta}}$, which corresponds to the true price of the amenity in this model. Alternatively, the exact estimate could have been found by obtaining predicted values from the regression of observed ability on wage and amenity (denoted $\hat{ability}$), then regressing wage on amenity with those predicted values as a control. Approximate levels of $\hat{ability}$ appear on the graph as colored bandwidths.

The estimating equation is as follows, with coefficient of interest $\hat{\beta}$:[7]

$$\hat{E^*}[wage|amenity, obs.\ ability] = \hat{\beta} amenity + I[obs.\ ability]$$

Because I observe workers' true abilities with noise,[8] within every level of observed ability there is still an upward-sloping relationship between wage and amenity. This explains the previous upward-sloping relationship of Figure 1, even with controls for observed ability.

In contrast, Panel B sketches two ways to implement my proposed estimation strategy in a linear framework. Both approaches begin the same way. Rather than using observed ability as a control, it is the outcome of a regression on both wage and amenity. I call this Step 1 on the graph:

$$\hat{E^*}[obs.\ ability|wage, amenity] = \hat{\delta} wage + \hat{\pi} amenity$$

Although the regression of Step 1 is not structural, it does provide some intuition. If wage and amenity are two forms of compensation that workers enjoy, I would expect both $\hat{\delta}$ and $\hat{\pi}$ to be positive. The relative magnitudes of the coefficients reveal the tendency for higher-ability workers to get more wage and more amenity. The regression also has a graphical interpretation in wage-amenity space. The line points in the direction in which observed skills – and thus by assumption offer sets – are increasing. This direction is orthogonal to what I want to estimate; I seek to estimate the wage-amenity trade-off when the offer set is held *constant*. The orthogonal direction is labeled as Step 2 in the graph, and it can be obtained either from a ratio of coefficients or using predicted values from Step 1. In a linear model, the two approaches produce numerically equivalent estimates (much like instrumental variables and control functions).

The simplest approach is to read off the perpendicular line to the regression line of Step 1. If Step 1 told us the direction in which skill increases, the orthogonal direction must describe wage-amenity variation that holds skill constant, which is the relative price a worker faces. Because the regression line of Step 1 has a slope of $\frac{\hat{\delta}}{\hat{\pi}}$, the slope of the orthogonal line and thus the price estimate is $-\frac{\hat{\pi}}{\hat{\delta}}$. I refer to this procedure as the ratio-of-coefficients approach.

An alternative approach uses predicted values from the regression of Step 1. The strength of the predicted-values approach, as I elaborate on in Section 3, is that it can easily accommodate a non-parametric estimator in place of the linear regression in Step 1. Referring to these predicted values as $\hat{ability}$, Step 2 of the predicted-values approach is to regress wage on amenity with the inclusion of the constructed control:

$$\hat{E^*}[wage|amenity, \hat{ability}] = \hat{\beta} amenity + \hat{\gamma} \hat{ability}$$

This regression also need not be linear, although $\hat{\beta}$ from the linear regression is numerically equivalent to the ratio of coefficients, $-\frac{\hat{\pi}}{\hat{\delta}}$.[9] The predicted values of the first stage serve as an ideal control for the offer sets. If the ability proxy was a noisy measure of the offer set, then that noise is purged in taking the predicted values. Critically, in order to correctly write the offer set as a function of wages and amenities, it must be the case that there is only one latent variable in the estimating equation that relates workers' wages to amenities. I refer to this latent variable interchangeably as the level of the offer set or the total compensation level of the worker. It is this sole latent variable that will be proxied by observed qualifications.[10]

[7] I use notation $\hat{E^*}$ to denote regressions throughout the paper (the linear conditional expectation estimator).

[8] The ability bias problem can be framed either as measurement error (observed variables like education are effectively noisy measures of the offer set) or analogously as omitted variables (some relevant controls are not observed).

[9] In this case, because $\hat{ability}$ is a linear combination of the outcome and a regressor, this approach yields a perfect fit; standard errors from the second-stage regression are not meaningful (I discuss identification-robust inference in Section 4.2).

[10] To build further intuition for this predicted-values approach, recall that the basic structure of the problem is

Whereas either estimation approach overcomes the problem of ability bias in observational data, a key limitation remains: like with typical hedonic approaches, my approach is really only suitable when the researcher feels they observe a sizable set of utility-generating job attributes. In other words, my method provides no new way to price characteristics that the researcher cannot observe. Real workers make choices taking into account many aspects of jobs, and my estimator readily extends to pricing many amenities jointly. As I discuss in Section 5.4, the likely "bundling" together of observed and unobserved amenities would cause my price estimates to be larger than those that priced the full set of amenities. Though the problem of unobserved amenities calls for caution when interpreting price estimates, a positive correlation between observed and unobserved amenities if anything would lead one to interpret my analysis of the large role of costly amenity substitution in the gender pay gap as a *lower* bound.

1.2 Recent Applications and Prior Approaches

While the empirical application contained in this paper shows one way that the anti-instrument method can be useful, several other works citing earlier drafts of this paper have now also used this method to yield important economic insights. Most applications of the method to date have been in labor economics. Burbano et al. (2023) estimate the price of meaningful work using multiple methods, largely finding it to be a disamenity; using their anti-instrument estimates, they show that job meaning accounts for 9.5% of Sweden's gender pay gap (that their result is smaller than mine is likely due to pricing fewer amenities). Folke and Rickne (2022) apply it to Swedish data, finding no compensating differential for sexual harassment risk, consistent with broader evidence of gendered sorting. De Schouwer et al. (2025) apply the method to the work meaning data collected by Maestas et al. (2023), while De Schouwer and Kesternich (2024) use Dutch data to study the flexibility puzzle, showing that flexibility is traded off against other amenities like meaning, a pattern not detectable with prior approaches. Ferreira et al. (2024) apply an anti-instrument approach to European survey data to value job safety, finding that using job satisfaction as a proxy yields greater precision than education. Maideu-Morera (2024) adapts the method to the macroeconomic context to test a hypothesis of "amenity-biased" technical change. The method has also been developed beyond labor markets. Bell et al. (2025a) establish identification and econometric properties of the anti-IV approach for hedonic price functions of the housing market, and apply it to pricing residential amenities such as flood risk. Bell et al. (2025b) develop a migration-based proxy for neighborhood quality to study the pricing of air quality.

The key difference between the anti-instrument method and prior approaches is in the treatment of the residual when regressing wages on amenities. My paper is the to explicitly model that residual as vertical differentiation, or job quality, and to acknowledge that this index is partially latent. The seminal model of Rosen (1986), on which my model is largely based, essentially assumes away any residual when measuring the tradeoff between wages and amenities with a regression. Although this "no-residual" model has been replaced in much of economics by newer perspectives on search frictions and other essentially idiosyncratic forces (Mortensen, 2003), it still remains the dominant lens through which researchers analyze job choice from observational data (Lucas, 1977; Brown, 1980; Mas and Pallais, 2017; Sorkin, 2018; Mas, 2025).

---

that workers' wages and amenities are both caused by differences in offer sets and preferences. The goal is to build a control variable for offer sets such that at a given level of the control, workers still obtain different combinations of wages and amenities (due to preferences). Thus, the ideal control variable for the offer set can be written as some combination of wages and amenities. A prediction exercise of observed ability (in this case, a linear regression) guides the construction of that ideal control variable. In a large enough sample, I would obtain the same predicted values if either true offer sets or a noisy proxy for the level of the offer set were used as the dependent variable in Step 1.

Additional assumptions can yield identification of additional parameters (Taber and Vejlin, 2020; Lamadon et al., 2022). In contrast, my method requires no distributional or functional form assumptions to achieve identification, providing researchers with estimates of wage-amenity tradeoffs that are more easily interpreted.

A key advantage of quasi-experiments, relative to the present paper, is that they can recover price estimates that control for unobserved amenities. Quasi-experiments have shown some promise to identify wage-amenity tradeoffs faced by particular groups of workers, but are limited in that they typically can estimate prices only for very specific amenities—typically physical safety—for very specific groups of workers (Ashenfelter and Greenstone, 2004; Greenstone et al., 2015; Lavetti, 2018).

A related but different literature has focused on estimating workers' valuations of amenities, rather than market prices (Stern, 2004; Mas and Pallais, 2017; Wiswall and Zafar, 2018; Mas and Pallais, 2019; Maestas et al., 2023). To clarify this distinction, a group of workers may value an amenity a great deal, but this fact alone does not imply they sacrifice wages to get it; that depends on the price. Rather, a price is an equilibrium between workers' heterogeneous valuations of an amenity and firms' heterogeneous costs of providing it.[11] Although willingness to pay has proven easier to estimate, the more policy relevant parameter when understanding income inequality is really the price of the amenity. Most of these papers do not attempt to estimate prices, or as in the case of Mas and Pallais (2019), do so only using a very explicit no-residual model.

Finally, the assumption that a single imperfectly observed variable exists in the neoclassical amenity pricing equation connects tangenetially to a growing literature on selection corrections, in particular as put forward by Altonji et al. (2005) and Oster (2019). In the presentation by Oster (2019), the estimator that I propose, motivated by a single-index assumption, is numerically equivalent to a selection correction that sets $R_{Max} = 1$, which was also the assumption of Altonji et al. (2005).

2 Economics of Compensating Differentials

My conceptual framework follows at a high level the model of Rosen (1986) in the Handbook of Labor Economics. The key divergence is that I allow for a single dimension of vertical job quality as well. The intuition behind this addition is few if any workers actually choose between being a coal miner and being a professor. Rather, factors exist that segment the labor market such that workers who choose to be coal miners typically cannot become professors. These vertical factors go by many names — skill, luck, search frictions, discrimination — but the important aspect is that they are, to a large extent, not observable to the researcher.

This section provides an intuitive background for such a model, while Appendix B provides a more formal treatment.

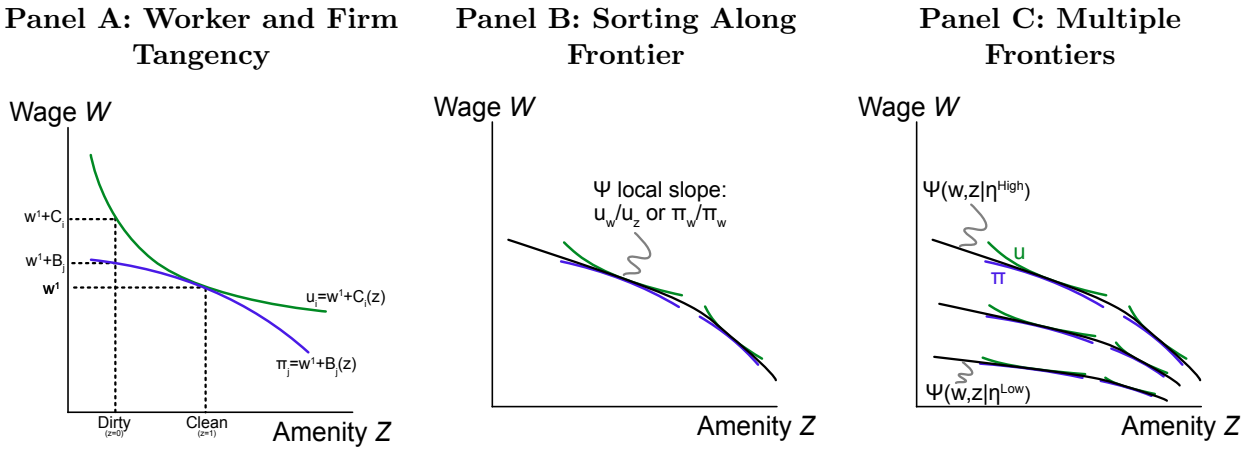
The two key parameters are horizontal and vertical heterogeneity, which are represented in Figure 3. Vertical heterogeneity is the tendency of some workers to get better jobs (i.e. the professor versus the coal miner referenced above), whereas horizontal heterogeneity is a tendency toward different jobs (e.g., professor versus consultant, for instance). Suppose only one amenity exists, such that horizontal heterogeneity can be indexed at the individual level by a willingness-to-pay parameter, θ.[12] Firms also have a cost of providing the amenity, which leads to sorting

[11] To illustrate this point, suppose that a stated choice experiment revealed that workers' willingness to pay to sit at work at a blue versus pink desk differed dramatically by gender. But it would be silly to attribute the gender wage gap to this amenity given the price of blue versus pink paint is the same.

[12] Classically, θ_i is sometimes also associated as a preference parameter of a utility function, though the more

on the continuum of worker and firm parameters, resulting in an equilibrium price schedule. But in the midst of this horizontal sorting is the vertical dimension: some positions differ in their exogenous staffing needs and pay out more total compensation, composed of wages and amenities, in accordance with what type of applicant they seek to attract. I index that dimension as η, and the multiple price frontiers, conditional on η, are shown in Panel C of Figure 3.

Figure 3: The "Kissing Equilibrium"



Notes: This figure graphically represents the equilibrium between labor supply and demand discussed in Section 2, and illustrates how the case of a binary amenity can be thought of as a discretization of a continuous model. In Panel A, an indifference curve for a particular worker is drawn in green, such that C_i represents the amount of wage required to make the worker indifferent with a one-unit change of amenity. The blue curve depicts the analogous level-set of profits for the firm that has matched with the worker. The particular level sets of utility and production are chosen such that they meet at a tangency. Panel B shows the tangency for this pair and for an additional worker-firm pair. The frontier ψ is drawn in black such that it passes through the tangencies for all pairs, and such that ψ is tangent to each worker/firm indifference curve (which restricts which level-sets each pair chooses). Appendix Figure A.5 illustrates that if a worker-firm pair chose a level of wages and amenities such that the frontier did not satisfy the latter tangency condition, then the match would not be stable. Finally, Panel C demonstrates the same sorting equilibrium for multiple sub-markets that are ordered by how much compensation they entail. For a single worker of a particular ability level η, the corresponding ψ frontier constitutes the set of obtainable jobs, similar to a budget set. The local slope of ψ is viewed by the worker as the price of the amenity, ins wage units.

Naming the vertical dimension as η allows me to be more precise in relation to the prior literature. What I have referred to as the "no-residual" model makes the assumption that η either is perfectly measured (no residual) or this residual for some reason affects wages, but does not affect amenities. If this vertical dimension did affect amenities, and was imperfectly measured, then it would have caused omitted variables bias in the typical approach.

The key innovation of this paper is to re-cast the way researchers should view latent job quality as affecting not only workers' wages, but *also* their amenities.

Aside from this one new and improved assumption about residual wage-amenity variation, all other assumptions are similar to what one would need to draw meaning from a regression of wages on amenities with controls. Like earlier hedonic methods, my framework still requires the researcher to observe a sufficiently rich set of amenities. Also like earlier methods, the researcher must have what Angrist and Pischke (2009) would refer to as a "good control" for the vertical dimension η, although it is an improvement that the researcher does not need to observe all of η. I refer to this

general terminology of "horizontal" differentiation is in some applications also more accurate because it reflects that workers may have constraints from outside the labor market (like childcare) that influence job choice.

variable as h, which I also refer to as an anti-instrument because it is being used to hold constant the (latent) index η. Mirroring the "Conditional Independence Assumption" of Angrist and Pischke (2009), the exclusion restriction that identifies the model I have outlined can be written in terms of structural primitives: $(h \perp \theta)|\eta$ (1)
or in a more reduced-form way: $((w, z) \perp h)|\eta$ (2)

Identification will rely on a variable that is informative about how high of a compensation frontier the worker lies on (η), but adds no *further* information about the location of the worker along that frontier. Put another way, h would be irrelevant to w and z if the researcher could observe and condition on η. Appendix B provides a more detailed mapping to the notion of "bad controls" in the typical sense.

In the empirical application, I will argue that score on the Armed Forces Qualification Test (AFQT) is correlated with the dimension of job quality for which we would ordinarily try to control, but is independent of workers' locations along their particular frontiers. The AFQT score is the primary summary score from the Armed Services Vocational Aptitude Battery, which was administered to respondents of the NLSY79 in 1980. As the purpose of the test is to measure basic skills that are applicable to a wide range of vocations, the score by construction should be independent of how employers located at various positions on a wage-amenity frontier view a worker. Additional discussion of this choice is in Appendix B; we will also see that results are invariant to several other candidate anti-instruments from diverse realms including physical, socioemotional, and intellectual traits.

3 Non-Parametric Identification

In this section, I prove that the Rosen frontier can be non-parametrically identified under the equilibrium exclusion restriction of Condition 2, in which ability proxy h is assumed irrelevant to wage-amenity sorting conditional on true ability η. In the next section, I apply the non-parametric result to a data-generating process assumed to be linear, in which I discuss issues of potential bias.

3.1 Object of Interest

Denote wage by w, amenity by z, and total compensation by η. The goal is to learn what wage and amenity values are associated with each level of total compensation, denoted in Section 2 as the function $\psi(w, z|\eta)$.

3.2 Assumptions

I make the usual standard assumptions: an IID sample of w, z, and h; and, all variables are assumed continuous. The substantive assumptions are:

1. Conditional independence of h: $((w, z) \perp h)|\eta$
2. Monotonic relevance of h: $E[h|\eta]$ is everywhere strictly monotone in η
3. η is degenerate given values of w and z; i.e., there exists a non-stochastic function g such that $\eta \equiv g(w, z)$

Assumption 1 is what I have referred to as the "identifying assumption," although it was also an identifying assumption for the typical approach that controlled for measures of the offer set. Assumption 3 is the assumption that enables identification by the anti-instrument variable. The global strict monotonicity of Assumption 2 can be relaxed; Online Appendix D discusses identification under a more local form of monotonicity.

3.3 Proposition

Under these assumptions, $\psi(w, z|\eta)$ is identified for all levels of η.[13]

3.4 Estimator

Define the random variable $\hat{h}$ as an approximation to $E[h|w, z]$. For simplicity, $\hat{h}$ may simply be a conditional mean of h in given bins of w and z.[14] Denote with subscripts η_i as quantile i of the unobserved variable η and $\hat{h}_j$ as quantile j of the constructed variable $\hat{h}$. For all observed w and z, $\psi(w, z|\hat{h}_j)$ approximates $\psi(w, z|\eta_i)$ for some unknown quantiles i and j.

3.5 Proof of Identification

The first step is to note that the constructed variable $\hat{h} = \hat{E}[h|w, z]$ can be equated, in probability limit, with $E[h|w, z, \eta]$. This is because η is assumed degenerate given w and z, so adding it to the conditioning is redundant (Assumption 3).

In the second step, I rely on conditional independence of Assumption 1 to equate $E[h|w, z, \eta]$ from above with $E[h|\eta]$. This is intuitive, since the wage-amenity split is by assumption not related to h.

At this point, I have shown that under the conditional independence assumption, $\hat{h}$ approximates $E[h|\eta]$. My last assumption is that $E[h|\eta]$ is strictly monotone in η (Assumption 2). If this is the case, then each quantile of the control function corresponds to exactly one quantile of η. Conditioning on an arbitrarily small range of the control function makes the correspondence to the level of η arbitrarily close.

Quantiles of $\hat{h}$ can be mapped to quantiles of η as follows. Consider a worker i with observed w_i and z_i and an unknown level of compensation, η_i. The researcher is interested in learning what other observed values of of w and z are associated with this particular amount of total compensation and are thus also in the worker's choice set; I denote this by the function $\psi(w, z|\eta_i)$. To find the quantile j of $\hat{h}$ such that $\psi(w, z|\hat{h}_j)$ approximates $\psi(w, z|\eta_i)$, all that one needs to do is find the $\hat{h}_j$ such that $\psi(w_i, z_i|\hat{h}_j) = 1$. Since $\hat{h}$ will be defined on any point for which w and z are observed, for all observed w and z, $\psi(w, z|\hat{h})$ identifies $\psi(w, z|\eta)$ for all levels of η.

4 Estimation in the Linear Model

For the purpose of discussing potential biases, it is now useful to restrict the model to one in which price can be summarized by a single parameter. The non-parametric estimator entails nothing more than plotting conditional means of the h variable by bins of wage and amenities. The linear proposed estimator is best understood as a special case of that non-parametric estimator, in which conditional expectations are allowed to be replaced with their linear counterparts and can thus be estimated by linear regression. Relative to the non-parametric framework, I will introduce two additional assumptions of linearity: linearity of the Rosen frontier and a linear relevance of h to η.

In the non-parametric framework, the goal was to identify a function of w and z given unobserved η. Here, I will take the object of interest to be the coefficient on z from a hypothetical regression of

[13] To clarify, the units of η are to be viewed as a nuisance parameter. The goal is to learn the values of w and z associated with the same η, but will remain agnostic about the cardinal units of that η.

[14] Alternatively, optimal kernel estimators might be considered for efficiency, which is not the present focus.

w on z controlling for η. That is, I will assume $E[w|z,\eta]$ is linear in z, with homogeneous parameter β, such that each Rosen frontier can be written of the linear form:

$$w - \beta z = \eta \tag{3}$$

I have omitted a scale coefficient on η as it makes no difference since η is unobserved and I have normalized the price of wage to 1. The content of Equation 3 is simply that I need to assume that conditional on η, the values of w and z form a line if I am to summarize this relationship with a single parameter. The second assumption of linearity is more nuanced.

In the non-parametric framework, I assumed $E[h|\eta]$ was strictly monotone in η. Now, I will strengthen monotonicity to linearity. In other words, $E[h|\eta]$ can be written $\gamma\eta$ for some unknown γ.

With these additional assumptions, one can modify the proof of identification in Section 3 to replace all conditional expectations with linear regressions. The table below provides a side-by-side comparison of the non-parametric and linear estimators. I have already shown graphical intuition for the linear approach using predicted values in Figure 2; the non-parametric approach simply uses predicted values that are not required to be linear functions.

Non-Parametric	Linear
1. $\hat{h} = \hat{E}[h\|w,z]$	1. $\hat{h} = \hat{E}^*[h\|w,z]$
2. Plot w on z at various quantiles of $\hat{h}$	2. Regress w on z controlling for $\hat{h}$

4.1 Bias when $(w,z) \not\perp h|\eta$

In the linearized model, the bias in $\hat{\beta}$ takes the sign of $cov^*(w,h|\eta)$. Online Appendix G contains a formal proof, but Figure 4 conveys the key intuition in the linear model. No bias occurs when the h variable is evenly distributed along each Rosen frontier. To the extent that high-h workers bunch on one end of the frontier, they will pull the estimate with them.

For instance, perhaps one suspects that a candidate h variable is positively correlated with workers' preferences for amenities (relative to wage). The excess bunching of high-h workers at high amenity levels would cause us to estimate that this is a more costly amenity than it actually is. This argument matches what is depicted in the figure, as would a story in which h was a skill that was particularly valued in high-amenity jobs (relative to high-wage jobs).

4.2 Finite-Sample Inference

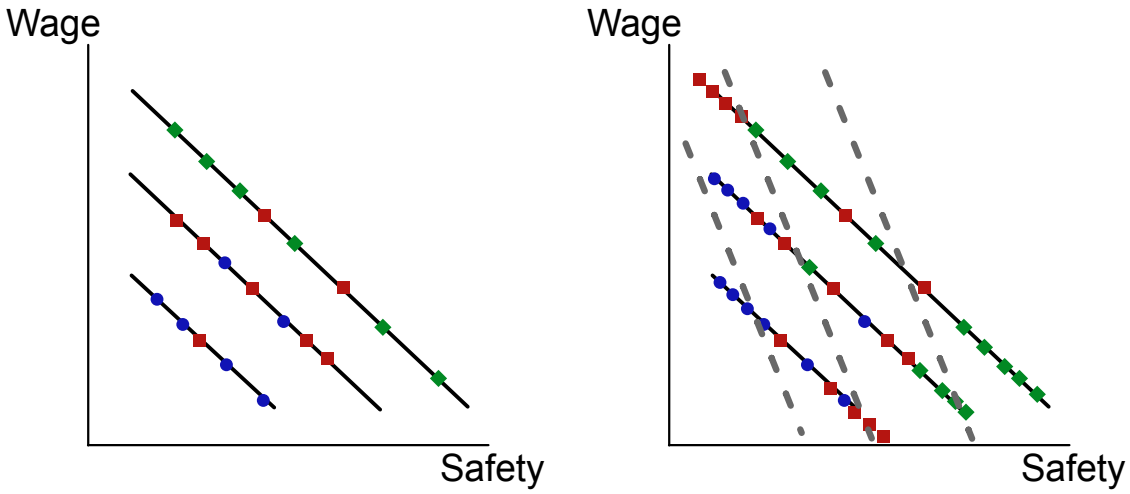
In general, the linear proposed estimator is vulnerable to a finite-sample problem similar to that of weak instrumental variables whenever the relationship between h and w is difficult to detect empirically. The problem is the result of reporting the price of the amenity per unit of wage, which puts the wage coefficient into a denominator.

Recall the ratio of coefficients estimator for the price is $-\frac{\hat{\pi}}{\hat{\delta}}$ from a regression of h on w and z: $\hat{E}^*[h|w,z] = \hat{\delta}w + \hat{\pi}z$. Whenever $\hat{\delta}$ may dip close to 0, the ratio of coefficients explodes, causing price estimates to skew large. In light of the potential for bias in finite samples, I report Anderson-Rubin confidence intervals when appropriate, following the suggestion and approach of Andrews et al. (2019).

4.3 Measurement Error in w and z

This estimator is affected by measurement error in a very different way from the typical OLS approach. Like the finite sample bias, the bias from measurement error is best understood from

Figure 4: Illustration of Bias



(a) Unbiased (b) Biased

This figure discretizes a hypothetical labor market into three segments of high, medium, and low compensation, represented by linear Rosen frontiers drawn as heavy black lines. Colors are used to depict levels of observed h; in ascending order blue then red then green (circles, squares, diamonds). No bias is present in Panel A because within each η-level frontier, h is distributed independently of w and z. When the ability channel is normalized to this h variable, the estimated price would lie on-top of the true frontiers as desired. In Panel B, additional observations have been added at the edges of the frontiers such that h is clearly not independent of w and z conditional on the frontier. Matching the example discussed in Section 4.1, on each frontier workers with the highest h now also have the highest z, and likewise those with the lowest h have the highest w. The positive conditional covariance of h and z leads to a rotation of the price estimate, shown in dashed lines. The slope estimate is more negative than the true one. In other words, the estimator would suggest that workers give up more wage for this amenity than is truly the case.

the ratio-of-coefficients approach, which runs the regression $\hat{E}^*[h|w,z] = \hat{\delta}w + \hat{\pi}z$ and takes the ratio $-\frac{\hat{\pi}}{\hat{\delta}}$.

Classical measurement error in h causes no particular bias on $\hat{\pi}$ or $\hat{\delta}$. This is a key advantage of my estimator; noise in the dependent variable of a regression only causes imprecision but not bias. (Although, as previously discussed, this can itself cause finite-sample bias.)

Measurement error in a regressor, on the other hand, causes the coefficient on that regressor to be attenuated. If w is measured with noise, then $\hat{\delta}$ is too small, so the estimated price of the amenity is too large in magnitude. If z is measured with noise, then $\hat{\pi}$ is too small, and the price of z is too small. Because all components of compensations are regressors, measurement error in any piece of compensation will simply make it appear less important or valuable. Of course, because measurement error in wage and amenity have the potential to balance each other out, there is also the case in which both variables are measured with a balanced amount of noise such that the resulting price estimate is unbiased. A fuller treatment of measurement error is given in the reduced-form statistical model of Online Appendix E.

5 Price Estimates

5.1 Data

The primary data sources for the analysis are the NLSY79 and O*NET. The analysis sample is NLSY respondents employed in 2012 linked at the occupation level with O*NET measures of amenities.

The NLSY79 is a longitudinal survey dataset constructed by the Bureau of Labor Statistics. The first wave, administered in 1979, sampled 12,686 men and women between the ages of 14 and 22. Subsequent survey waves were administered annually until 1994, and once every two years since. The NLSY contains three sub-samples. The primary sample (6,111 people) is "designed to represent the non-institutionalized civilian segment of young people living in the United States in 1979 and born January 1, 1957, through December 31, 1964." The NLSY contains two additional samples; the first (5,295) is a sub-sample of racial minorities and economically disadvantaged people, and the second (1,280) is designed to represent those serving in the military as of September 30, 1978. To maximize power, I include all three samples.[15]

I focus on data in 2012, as the occupation data has changed across years. By 2012, only 7,300 of the initial NLSY respondents were retained. Of the 5,386 not interviewed, the biggest driver has been administrative decisions to reduce the sizes of the non-representative sub-samples.[16] My final analysis sample consists of 4,868 NLSY respondents who reported being employed in 2012, when they were aged 48-55. This number excludes 152 individuals who I drop due to missing job amenity fields. To be employed, the respondent must have worked at least ten hours per week and at least 9 weeks since the date of the last interview, or be self-employed. Although I conduct the main analysis on this broad sample, I also test robustness to selecting on the full-time employed population. In the analysis of demographic pay gaps, I divide the NLSY by gender and by race. Gender and race both correspond to the respondent's data in 1979, at the start of the survey, as determined by an interviewer. Sex and gender were not coded separately at this time, so I use them interchangeably. The racial/ethnic categories are as follows: "Hispanic," "Black," or "non-Black/non-Hispanic." For

[15] In the final analysis sample, only 6 workers list Armed Forces (Census code 984) as their occupation in 2012.

[16] 84% of the military sub-sample was dropped after the 1984 survey, and in 1990 31% of the disadvantaged sub-sample was dropped (poor White workers). After the 2,722 individuals dropped for one of these reasons, the most common reasons for attrition were refusal (903), can't locate (466), deceased (689), other (481), and difficult cases (125).

the analysis by race, I will compare the Black group to the non-Black/non-Hispanic group, which I will at times refer to by the approximate term "White" for brevity.

For additional results on income inequality by parent background, a PSID sample was constructed to mirror the sample definitions of the NLSY. The PSID sample contains 929 workers in 2013 linked to their parents. For each child, I define a measure of parental income to be the father's income averaged over all years for which the father was between 30 and 50. Because of the requirement that I observe income for both workers and their fathers, this sample of workers is substantially younger than the NLSY one, between ages 30 and 40.

O*NET, which stands for Occupational Information Network, is a database of occupational characteristics maintained by the US Department of Labor. The data are derived primarily by occupational surveys and supplemented in some cases by occupational experts. I draw primarily on the O*NET Work Context table, which includes attributes that workers might plausibly have varying preferences over. Examples include how regular the schedule is, how often minor cuts and burns arise on the job, whether the work is indoors or outdoors, or how much freedom there is to make decisions. As the scales of most of these measures are difficult to compare, I translate income and all amenity measures to standard deviation units across workers in the sample. Table A.1 describes the process by which several amenity measures were constructed from O*NET.

I make two key deviations from the logic of the simple conceptual framework in which workers choose jobs indexed by a wage and amenity level. First, whereas the canonical framework allows workers to choose particular jobs, in my data many amenities are only observed at the occupation level via the O*NET occupation survey. The difficulty of measuring amenities at the job rather than occupational level has been discussed in prior work (e.g., Goldin (2014); Denning et al. (2019)). Due to this limitation, I aggregate the entire analysis to the occupation level, rather than job level, and weight by the number of workers in each occupation.[17] I also show robustness of the final results to estimation without aggregation. Collapsing the data to occupation level entails a loss of power. After mapping the O*NET SOC occupation codes to the same Census format as is used in NLSY in 2012, I am left with only 397 occupations that are represented in the NLSY in 2012 (451 across all years). This means that 397 is the effective sample size, or number of clusters, for most of the analysis.

Second, while the basic framework considers the trade-offs workers face between wage and amenities, in practice most workers in the sample do not report being paid an hourly wage (Figure A.2). Although it is possible to impute an hourly wage based on workers' estimates of hours worked, to more accurately mirror the level at which most workers are making decisions, I use yearly labor income rather than hourly wages. Labor income in the NLSY is defined as total pre-tax income from wages and salary in the past calendar year, but if a worker holds more than one job then it is aggregated over all jobs. The amenities assigned, however, will correspond to the worker's "primary" job. Assuming that any additional jobs held by the worker have similar amenities to the primary job, then this data definition can be viewed as the relevant yearly income and typical amenity levels consumed by the worker.

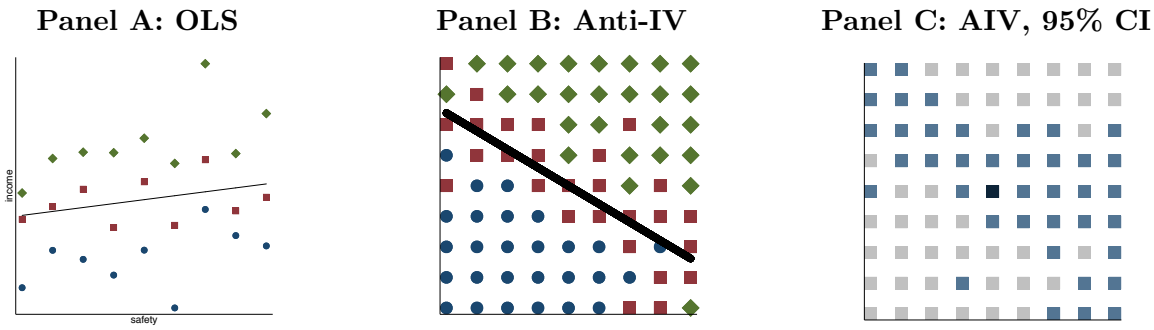
5.2 Functional Form & Relation to Prior Approaches

As a first look at empirics, Figure 5 compares the OLS approach with the proposed approach, overlaying linear estimates on-top of their non-parametric counterparts. Similar to Figure 1, the OLS approach in Panel A entails regressing income on safety, controlling for the ability proxy

[17] Appendix Table A.2 contrasts the findings of the typical OLS approach on data at the individual versus occupation level; even with OLS, aggregation of the entire analysis to the occupation-level does seem to remove some, but not all, of the wrong-signed bias.

of AFQT by way of terciles. Within terciles of AFQT, the income-safety relationship is upward sloping. In contrast, the proposed approach in Panel B uses precisely the same data but plots mean AFQT by bins of income and safety. Despite identical underlying data, the finding is the opposite: within each tercile of mean AFQT, which should be thought of as the empirical stand-in for the offer curve, the income-safety relationship is now downward sloping, implying that a trade-off exists in the labor market between income and safety.

Figure 5: Non-Parametric & Linear

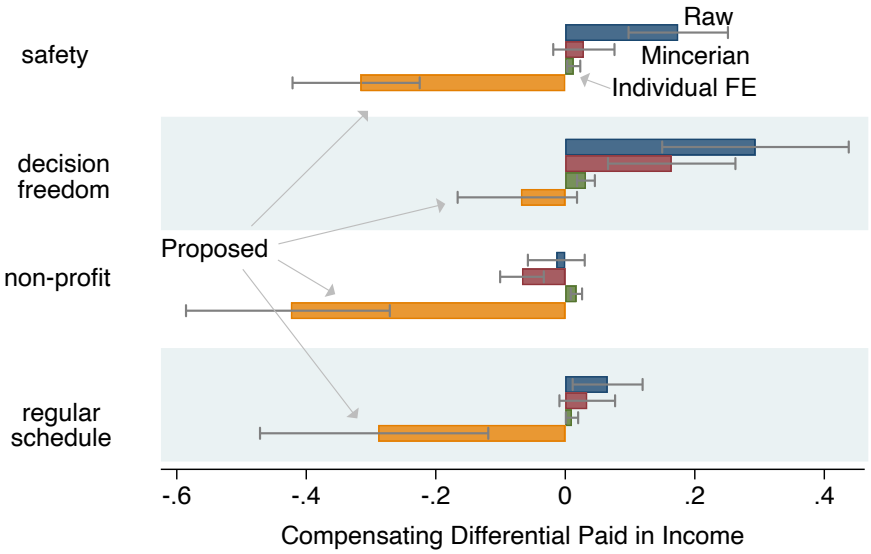


Notes: Panel A is a binned scatterplot of income by bins of safety, separately by terciles of AFQT, in the NLSY analysis sample at the occupation level. Green diamonds are workers with the highest AFQT, then red squares, and blue circles are the lowest-AFQT workers. The black line is the linear best-fit: on average, workers with more safety also earn more income, even within skill bins. Panel B plots mean AFQT by income and safety; these means have been discretized into terciles for visualization. The black line is the linear best-fit of income on safety controlling for the bin's mean AFQT. Panel C is similar to Panel B, except the dots are shaded in the following way: first, the center dot is made dark to indicate it is the reference point. Then, two-sided t-tests are performed to test whether the mean in the reference bin is significantly different from each of the other bins. Bins in which mean AFQT is indistinguishable at the $\alpha = .05$ level are colored in blue, and the rest in grey. For all panels of this figure, income and safety are in percentiles rather than standard deviations. All panels use 9 bins to allow for a centered dot in Panel C. Additional sample details are found in Section 5.1.

Just how steep is the trade off, and is it well characterized by a single parameter? My answers to these questions are limited by the precision of the data. In Panel B, for instance, although a downward trend is apparent, some bins fall off the trend. This is likely due to sampling variation. In Panel C, a formal statistical test is applied to do inference on the graph. Because every bin represents a sample mean of AFQT, I can apply a standard two-sample t-test for difference in means to help determine which regions of the graph correspond to different offer curves. I have arbitrarily picked the center point as the reference point, by which all other means are tested. The test constructs a graphical confidence interval for the price, allowing the researcher to evaluate potential non-linearities around the center point. Although the graphical test can reject many upward-sloping patterns that the wage-amenity frontier passing through the center point might have taken on, there is still a great deal of uncertainty about the exact shape of the frontier. Due to the low level of precision that can be obtained from this relatively small sample, for most of the remainder of the paper, I simply estimate linear average prices rather than attempting to gain non-parametric precision.[18]

[18]Related to the problem of imprecision, the potential for finite-sample bias when summarizing the data linearly can also be detected in Figure 5. Because my data has sampling variation, some of the points in Panel B appear to lie off trend. When the viewing this graph, the noise in the measure of the offer curve level is apparent. Now, consider a regression of income on safety controlling for the estimated level of the offer curve, which contains sampling variation. For any level of the noisily estimated offer curve, there is a weaker relationship between wage and amenity due to the sampling variation, so the linear estimate of the slope of interest looks more like a vertical line that it should. This is the graphical interpretation of the finite-sample bias that was shown algebraically in Section 4.2.

Figure 6: Raw, Mincerian, Fixed-Effects, and Proposed Estimates



Notes: Each data point on the bar graph is a separate (bivariate) regression of the price of the amenity versus wage, controlling for no other amenities, in the NLSY sample at the occupation level. Income and all amenities are in standard deviation units. Mincerian includes indicators for years of education, exact AFQT score, and age. 95% CI's are calculated at the occupation level, and are identification-robust for the proposed estimates using AFQT. All series use the NLSY sample in 2012 except for the individual fixed effects specification, which extends the years to 2004-2014. Additional sample details are found in Section 5.1.

In a linear model, Figure 6 compares several of the typical strategies with the proposed estimator. In this case, I define the Mincerian estimator to be the regression of wage on the amenity with controls for education, AFQT, age, and year. I study the price of amenities one-by-one in relation to income. For instance, examining the amenity of safety, the raw, Mincerian, and individual fixed effects estimators all produce price estimates indicating that the average price of safety is negative (it is a disamenity), similar to what was reported by Brown (1980). On the other hand, the proposed estimator for the price of safety when the h variable is AFQT lies on the other side of 0 and yields a result contrary to all other established methods in observational data. The point estimate indicates that on average, workers must give up .31 standard deviations of income to obtain a typical job with a standard deviation more of safety.[19] Similar descriptive compensating differentials can be found for jobs that have more regular schedules (.29 SD less income), those with more freedom to make decisions (.07 SD less income), and jobs that tend to be in the non-profit sector (.42 SD less income). Next, I turn to examine how results from the proposed estimator compare when proxies of workers' productive abilities other than AFQT are used.

[19] As discussed in the introduction, this is not precisely the "price" of safety, as it does not control for other amenities. The price of safety would describe the trade-off holding all other attributes fixed. Rather, this estimate describes the trade-off between income and a typical safer job.

5.3 Sensitivity to h variable normalization

A strength of the NLSY is its preponderance of measures of skills. There is no way to prove empirically that any potential h variable fits the identification assumption. Whereas each potential ability proxy may always be subject to its own nuanced critiques as to why it may not fit the identification assumption, an opportunity exists to compare results from many different independent measures of skill. If very different proxies for the ability channel yield similar results, this can lend some credibility.

5.3.1 Data on h variables

I have argued that a cognitive test such as the AFQT, though by no means "innate," may be an ability proxy not contaminated by wage-amenity sorting. This would be plausible so long as workers do not systematically manipulate their AFQT in order to gain particularly more of either wage or amenity, rather than total compensation. The AFQT was administered to NLSY respondents in the 1981 survey year.

A variable that may be more pre-determined is the worker's height, which is surveyed in the NLSY in several years including 2012. The positive relationship between height and earnings, which holds even in developed nations, has been well documented though its source is a topic for debate. The cause of this association does not seem to be that height is a skill that is useful in certain physically demanding jobs – i.e., it does not seem to be the case that warehouse firms compete for taller workers who can reach higher boxes. Evidence against this explanation is presented by Persico et al. (2004), who show that the height premium is entirely distinct from a worker's contemporaneous height once height during adolescence is controlled. Thus, the returns to height must be running through an unobserved variable correlated with adolescent height. Persico et al. (2004) propose that the skills and confidence stemming from increased social and athletic experiences during high school may partly be at play. However, more recent work by Case and Paxson (2008) makes a convincing case that taller adults earn more because they have higher cognitive skills. Since physical development is known to be influenced partly by *in utero* and early childhood environmental factors such as disease and nutrition, such environmental factors would also act in the same direction on cognitive development. Whichever combination of social, cognitive, or even taste-based discriminatory explanations one wishes to believe, the evidence indicates that height proxies for an ability channel that is likely to be irrelevant to where workers locate along the wage-amenity frontier. To better isolate the component of height accounted for by environmental rather than genetic factors, I follow the recent literature by residualizing height on sex.[20]

Other plausibly exogenous variables include education and measures of what might be called social skills. Education, measured in the NLSY as highest grade completed by 2012, is a variable often used as one of several controls for ability in the labor economics literature, and has often been included as a control when researching compensating differentials. As discussed in Appendix B.5, any valid ability control from the typical OLS or "Mincerian" specification is a valid h variable here. I also use three scores contained in the NLSY from tests that have been referred to by previous literature as "social" or "non-cognitive" skills that are predictive of compensation (e.g., Heckman et al. (2006); Deming (2017)). These variables are self-esteem, mastery, and locus of control. The self-esteem measure, pioneered by Rosenberg (1965), is based on a 10-item survey that gauges agreement with statements about attitudes toward oneself (e.g., "I have a number of good qualities."). Mastery is a similarly constructed variable based on a survey developed by Pearlin et al. (1981) to study factors that mediate stressful life events such as job losses; it is a

[20] Results are similar when height is also residualized on race.

Table 1: Correlations Among Income & Proposed h Variables

	Income	Educ.	AFQT	Self-Est	Mastery	Loc. C.	Height
Income	1						
Years Educ.	0.3685	1					
AFQT	0.3579	0.5426	1				
Self-Esteem	0.1871	0.2803	0.3175	1			
Mastery	0.1816	0.2392	0.2084	0.2922	1		
Locus of Control	-0.1501	-0.2117	-0.3095	-0.2814	-0.1605	1	
Height (res. sex)	0.066	0.1185	0.1486	0.0593	0.0577	-0.0492	1

Notes : This correlation matrix summarizes the main NLSY analysis sample with all variables non-missing (4,257 workers). Additional sample details and definitions of h variables are contained in Section 5.1.

self-reported measure of a worker's ability to solve problems that affect him or her. Finally, locus of control is a variable constructed from an abbreviated version of the 60-item scale originally developed by Rotter (1966). It is higher when individuals agree with statements that their life is determined primarily by fate or factors outside of their control (external control) and lower when people respond that they direct their own lives (internal control). I use the measure of self-esteem from 1980, mastery from 1992, and locus of control from 1979.

Table 1 displays pairwise correlations among income and the proposed h variables. All correlations are significant when run as regressions between income and the h variable. Years of education and AFQT have the strongest correlations with income. Height has the weakest, though still positive. Higher self-esteem and mastery are predictive of higher income, and workers with a more external locus of control have lower income.

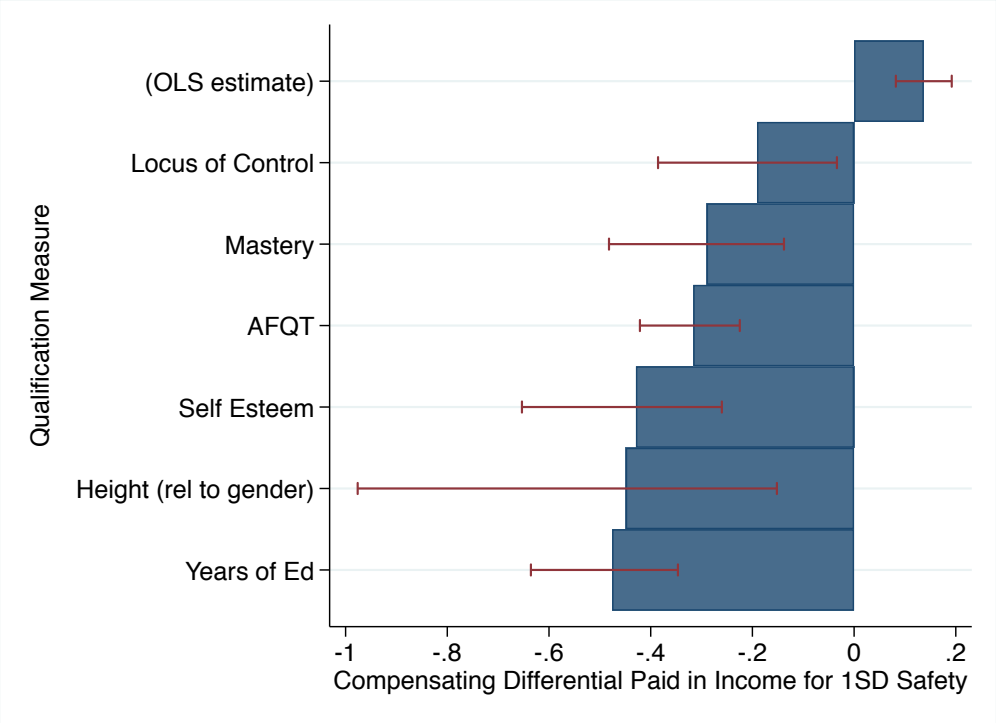
To reinforce the point that not any measure of skill will be an unbiased shifter for the Rosen frontier, I also re-run the analysis using several skills measured in the NLSY that are intended to measure occupation-specific skills rather than those useful in a broad variety of occupations. For these occupation-specific skills, I turn to the section-specific scores of the Armed Services Vocational Aptitude Battery. The ASVAB is a battery of 10 tests that measure skills ranging from word knowledge to automobile repair skills (e.g., "what type of wrench is pictured?"). The logic against using a variable like automobile repair skills as the direction by which to normalize the ability channel is that it is likely to contain information about a worker's job preferences by way of specialization in job-specific skills. It thus violates the identification assumption because it is correlated with how workers split their compensation into wage and amenities.

5.3.2 Results

The main result of this section is presented in Figure 7, which compares estimates for the price of safety using each of the six proposed h variables. These variables span widely different spaces of physical, cognitive, and social differences that workers bring to the labor market. Yet, they all lead to very similar magnitudes of price estimates. All six confidence intervals overlap with each other, and none overlap with the OLS coefficient of income on safety.

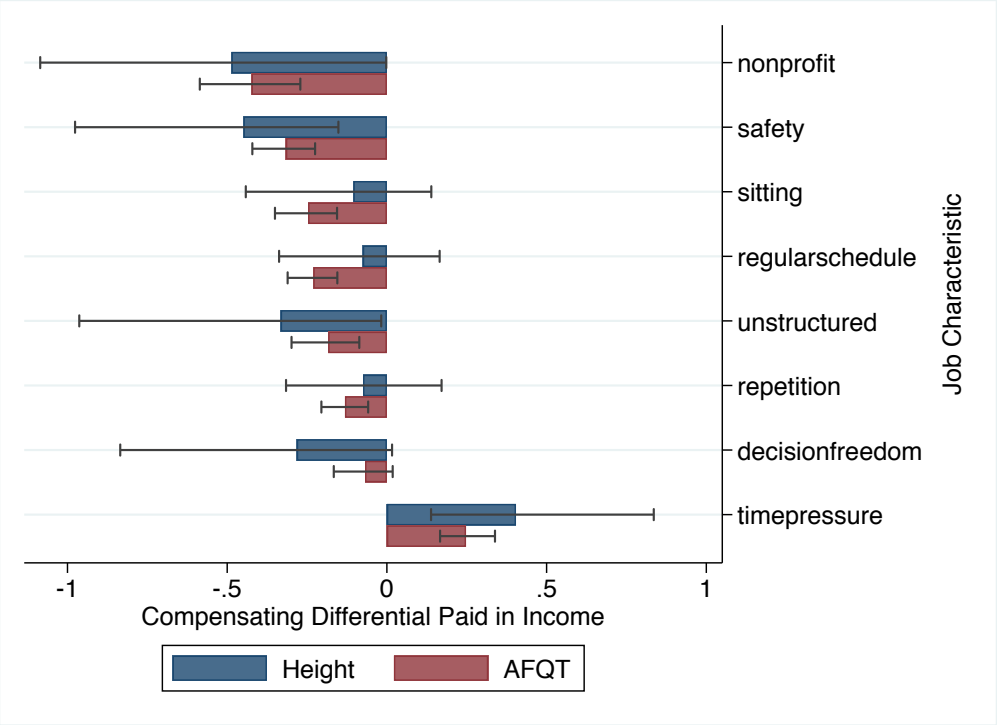
In contrast, Figure A.1 additionally includes the 10 component scores of the ASVAB, and thus plots 16 estimates using the proposed estimator. These tests are developed to measure relative skills, rather than the general ability channel that might proxy for total compensation. As predicted, not all of the ASVAB components even yield the same sign, with the ASVAB Auto & Shop score

Figure 7: Revisiting Safety: Comparison of h Variables



Notes: The horizontal axis represents the coefficient on the amenity in a regression of income on the one amenity controlling for $\hat{h}$, the predicted values from a regression of h on income and the amenity. Negative values indicate that workers on average lose income to acquire the amenity. All regressions are on the NLSY sample and run at the occupation level, weighted by number of workers. Anderson-Rubin 95% confidence intervals shown at the occupation level. Additional sample details are found in Section 5.1.

Figure 8: Detailed Comparison of Height vs AFQT



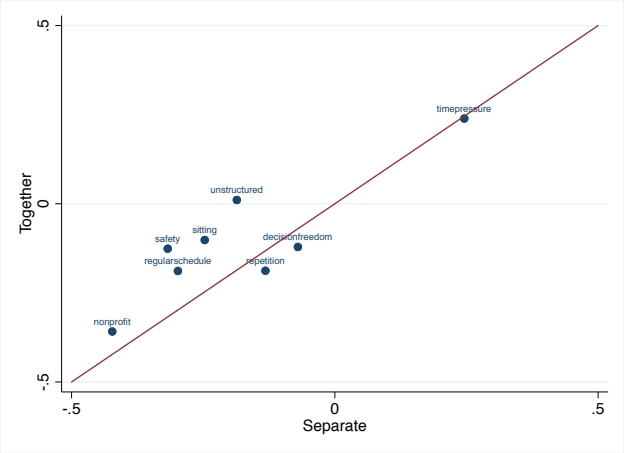
Notes: This graph compares the results when the h variable used is AFQT versus height residualized by sex. The correlation between residualized height and AFQT in the NLSY is 0.1504 at the individual level and 0.3974 at the occupation level. The horizontal axis represents the coefficient on the amenity in a regression of income on the one amenity controlling for mean h in the NLSY sample at the occupation level. Additional sample details are found in Section 5.1.

yielding an estimate even more extreme than the OLS estimate. Such a finding is to be expected given that this skill measures aptitude for a mechanic job, which is one of the most dangerous occupations by the O*NET measure. On the other end of the estimates is "coding speed," a test which provides a table of words and numbers and asks which number corresponds to a certain word. Dropped from the ASVAB in 2002, the score was previously used to classify military recruits into clerical jobs (Held and Carretta, 2013). Clerical jobs tend to be relatively safe.

Among several skills that are plausibly relevant to a general set of occupations, I obtain similar price estimates for safety. Such proxies include the six pre-specified h variables, and ex-post, even some ASVAB component scores (e.g., math) also yield similar results. On the other hand, if the ability channel is normalized to a skill that is designed to measure aptitude in a very narrowly defined set of occupations (e.g., car repair skills), I obtain a very different estimate. This is a normalization that the researcher must think critically about.

To examine the extent to which different h variables produce similar results for a wider range of amenities, Figure 8 zooms in on two leading candidate h variables: height and AFQT. Estimates for several amenities, priced separately, are presented to investigate whether the two candidate h variables, which share a correlation of only .1486 at the individual level, produce similar estimates. Although height is much less precise, there is no evidence to indicate that these variables are leading to different estimates. The empirical analysis of this section is consistent with my claim that any of several observed h variables can serve as proxies for the factors that determine a worker's level

Figure 9: Separate & Multivariate Amenity Prices



Notes: The scatterplot compares linear price estimates when amenities are priced separately (on the horizontal axis) versus in a multivariate framework (on the vertical axis). The h variable is AFQT. The red line is the 45° line. Additional sample details are found in Section 5.1.

of total compensation.

5.4 Multivariate prices

For simplicity, all of the empirics so far have priced amenities against income in a bivariate nature. However, the estimator readily extends to a multivariate pricing equation by including the additional amenities in both the first stage (generating predicted values) and in the second stage (calculating prices). Mechanically, this is equivalent to the ratio-of-coefficients approach in which h is regressed on all forms of income and amenity compensation. The negated ratio of each amenity coefficient over the income coefficient is equivalent to the corresponding price estimate from the predicted values approach.

Figure 9 compares linear price estimates when amenities are priced in isolation as opposed to in the multivariate framework. In general, the magnitudes of amenity prices shrink when they are priced jointly. That the multivariate coefficients are less than the bivariate ones can be seen in the scatterplot of Figure 9 as a slope of the points less than the red 45° line. This finding is consistent with the theory that good job attributes come bundled together. That is also what would be expected if income and amenities are viewed as two forms of compensation firms use to attract more productive workers.

6 Demographic Gaps

With estimates of the price of various amenities in the US labor market, I next turn to the question of quantifying the extent to which these amenities, at current prices, contribute to widely studied demographic pay gaps. This exercise is a decomposition of income inequality into an absolute component due to differences in offer curves and a relative component due to workers' locations across the income-amenity frontier. Before accounting for a wide array of amenities, it is instructive to first make the question concrete in the context of just one amenity.

6.1 Graphical Intuition: Regular Schedules

To begin with some graphical intuition, I focus on a variable describing how regular the job's schedule is. Although safety seemed to be the quintessential job amenity in the days of Adam Smith, this variable seems to more closely mirror modern scholarship surrounding amenities in the gender gap.

The NLSY asks workers whether they "usually work a regular daytime schedule or some other schedule" at their current job. Of those who do not work a regular schedule, the most common alternatives are either irregular schedules arranged by employers or regular schedules containing evening or night shifts. About 80% of workers report working a regular, daytime schedule. Significant differences exist by race and gender; women and White workers each work more regular schedules than their comparison groups. Admittedly, this variable may not be an ideal measure of the more nuanced job attributes relating to scheduling discussed in the literature on the gender gap, such as whether the nature of the work entails a convex return to hours as in Goldin (2014). However, if women tend to have have higher valuation for a regular daytime schedule – for instance, to be home for children in the evenings – then it may still be interesting to see how costly this noisy and narrowly defined amenity turns out to be in the data.

For the purposes of graphical exposition, let us assume that the true model is linear, and I am only interested in pricing this one amenity in relation to income. (That is, I will obtain the price of this amenity given how other unobserved amenities covary.) This price of the amenity was already reported in Figure 8: 1 SD of schedule regularity when AFQT is used as the h variable costs .29 SD of income.[21] This price estimate corresponds to the slope of the Rosen wage-amenity frontiers. That is, if I observe two workers who have different bundles of wage and amenity, the slope from the price estimate is enough to discern whether the workers lie on the same or different frontiers. I can thus decompose income inequality between men and women into a component that is due to sitting on different Rosen frontiers (differences in total compensation) versus differences in locations along the wage-amenity frontiers (differences in valuations).

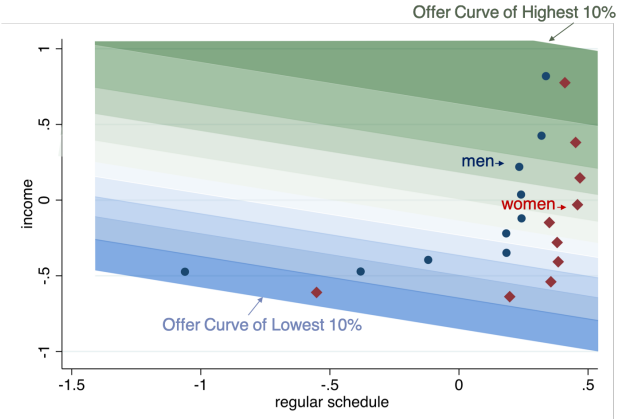
Panel A of Figure 10 shows the results of this analysis by binning the labor market into deciles of total compensation. There is a clear trend: at any level of compensation, men on average have higher earnings and less regular schedules than women. The difference is substantial; averaging over all the bins men earn about .08 SD more than women conditional on total compensation. To put the number in perspective the overall gender income gap in this sample is .54 SD. This simple result implies that $.08/.54 = 14\%$ of the gender income gap can be attributed to this trend of costly substitution by women to jobs that, roughly measured, have regular schedules (and other features that go with it).

Panel B shows the identical analysis, but grouped by race rather than by gender. The pattern is vastly different. Regardless of the well-known income differences between racial groups, their mean locations are essentially indistinguishable once total compensation is conditioned on. On average, a Black and a White worker who face the same job prospects do not make meaningfully different choices in the same way that men and women do. At least, Black workers and White workers do not meaningfully sort along the dimension measured by how regular the schedule of the job is.

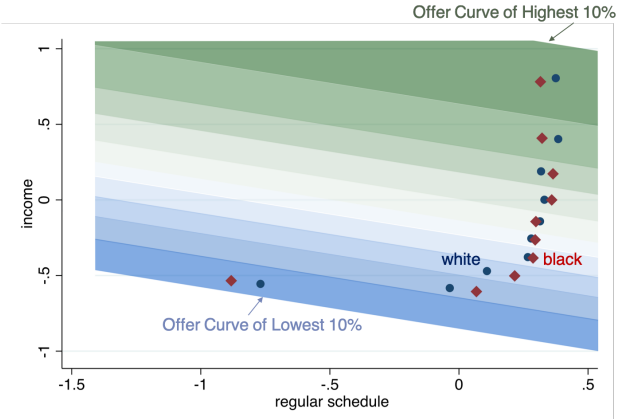
[21] In the previous section, I priced all amenities at the occupation-level because that was the level at which many were observed. This amenity is observed at the individual-level. Qualitatively, the result of this analysis are similar whether the amenity is priced at the occupation- or individual-level, although the price is smaller and less significant when this amenity is priced at the individual-level. This difference is consistent with the binary amenity measure containing some measurement error at the individual-level, causing attenuation, which is mitigated when the analysis is aggregated to the occupation level.

Figure 10: Valuing Regular Schedules

(a) Differences by Gender

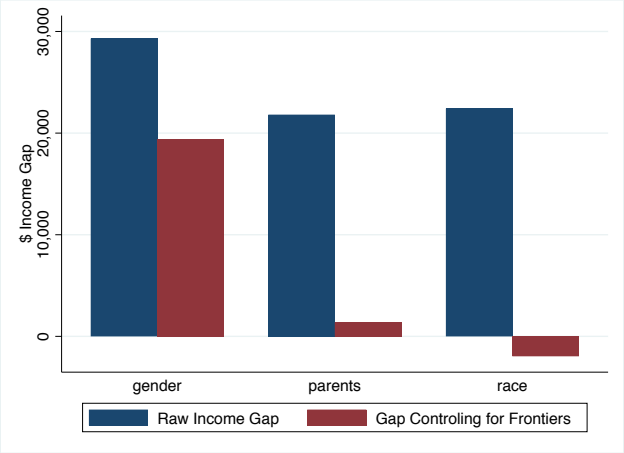


(b) Differences by Race



Notes: Both panels depict superimposed gradients of slope -.29, corresponding to the linear price estimate of this amenity when AFQT is used as the h variable; both axes are in standard deviations. The widths of each of the 10 bandwidths are not drawn to scale, but loosely corresponds to the region of the graph in which workers who obtain a particular decile of compensation would fall. The exact bandwidths used for calculations were calculated in the following way. First, AFQT was regressed on income and regular schedule at the occupation level weighted by number of workers; then, at the individual level, predicted values were calculated. The predicted values were grouped into 12 quantiles, and the top and bottom quantiles were dropped. This left 10 quantiles of predicted values, which correspond to (winsorized) quantiles of total compensation. The data points were constructed in the following way using the exact quantiles. For Panel A, I plot the average income and amenity values for men and women, separately, by decile of total compensation. Mens' locations are summarized by blue dots, and women's locations by red diamonds. Panel B is constructed in the exact same way, but separately by race rather than by gender. Red diamonds summarize the locations of "Black" respondents, and blue dots refer to respondents who identify as "non-Black, non-Hispanic." The data underlying the figure is the main NLSY analysis sample; additional sample details are found in Section 5.1.

Figure 11: Income Gaps vs. Gaps in Frontiers



Notes: The height of the blue bar is the unadjusted income gap among all employed respondents. Data by gender (men minus women) and race (non-Black/non-Hispanic minus Black) are from the NLSY in 2012, and data by parents (above minus below median parent income) are from PSID in 2013. The red bar is the coefficient on the demographic group in a regression of income on the group with a linear control for the multi-dimensional Rosen frontier. The frontier prices the following amenities: nonprofit (NLSY only), sitting, unstructured, repetition, decision freedom, regular schedule (NLSY only), safety, and time pressure. In the NLSY, AFQT is used as the h variable; for the PSID results by parent background, the h variable is education. Additional sample details are found in Section 5.1.

6.2 Pay Gap Accounting

To get a more holistic picture of the contribution of job amenities to the demographic pay gaps, I next consider a multivariate pricing equation of, in addition to regular schedule, the following job characteristics: nonprofit, sitting, unstructured, repetition, decision freedom, safety, and time pressure. These are the same characteristics priced separately and together in Figure 9. Two of the job characteristics (nonprofit and regular schedule) are not available in the PSID, so are not included for the analysis of inequality by parent background. For the analysis in the NLSY, I use AFQT as the h variable; as this variable is not available in the PSID, I use years of education to price amenities among that group. I re-price amenities in the PSID because after restricting the PSID to respondents for whom I observe parental earnings, I end up with a much younger sample than that of the NLSY; the samples thus may not be representative of each other.

Effectively, the many dimensions of job heterogeneity allow me to construct multi-dimensional Rosen frontiers. If the result of the previous section was that men tend to sit on the income side of the income-regular schedule frontier, the question I now ask is which side of a high-dimensional income-amenity contour men sit on.

As usual, I define income gaps by sex (male minus female) and race (non-Black/non-Hispanic minus Black). To operationalize inequality by parent income in binary terms, I define an income gap between those in the PSID based on whether one's parents earned below or above the in-sample median. Figure 11 displays the results of this analysis, as well as the overall demographic gaps in the relevant sample. I first start in the main analysis sample, which has no full-time hours restriction, and then examine robustness of the results to alternative specifications. As reported in Table 2, in the NLSY analysis sample men earn on average almost $30,000 more than women, and

Table 2: Income Gaps

	(1)	(2)	(3)	(4)	(5)	(6)
Male	29296.4	19396.0				
White			22407.1	-1959.1		
Above-Median Parents					21813.6	1409.2
Constant	41250.9		41992.6		39988.8	
Frontier Control		X		X		X
N	4868	4868	3977	3977	929	929

Notes: All columns report linear regressions in which the dependent variable is income in 2012 (or 2013, for the parental sample in columns 5 and 6). The sample is the main NLSY analysis sample with no hours restrictions for columns 1-4, and the PSID analysis sample for columns 5 and 6. Additional sample details are found in Section 5.1. Standard errors are not included for this table due to the difficulty of calculating them for the specifications controlling for the frontier. The controls for Rosen frontiers in columns 2, 4, and 6 are constructed as described in Section 6.2, in which the frontier prices the following amenities: nonprofit (NLSY only), sitting, unstructured, repetition, decision freedom, regular schedule (NLSY only), safety, and time pressure. In the NLSY, AFQT is used as the h variable; for the PSID results by parent background, the h variable is education.

the race gap is $22,000. The PSID reports a gap by parent income of $22,000.

Consistent with the graphical intuition of Figure 10, the race gap is not driven by differences in income-amenity substitution. Once offer curves are included as controls, there is essentially no income divide between Black and White workers. Similar can be said of the parent income gap. Absent a method to calculate proper confidence intervals for the results, a great deal of caution is warranted in reading too much into the point estimates. Still, if one were to interpret the directions of the point estimates as signal rather than noise, the story would be as follows. Although unconditionally Black workers earn lower income (which is not shown on the graph), conditional on an amount of total compensation, Black workers also earn slightly more income and worse amenities. That is, Black workers substitute toward income, and thus the racial pay gap exists in spite of Black workers' willingness to take on work with worse amenities for higher pay. The point estimate for the gap by parent income tells the opposite story, in which kids from poor families are more likely to substitute toward better amenities. Although quantitatively small and likely to be imprecise, the point estimate is consistent with the logic of Jencks and Tach (2005), who noted that if preferences for work characteristics are passed across generations, some part of the intergenerational income gradient is likely due to amenity substitution. Across alternative specifications (Appendix H, discussed more shortly) the directions in which amenities point for both the race and parent background gap seems to be relatively stable.

In contrast to the nature of inequality by race and parent background, which seem better attributed to general-compensation factors such as skills or discrimination than to income-amenity substitution, the bulk of the gender gap still exists within observable Rosen frontiers. These findings are directionally similar to the analysis using only the regular schedule amenity, which found 14% of the gender gap due to the amenity, but are quantitatively much larger. The analysis with more amenities indicates that approximately two-thirds of the gender income gap in this dataset is due to variation along the wage-amenity frontier. Whether this result is large is a matter of perspective. On the one hand, if the remaining one-third of the gender gap were due to discrimination in the labor market, this would certainly be cause for concern. But on the other hand, given the limited

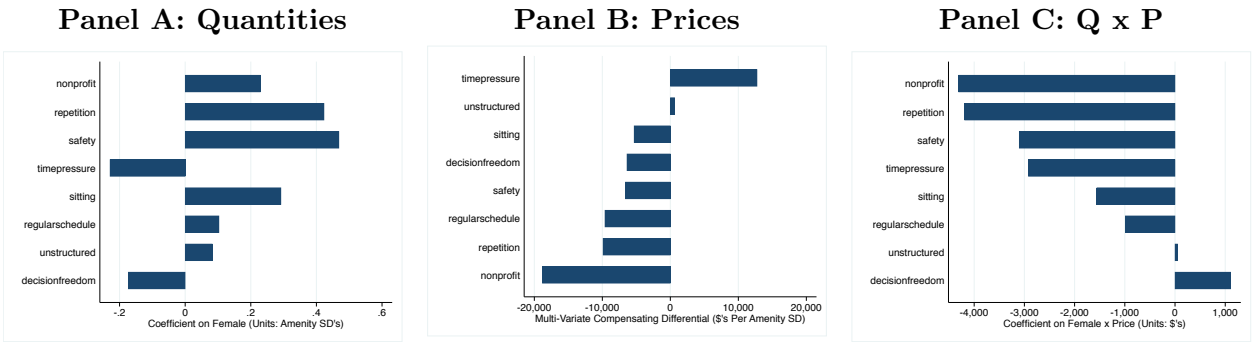
amount of information observable about the job, it is possible that if further amenities could be priced, even more of the gender pay gap might be explained.

At the very least, the findings indicate that the gender pay gap is to a large extent a relative one, whereas those by race and parent background are absolute gaps in total compensation. The absolute differences in compensation may be due to differences in general-purpose skills such as education, which may themselves be due to discrimination, for instance along the lines of Neal and Johnson (1996). The absolute differences in job quality may also stem from discrimination in the labor market, as in search model of Black (1995). In contrast, much of the gender gap does not seem to be explained by such general channels. As I outline next, the gender gap appears to be best explained by the fact that the average man and average woman, in addition to having very different incomes, also have observable different jobs in other ways. And differences in job characteristics command prices in the labor market.

6.3 Amenity quantities and prices

Figure 12 provides intuition as to how one might decompose the result on the gender pay gap into prices and quantities of amenities.[22] The amenity quantities in Panel A are a simple but instructive way to think about the implications of this paper for the gender pay gap by describing the extent to which the jobs of men and women tend to differ among observable dimensions, without taking a stance on what types of jobs are "better" or "worse." For instance, although average income of women is lower than men, the average quantity of on-the-job safety is .47 SD higher among women.

Figure 12: Gender Pay Gap in Quantities and Prices



Notes: Panel A plots the coefficients from separate regressions of the amenity on an indicator for female in the NLSY-ONET sample. The ordering is in accordance with the result in Panel C. To construct Panel B, I first obtained predicted values from a regression of AFQT on income and all amenities shown at the occupation level, weighted by number of workers. The bars in Panel B correspond to the coefficients from a regression of income on these amenities controlling for predicted values. For this figure, yearly income is in dollars rather than standard deviations. Panel C plots the data in Panels A and B multiplied. Additional sample details are found in Section 5.1.

Panel B displays multi-variate price estimates for the same bundle of amenities shown in Panel A. Prices allow us to interpret the differences in quantities. In the example of safety, the amenity is valued by the labor market – a positive price, displayed here as a negative compensating differential. Holding fixed other observed job characteristics, workers on average lose $6,622 of income to obtain

[22] It is worth reiterating that the methodology does not solve issues relating to omitted amenities that are correlated with observed ones. While the results give credence to the idea that job features are important, the question of which job features matter is inherently subject to which job features the researcher includes. The goal of this paper is not to argue that a definitive set of job features has been found, but rather to propose a methodology that many researchers can apply with other sets of job features. This exercise should be viewed as instructive rather than exhaustive.

1 SD of safety. A simple back-of-the-envelope calculation thus suggests, from gender differences in safety alone, one would expect to see a pay gap of .47 x -$6,621 = -$3,103.[23]

Panel C systematically applies the same calculation to the rest of the amenities, and is ordered by the product of quantity differences times the price. Across many of the amenities studied, the finding is that gender differences of the amenity at current prices generate substantial pay gaps in favor of men. Notably, not all amenities included in this analysis follow this trend. For instance, men on average tend to have jobs with more freedom to make decisions, and decision freedom is a job attribute estimated to be costly. However, this countervailing effect is not large.

Further robustness checks in Appendix H show that the main findings are not sensitive to alternative specifications. Results are similar across multiple choices of the proxy for total compensation (h variable), when restricting the sample to full-time workers, when estimating amenity prices separately by gender, and when adding occupational prestige as an additional amenity. Pricing amenities at the individual rather than occupation level slightly amplifies the results. Across all specifications, the contribution of amenity substitution to the gender pay gap remains large, while its role in explaining race and parent background gaps remains limited.

6.4 Confirming prior results

This paper puts forward a new way to look at typical data on income, amenities, and observed skills. And yet, the findings of this section are in line with the broader literature.

A great deal of literature on gender inequality has increasingly moved toward understanding compensating differentials. A sweeping analysis of the gender pay gap by Goldin (2014) pointed a finger toward pay schedules that are non-linearly increasing in hours, particularly in male-dominated occupations. In a recent review of literature on the gender wage gap, Blau and Kahn (2017) concluded that in the current labor market, with the gender education gap now reversed, conventional human capital variables explain "little of the gender wage gap" while "gender differences in location in the labor market—distribution by occupation and industry—continued to be important in explaining the gap." These differences between where men and women tend sit in the labor market is a "factor long highlighted in research on the gender wage gap." Much recent work using stated and real choices has also drawn attention to the fact that men and women have different distributions of valuations for various workplace amenities, which, along with prices, drives my results. Recent work by Juhn and Rubinstein (2022) also highlights the role of job sorting between men and women in pay gaps. While their analysis treats AFQT as a control for ability rather than as an anti-instrument, their model and results yield additional insights about how the tendency for abstract jobs to require more work hours than routine jobs contributes to gender pay differentials through sorting.

Less work has sought to analyze the effect of amenity substitution in other demographic pay gaps. A notable exception is work by Jencks and Tach (2005) to systematically review available evidence on the sources of intergenerational income inequality. Some evidence does exist on the propensity for kids to take on similar roles in the workplace to their parents – for instance, the intergenerational correlation in hours worked documented by Altonji and Dunn (2000) and studies documenting trends in self-employment across generations. Jencks and Tach (2005) offer an educated guess that at least 10% of the intergenerational income correlation might be due to the effect of a broad class of values passed across generations, including hours worked, location, and choice of non-monetary work characteristics. In other words, they argue that part of the intergenerational income gap is simply due to the tendency of children from lower-income backgrounds to

[23] If the analysis had indicated that jobs with higher safety command a compensating differential of $0, then I would have been led to conclude that this amenity does not contribute to the gender pay gap. And if the price were of the opposite sign, then the conclusion would be that the amenity at current prices functions to close the gender pay gap.

choose lower-paying jobs with better amenities. My results are most consistent with this theory. A contrary theory, which I have not encountered in the literature, is that children from low-income families would substitute toward higher-paying jobs with worse amenities so as to compensate for lost resources of parent background. Under such a theory, the intergenerational income gap in income terms would be even larger if kids from poor backgrounds had the same preferences as those from richer ones; my results are less consistent with that theory.

Although the findings of this section are not entirely new, my procedures entails a significant deviation from the literature. The mounting scholarship on the role of job amenities in the gender pay gap stands in discord with a literature focused on estimating the prices of such amenities. The central finding of this literature, to echo the words of Bonhomme and Jolivet (2009), appears to be a "pervasive absence of compensating differentials." But these findings demonstrate that when more versatile econometric techniques are used to estimate compensating differentials, a key finding on the importance of job amenities in the gender pay gap in the current US labor market can also be obtained from widely available datasets. The econometric techniques can now be applied to datasets of other countries and other decades in order to better understand the role of compensating differentials in other income inequalities of interest.

7 Conclusion

This paper proposes a new method for estimating compensating differentials in the presence of unobserved ability, a limitation that has long complicated efforts to interpret observational wage-amenity trade-offs. While past work has warned of bias in cross-sectional regressions of wage on job characteristics, many widely cited papers continue to rely on this flawed approach. The method introduced here provides an alternative: by leveraging an "anti-instrument," or noisy proxy for the ability channel, it becomes possible to estimate amenity prices while holding ability constant—even when ability is never directly observed.

Applied to classic datasets in the compensating differentials literature, the estimator recovers wage-amenity trade-offs that are more consistent with both quasi-experimental evidence and economic intuition. The revised estimates imply that amenities account for a large share of the gender pay gap—on the order of two-thirds—but little of the gaps by race or by parent income. These findings help reconcile past debates about whether observational estimates are too attenuated to be useful, as argued by Hwang et al. (1992), and whether any meaningful signal about amenity prices could still be extracted. With improved estimates, it becomes possible to more confidently assess how the structure of jobs—and their costs—contribute to income inequality.

From a policy standpoint, the findings do not suggest that women are paid less *because* they value different job features, but rather that when those job choices interact with non-zero prices, substantial wage gaps emerge. Efforts to make those amenities more accessible—through regulation or institutional reform—could reduce inequality without suppressing individual choice. The potential is especially large in areas already governed by public policy, including workplace safety, flexibility, and scheduling.

Although the empirical exercises in this paper are modest in scope, the estimation strategy is simple and broadly applicable. Already, researchers have begun adapting the approach to study flexibility in Dutch labor markets, meaningful work in Sweden, and occupational safety across Europe. With better individual-level data and richer measures of job attributes, the estimator can be used to explore new questions: how amenity prices changed during the COVID-19 shift to remote work; how valuations evolve over the lifecycle; how patterns differ across countries and policy environments. My hope is that this paper will support a shift in empirical practice—away

from regressions that confound preferences and productivity, and toward estimators that better isolate the price of work itself.

References

Altonji, J. G. and T. A. Dunn (2000): "An Intergenerational Model of Wages, Hours, and Earnings," *The Journal of Human Resources*, 35, 221.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113, 151–184.

Andrews, I., J. H. Stock, and L. Sun (2019): "Weak Instruments in Instrumental Variables Regression: Theory and Practice," SSRN Scholarly Paper ID 3445885, Social Science Research Network, Rochester, NY.

Angrist, J. D. and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press, 1 edition ed.

Ashenfelter, O. and M. Greenstone (2004): "Using Mandated Speed Limits to Measure the Value of a Statistical Life," *Journal of Political Economy*, 112.

Bell, A., S. B. Billings, S. Calder-Wang, and S. Zhong (2025a): "An Anti-IV Approach for Pricing Residential Amenities: Applications to Flood Risk *," .

Bell, A., S. Calder-Wang, and S. Zhong (2025b): "Measuring Housing Quality Using Revealed Preference: A Geographic PageRank Approach," .

Black, D. A. (1995): "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics*, 13, 309–33.

Blau, F. D. and L. M. Kahn (2017): "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 55, 789–865.

Bonhomme, S. and G. Jolivet (2009): "The pervasive absence of compensating differentials," *Journal of Applied Econometrics*, 24, 763–795.

Brown, C. (1980): "Equalizing Differences in the Labor Market," *The Quarterly Journal of Economics*, 94, 113–134.

Burbano, V. C., O. Folke, S. Meier, and J. Rickne (2023): "The Gender Gap in Meaningful Work," *Management Science*, publisher: INFORMS.

Case, A. and C. Paxson (2008): "Stature and status: Height, ability, and labor market outcomes," *The journal of political economy*, 116, 499–532.

Combes, P.-P., G. Duranton, and L. Gobillon (2021): "The Production Function for Housing: Evidence from France," *Journal of Political Economy*, 129, 2766–2816, publisher: The University of Chicago Press.

De Schouwer, T., T. Deneus, and M. Forti (2025): "How to Make Work Meaningful?" .

De Schouwer, T. and I. Kesternich (2024): "Work Meaning and the Flexibility Puzzle | Publications | CESifo," .

Deming, D. J. (2017): "The Growing Importance of Social Skills in the Labor Market," *The Quarterly Journal of Economics*, 132, 1593–1640.

Denning, J. T., B. Jacob, L. Lefgren, and C. v. Lehn (2019): "The Return to Hours Worked Within and Across Occupations: Implications for the Gender Wage Gap," *National Bureau of Economic Research Working Paper Series*.

Diegert, P., M. A. Masten, and A. Poirier (2023): "Assessing Omitted Variable Bias when the Controls are Endogenous," ArXiv:2206.02303 [econ].

Ferreira, S., S. Martinez-de Morentin, and A. Erro-Garces (2024): "Measuring Job Risks When Hedonic Wage Models Do Not Do the Job," .

Folke, O. and J. Rickne (2022): "Sexual Harassment and Gender Inequality in the Labor Market*," *The Quarterly Journal of Economics*, 137, 2163–2212.

Goldin, C. (2014): "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 104, 1091–1119.

Greenstone, M., S. P. Ryan, and M. Yankovich (2015): "The Value of Statistical Life: Evidence from Military Retention Incentives and Occupation-Specific Mortality," .

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24, 411–482.

Held, J. D. and T. R. Carretta (2013): "Evaluation of Tests of Processing Speed, Spatial Ability, and Working Memory for use in Military Occupational Classification," Tech. Rep. NPRST-TR-14-1, Navy Personnel Research, Studies, & Technology Bureau of Naval Personnel.

Hwang, H., W. R. Reed, and C. Hubbard (1992): "Compensating Wage Differentials and Unobserved Productivity," *Journal of Political Economy*, 100, 835–858.

Jencks, C. and L. Tach (2005): "Would Equal Opportunity Mean More Mobility?" SSRN Scholarly Paper ID 779507, Social Science Research Network, Rochester, NY.

Juhn, C. and Y. Rubinstein (2022): "Task Requirements, Time Requirements, and the Gender Gap in Jobs and Pay," Tech. rep.

Lamadon, T., M. Mogstad, and B. Setzler (2022): "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, 112, 169–212.

Lavetti, K. (2018): "The Estimation of Compensating Wage Differentials: Lessons From the Deadliest Catch," *Journal of Business & Economic Statistics*, 0, 1–18.

Lucas, R. E. B. (1977): "Hedonic Wage Equations and Psychic Wages in the Returns to Schooling," *The American Economic Review*, 67, 549–558.

Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2023): "The Value of Working Conditions in the United States and Implications for the Structure of Wages," *American Economic Review*, 113, 2007–2047.

Maideu-Morera, G. (2024): "Amenity-Biased Technical Change," .

Mas, A. (2025): "Non-Wage Amenities," .

Mas, A. and A. Pallais (2017): "Valuing Alternative Work Arrangements," *American Economic Review*, 107, 3722–3759.

——— (2019): "Labor Supply and the Value of Non-work Time: Experimental Estimates from the Field," *American Economic Review: Insights*, 1, 111–126.

McFadden, D. L. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in *FRONTIERS IN ECONOMETRICS*, New York: Academic Press:, 105–142.

Mortensen, D. (2003): *Wage dispersion: why are similar workers paid differently*, Zeuthen lecture book series, Cambridge, Mass: MIT Press.

Neal, D. A. and W. R. Johnson (1996): "The role of premarket factors in black-white wage differences," *Journal of Political Economy*, 104, 869.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Oster, E. (2019): "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 37, 187–204.

Pearlin, L. I., E. G. Menaghan, M. A. Lieberman, and J. T. Mullan (1981): "The Stress Process," *Journal of Health and Social Behavior*, 22, 337–356.

Persico, N., A. Postlewaite, and D. Silverman (2004): "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height," *Journal of Political Economy*, 112, 1019–1053.

Rosen, S. (1986): "Chapter 12 The theory of equalizing differences," in *Handbook of Labor Economics*, Elsevier, vol. 1, 641–692.

Rosenberg, M. (1965): *Society and the Adolescent Self-image*, Princeton University Press, google-Books-ID: mZAzQQAACAAJ.

Rotter, J. B. (1966): "Generalized expectancies for internal versus external control of reinforcement," *Psychological Monographs: General and Applied*, 80, 1–28.

Smith, A. (1776): *The Wealth of Nations*.

Sorkin, I. (2018): "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133, 1331–1393.

Stern, S. (2004): "Do Scientists Pay to Be Scientists?" *Management Science*, 50, 835–853.

Taber, C. and R. Vejlin (2020): "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 88, 1031–1069, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14441.

Wiswall, M. and B. Zafar (2018): "Preference for the Workplace, Investment in Human Capital, and Gender," *The Quarterly Journal of Economics*, 133, 457–507.

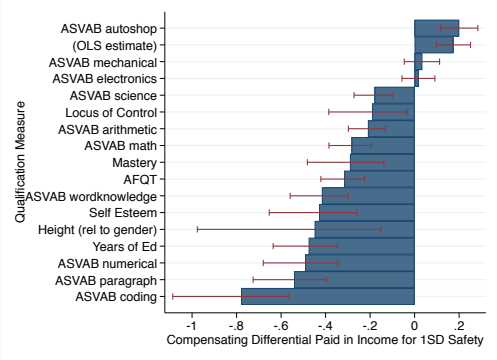
For Online Publication: Appendix to

"Job Amenities & Earnings Inequality"

by Alex Bell

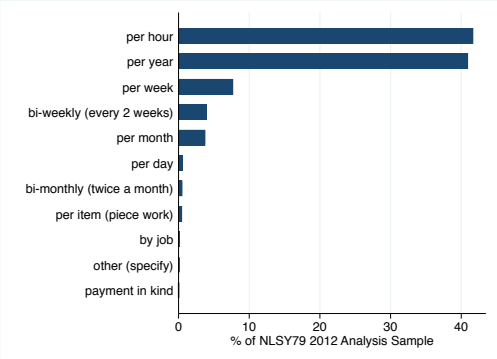
A Appendix Figures and Tables

Figure A.1: The Price of Safety: Comparison of h Variables, including ASVAB Components



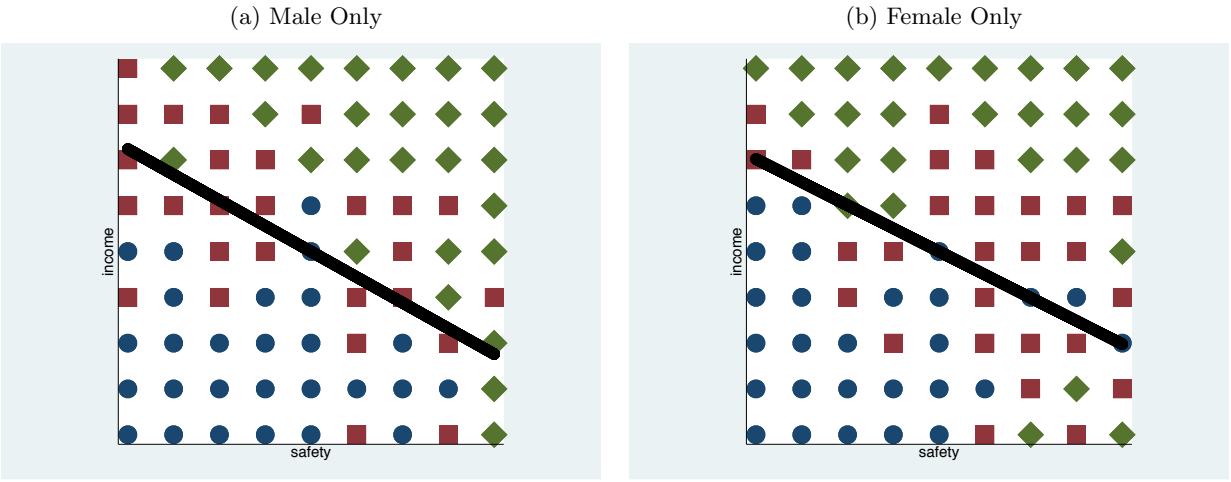
Notes: In both panels, he horizontal axis represents the coefficient on the amenity in a regression of income on the one amenity controlling for $h\hat{}$, the predicted values from a regression of h on income and amenity. All regressions are on the NLSY sample and run at the occupation level, weighted by number of workers. Anderson-Rubin 95% confidence intervals shown at the occupation level.

Figure A.2: Rate of Pay



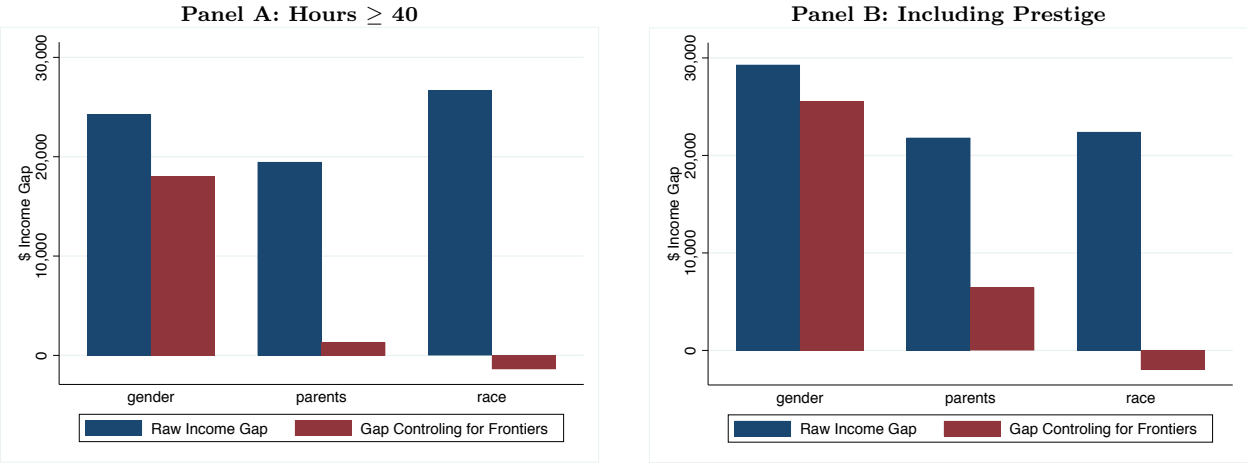
Notes: The histogram plots the rate of pay of the 4,899 individuals in the primary NLSY analysis sample. Asked as part of the yearly Employer Supplement, respondents are prompted "what is the easiest way for you to report your total earnings before taxes or other deductions: hourly, weekly, annually, or on some other basis?" and if clarification is requested, the researcher will add that this information is necessary to compare the amount that people earn in different jobs. After responding to this question, the respondent will also be asked to estimate the amount earned in that timeperiod (per hour, year, etc.). If the respondent is a teacher, a prompt will also be made that earnings should be reported only over the number of months for which the respondent was paid for that job, which is typically 9 or 10 months rather than per year.

Figure A.3: Non-Parametric Safety Estimates by Sex



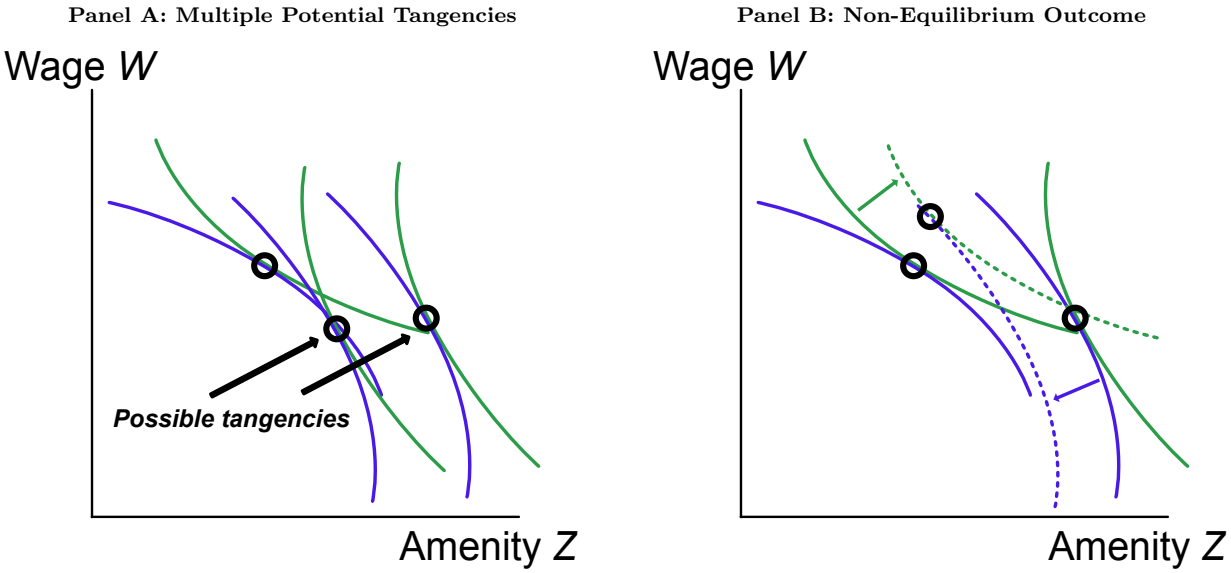
Notes: Both panels are constructed in an identical way to Figure 5, except that they split the sample on sex.

Figure A.4: Income Gaps vs. Gaps in Frontiers: Hours $\geq$ 40 and with Prestige



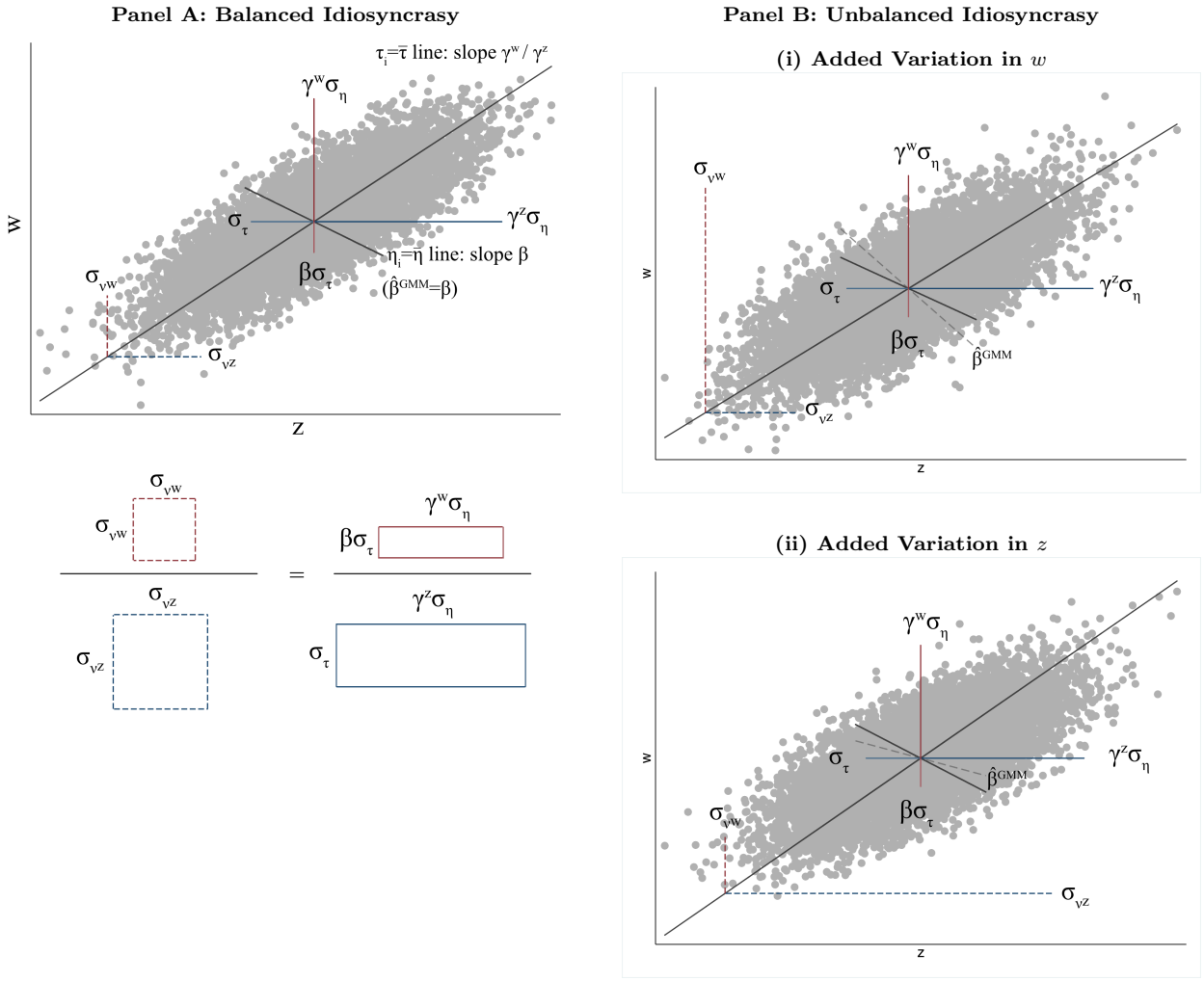
Notes: Panel A reproduces Figure 11 restricting to workers who report at least 40 hours of work per week. Panel B reproduces the same analysis but includes the additional amenity of prestige.

Figure A.5: Indifference Curves are Tangent to the Equilibrium Frontier



Notes: This figure illustrates that if the frontier of jobs chosen does *not* lie tangent to each worker's indifference curve at the point of the job, then the equilibrium is not stable. Panel A depicts two possible jobs; Panel B shows that choosing the non-tangent job leads to an off-equilibrium blocking situation.

Figure A.6: Balanced and Unbalanced Idiosyncrasy



Notes: Panel A shows balanced measurement error across wage and amenity dimensions, such that the true price is identified. Panel B shows unbalanced cases: in (i), excess idiosyncrasy in wage determination biases the estimate upward; in (ii), excess idiosyncrasy in amenity determination biases it downward. See Equation 12 and Section E for further discussion.

Table A.1: Construction of Amenity Measures

Variable Name	**Question**	**Operationalization**
regular schedule	Do you "usually work a regular daytime schedule or some other schedule?"	regular daytime
non-profit	Are you "employed by government, by a private company, or a non-profit organization" or self-employed?	non-profit
safety	How often does this job require exposure to minor burns, cuts, bites, or stings?	1 – "every day"
repetition	How important is repeating the same physical or mental activities (e.g., key entry) over and over?	"extremely important"
decision freedom	How much decision-making freedom, without supervision, does the job offer?	"a lot"
sitting	How much does this job require sitting?	"continually or almost continually"
unstructured	To what extent is the job structured versus allowing the worker to set tasks and goals?	"a lot of freedom"
time pressure	How often does this job require meeting strict deadlines?	"every day"

Table A.2: OLS: Individual- vs. Occupation-Level

	OLS	OLS	Proposed
	Outcome: Income SD's		
Safety SD's	.078**	-0.032	-.32***
	-0.026	-0.027	-0.062
AFQT	.014***	.030***	.072***
	-0.0017	-0.0039	-0.0071
Constant	-.51***	-1.20***	-2.98***
	-0.047	-0.15	-0.32
Level of Aggregation	Individual (None)	Occupation	Occupation

Notes: Standard errors in parentheses clustered at the occupation level. Sample is 4,899 workers in 397 occupations. Standard errors in the final column calculated by bootstrap. Results in Column 2 are identical if the outcome is maintained at the individual level, since none of the regressors vary within-occupation. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

B Economics of Compensating Differentials

My conceptual framework follows at a high level the model of Rosen (1986) in the Handbook of Labor Economics. To emphasize the problem of ability bias, I introduce the standard continuous Rosen model first in terms of workers that are homogenous in their ability (but differ in preferences). I then introduce ability as the second level of exogenous heterogeneity.[24]

B.1 Relative Taste Heterogeneity

To start, workers are equally productive and differ only in how their compensation is split into various forms. To illustrate the model, I will assume compensation is split between wage w and only a single amenity z, but the framework generalizes to substitution across many amenities.[25] Index worker i's valuation of the amenity by θ_i and firm j's cost of providing it by Ω_j. Worker utility is given by $u(w_i, z_i|\theta_i)$ and firm profits by $\pi(w_j, z_j|\Omega_j)$. The equilibrium price will be determined by the distributions of θ and Ω.

While the literature typically considers θ_i to be more or less "preference" variation, this may not always be the most accurate wording. To see why, consider a hypothetical case of a father and mother who might be able to obtain the same set of jobs as far as employers are concerned, but the woman chooses a job that comes with amenities that make it easier for her to also provide childcare. In this light, the differing choice seems less motivated by preferences, and if anything motivated more by a societal constraint (i.e., that women carry an added burden to provide childcare)

An alternative implication, at least at face value, seems more controversial: If women and men had similar valuations of certain job attributes, then there would be less income inequality by gender. A nuance of this conclusion is that it is still open for debate the extent to which differing valuations of job characteristics stem from preferences or from constraints outside of the labor market. For instance, women may face constraints relating to childcare. Men may face social norms and expectations to abstain from parental leave. Individual workers may be unable to remove such societal constraints.

Market clearing is defined as a matching of workers to firms and a wage and amenity paid out in each job. Efficiency dictates workers with the highest valuation of the amenity are matched to firms with the lowest cost of producing it: perfect assortative matching on θ_i and Ω_j. Each worker-firm pair determines a level of income and amenity that is Pareto optimal, which means that the (w, z) coordinate chosen will be at a tangency of the worker and firm indifference sets, specified by u and π. Famously, all workers' and firms' indifference sets of u and π are "kissing" in equilibrium. For the matching of workers and firms to be stable, the (w, z) coordinate chosen by a particular pair will also be one that is less desirable to neighboring workers and firms with similar valuations, relative to what those neighbors have chosen.[26] With a sufficiently continuous distribution of θ and Ω, the relative location choice of each pair is unique and constitutes the unique equilibrium within the market.[27]

[24] As a more realistic (but substantially more complex) alternative to this two-step model, Han and Yamaguchi (2015) derive conditions under which a stable worker-firm match arise when markets form under both dimensions of heterogeneity simultaneously. A simulation of a similar simultaneous matching process is also included in Appendix D of Bell et al. (2025).

[25] Although wage refers to an hourly concept, the logic of the model would also generalize to different earnings measures, such as yearly earnings among a group of full-time salaried workers. The model is intended to explain the worker's labor supply decision between different types of jobs, rather than of how many hours to supply.

[26] Figure A.5 illustrates that if the frontier deviates from this condition, it is not a stable equilibrium.

[27] While the location choice of all pairs *relative* to each other is unique, there is a level parameter that is not pinned down by this model. It can be thought of as homogenous bargaining weights that determine the way productive outputs are split between workers and firms. I take that process as exogenous.

From the perspective of a single participant who takes other participants' choices as given, the convex combinations of w and z that have been agreed upon function as a budget set. The local slope of this set – how much wage must be given up to obtain a small amount more amenity – is the local price of the amenity in wage units. A single worker's problem can be rewritten as maximizing $u(w_i, z_i|\theta_i)$ subject to the non-linear budget constraint that w_i and z_i are points along a particular frontier. All workers at this point must lie on the same frontier, because productivity does not yet differ. Firms, too, maximize profits subject to the feasible set of options given by this frontier. This is a partial-equilibrium interpretation of the frontier, but if many workers or firms changed their choices, the shape of the frontier could change in the new equilibrium. The first two panels of Figure 3 illustrate the relationship between the slope of the frontier and the distributions of preferences and firm costs that have contributed to its shape.

B.2 Absolute Heterogeneity

Finally, I extend the model to incorporate the fact that some jobs afford better compensation than others. To incorporate this feature, I model the amount of total compensation the worker can obtain as heterogeneous and determined exogenously, perhaps by factors determined prior to labor market entrance such as intelligence, motivation, and the unobserved facets of job search. Firms differ in their exogenous staffing needs and pay out more total compensation, composed of wages and amenities, in accordance with what type of applicant they seek to attract.[28] The multiple Rosen frontiers are shown in Panel C of Figure 3.

The critical modification is that total compensation, which I have denoted by η, is no longer the same for all workers.[29] Viewing η as pre-determined, the equilibrium can still be found by sub-dividing the entire labor market into sub-markets indexed by η. Markets indexed by the highest η tend to contain the workers with the most productive abilities, matched with firms that most value such human capital. Within each η-level sub-market, wage and amenity are determined in accordance with the Rosen framework in which all jobs entail an η amount of total compensation, but in different forms. The amenity price is now local to the sub-market. For a sub-market of compensation level η, the price is governed by a frontier that I will denote as a function:

$$\psi(w, z|\eta) = \begin{cases} 1 & w \text{ and } z \text{ entail total compensation valued at } \eta \\ 0 & \text{otherwise} \end{cases}$$

The ψ function is strictly monotone in both arguments, such that the frontiers of different compensation levels do not intersect each other.

B.3 Stochasticity of the Offer Curve

Decompose the determinants of worker i's compensation η_i into an observable part h_i and an unobservable part ε_i. The unobserved determinants of compensation in ε can include unmeasured skills, but also random search and matching frictions which result in observably identical workers obtaining better or worse jobs.

[28] A more complex model of the labor market would allow firms to have some elasticity of substitution between skilled and unskilled labor. However, the goal of the present paper is not to understand the factors determining why higher-skilled workers earn more than lower-skilled workers, but rather the income-amenity choices faced by workers of a similar skill level. Thus, I take the skill premium as determined outside of the model.

[29] Differences in η can capture not only productivity, but also to some degree due to luck or search frictions in the Mortensen (2003) sense.

To the best of my knowledge, all prior work on estimating market prices of amenities has made the assumption that wages are stochastic (caused by a variable like ε) but amenities are not.[30] To the extent that search frictions exist in the labor market (Goldberger, 1984; Mortensen, 2003), it does not seem plausible that they would affect wages but not amenities. A key deviation of my model from the literature is that I treat total compensation – and thus both wage and amenity – as stochastic.

This assumption relates to what can be learned from conditioning on workers' wages and amenities. In the model consistent with the OLS approach, there is agnosticism about whether two workers with the same wage and amenities actually share the same latent offer of total compensation for which we seek to control. (This is because at a given level of amenities and offer sets, latent variables are still assumed to influence wages.) In my model, if two workers obtain the same wage and the same amenities, I view them as necessarily sharing the same latent offer set. Therefore, a new assumption that I introduce into the empirical literature on job choice is that each value of wage and amenity should be consistent with only one level of total compensation. In the proof of Section 3, this translates to an assumption that η can bet written as a function of wage and amenities.

B.4 Equilibrium Exclusion Restriction

Among workers of a total compensation level η, h must be independent of worker and firm tastes, θ and Ω. Formally:

$$(h \perp \theta)|\eta \text{ and } (h \perp \Omega)|\eta \tag{4}$$

There is some redundancy here in that the condition on the relationship with θ is really the same as the one on the relationship with Ω. Under the sorting framework, at job k, $\theta_k = \Omega_k$ so $(h \perp \theta)|\eta$ is interchangeable with $(h \perp \Omega)|\eta$. However, I state both conditions to emphasize that one can evaluate the exclusion restriction from either the firm's perspective or the worker's perspective – and an argument in violation of either is sufficient to rule out a candidate h variable.[31]

Finally, the parameters θ or Ω in Condition 4 can be replaced with the joint distribution of w and z to generate a more reduced-form statement of the exclusion restriction.

$$((w, z) \perp h)|\eta \tag{5}$$

Identification relies on a variable that is informative about how high of a compensation frontier the worker lies on (η), but adds no $further$ information about the location of the worker along that frontier. Put another way, h would be irrelevant to w and z if the researcher could observe and condition on η.

B.5 No "bad controls"

To see why my assumptions are weaker than those of the typical OLS approach, note that the typical approach requires the researcher to select control variables for ability. I will show that

[30] When wage is regressed on amenities and proxies of the offer set, the regression typically has an R^2 far less than 1 (Hwang et al. 1992). This can only be explained by measurement error in wage or a structural error term ε. For the parameters of this equation to be identified by OLS, it must be the case that the structural error causes wage, but is unrelated to amenities.

[31] From the worker side, within the η-level workers, those that have high h can't have different valuations of the amenity (θ) from those that have low h. From the firm side, among firms that pay η-level compensation, those that employ high-h workers can't have different valuations of the amenity (Ω) from those that employ low-h workers.

any ability control that satisfies OLS exogeneity when regressing wage on amenities also satisfies Condition 5.

The typical amenity pricing approach seeks to estimate the joint distribution of w and z conditional on η, which I denote as $(w,z)|\eta$. For instance, this may be linearly approximated by a regression of w on z with controls. A control is a bad control if it can cause bias, e.g. if adding it to an already-identified model would cause identification of the structural parameters to be lost.[32] Defining $(w,z)|\eta$ as the identified distribution of interest, then, a *good* control would have the property that adding it to the conditioning does not change the w-z relationship of interest:

$$(w,z)|\eta,h = (w,z)|\eta$$

Rewrite the "good control" condition in terms of PDF's:

$$f_{wz|\eta,h} = f_{wz|\eta}$$

Re-writing the left-hand side:

$$\frac{f_{wzh|\eta}}{f_{h|\eta}} = f_{wz|\eta}$$

Re-arranging,

$$f_{wzh|\eta} = f_{wz|\eta} f_{h|\eta}$$

By definition of independence, I arrive back at Condition 5:

$$((w,z) \perp h)|\eta \tag{6}$$

In the typical analysis of wage on amenity with ability controls, Condition 5 is necessary to include h as a control variable. The only difference is that with the standard regression approach, I would also require the rest of the full set of ability controls. Therefore, the proposed estimator entails a weakening of the assumptions required by the OLS approach. Any single variable allowed to be included as a control for OLS should also be a valid h variable for my approach.

B.5.1 Pre-Determination

Related to the idea of "bad controls," one can conceptually test whether an h variable is valid by asking whether it would change in response to the equilibrium amenity price. Valid h variables should not respond to prices.

For example, following Goldin and Katz (2011) suppose that pharmacists, relative to lawyers, tend to be paid in family-friendly amenities. Workers sort into occupations on valuations of that amenity. Although score on a law-school entrance exam might be relevant to workers' levels of compensation, it would not be a valid h variable. It is not pre-determined because if the labor market changed such that lawyers were also paid in family-friendly amenities, one would expect a very different group of workers to move in to law. If anything, legal knowledge would be determined in *response* to the wage-amenity equilibrium. In contrast, a skill that is more widely applicable to a range of occupations (such as a cognitive aptitude measure) would not be expected to respond to the outcome of the wage-amenity sorting equilibrium.

[32] Angrist and Pischke (2009) write: "Some variables are bad controls and should not be included in a regression model even when their inclusion might be expected to change the short regression coefficients. Bad controls are variables that are themselves outcome variables in the notional experiment at hand. That is, bad controls might just as well be dependent variables too. Good controls are variables that we can think of as having been fixed at the time the regressor of interest was determined."

B.6 AFQT as an h variable

I argue that score on the Armed Forces Qualification Test (AFQT) is correlated with the level of the wage-amenity frontier, but is independent of workers' locations along their particular frontiers. The AFQT score is the primary summary score from the Armed Services Vocational Aptitude Battery, which was administered to respondents of the NLSY79 in 1980.

As the purpose of the test is to measure basic skills that are applicable to a wide range of vocations, the score by construction should be independent of how employers located at various positions on a wage-amenity frontier view a worker. Although AFQT is in no sense "innate," any effort by a worker to increase his or her AFQT score should result only in a movement to a higher compensation frontier, rather than a movement along a given frontier. Perhaps the most pressing concern is that the decision to invest in AFQT-relevant skills could for some reason be correlated with a preference for wage or for certain types of amenities. For instance, such an argument might be that workers who have a stronger affinity for certain amenities invest more into training for general-purpose skills. Still, such an argument is hard to rationalize with utility-maximizing individuals.

In a more reduced-form sense, AFQT is plausibly a noisy proxy for productivity, in which the noise is uncorrelated with preferences. Heterogeneity in income versus amenities among η-level workers and firms would be determined not by AFQT but rather by workers' and firms' heterogeneous valuations of the amenity. The key is that once we've conditioned on η, AFQT should be irrelevant to both the firms' and workers' problems.

Finally, several foundational papers in labor economics have used AFQT score from the NLSY as a noisy measure of worker productivity. Employer learning and statistical discrimination have been explored by Farber and Gibbons (1996) and Altonji and Pierret (2001), who used AFQT as a measure of worker skills that are unobserved to the employer. In a different vein, Neal and Johnson (1996) used AFQT as a "measure of basic skills that helps predict actual job performance," noting in particular its usefulness of measuring skills present prior to labor market entrance. The importance of AFQT in explaining the wage gap has often been cited as evidence of the racial "skill gap." A central assumption in prior literature has been that AFQT measures general-purpose skills, not preferences. For instance, although many interpretations of the findings of Neal and Johnson (1996) have emerged that do not involve a racial "skill gap," no arguments have been made that controlling for AFQT is revealing a "preference gap" for job amenities, which is precisely what would be needed to invalidated to estimate the neoclassical model of compensating differentials.

C Multiple Amenities

Consider the generalized equation for the multi-dimensional frontier with various z_k variables representing either different amenities or higher-order polynomials in each amenity:

$$w_i + \sum_k \beta_k z_{k,i} = h_i + \varepsilon_i$$

The identification assumption must hold for each z_k:

$$\frac{cov(\varepsilon_i, w_i)}{cov(\varepsilon_i, z_{k,i})} = \frac{cov(h_i, w_i)}{cov(h_i, z_{k,i})}$$

Then $\hat{\beta}_k$ can be represented similarly as before as a ratio of regression coefficients from

$$E^*[h_i|w_i, z_i] = \hat{\delta_w} w_i + \sum_k \hat{\delta_{z_k}} z_{k,i}$$

$$\hat{\beta}_k = \frac{\hat{\delta_{z_k}}}{\hat{\delta_w}}$$

D Non-monotonicity of h

Section 3 assumed $E[h|\eta]$ was everywhere strictly monotone in η. Since $E[h|\eta]$ was identified by $\hat{h}$, this implied that every quantile of $\hat{h}$ corresponded to a different quantile of η.

To convert $\hat{h}$ to quantiles of η in the absence of global monotonicity, the procedure is modified as follows. First, quantiles of $\hat{h}$ are constructed as ordinary. If $E[h|\eta]$ is non-monotonic, then it is also the case that the function $E[h|w, z]$ is non-monotonic. This means that when plotted in w-z space, some quantiles of $\hat{h}$ will be disconnected sets, i.e., they will contain (w, z) tuples that are separated from each other by another quantile of $\hat{h}$. The problem is now that no matter how small we divide quantiles of $\hat{h}$, there is still the possibility the same quantile of $\hat{h}$ may correspond to two levels of η; for instance, this would certainly be the case if η is strictly monotone in w-z space. But all that needs to be done is to modify the quantiles of $\hat{h}$ such that disjoint sets are coded as separate levels. Every level of the re-coded $\hat{h}$ quantiles will now correspond exactly to a different quantile of η.

E A Reduced-Form Model

In the model of Section 2, workers chose wages and amenities according to preferences and subject to a budget constraint with slope β, the level of which was indexed by η. The only unobserved component of the model was ε, the unobservable skill component. Although this model allows for a principled application of economic logic to the labor market, the simplifications limit the extent of the analysis. For instance, what if wage or amenities are observed by the researcher with imprecision? Such problems of measurement error are known to contribute to bias in OLS when affecting regressors. How would additional idiosyncratic variation in wage or amenity bias the estimator?

Whereas the baseline model allowed for unobserved variation only in the determination of the offer curve, this section extends the model to incorporate additional unspecified unobserved variation in both wage and amenity determination. The purpose of this exercise is to determine under what constraints on these idiosyncratic terms $\hat{\beta}^{GMM}$ still retains an interpretation as a relative price of the amenity.

The simplest statistical framework can be described as two equations governing worker i's choice of wage w_i and amenity z_i. First, each outcome is a function of skill level, η_i, which increases potential wage by γ^w and potential amenity by γ^z. Second, each outcome is affected by a sorting parameter indicating the worker's place along the frontier, which I refer to as τ_i. It is important to note that τ_i has no natural structural interpretation; in particular, it does *not* correspond directly to any parameter of the worker's utility function. I introduce this reduced-form parameter to capture the result from the structural model that – as a result of heterogeneous valuations of both workers and firms – workers will be scattered along the income-amenity frontier, and their place along that frontier is indexed by τ_i. Without loss of generality, the variables τ_i and η_i are independent and can be assumed to be mean-zero.[33] Equations 7 and 8 describe this statistical model.

$$w_i = \gamma^w \eta_i + \beta \tau_i \tag{7}$$

$$z_i = \gamma^z \eta_i - \tau_i \tag{8}$$

$$w_i + \beta z_i = (\gamma^w + \gamma^z \delta)\eta_i \tag{9}$$

Equation 9 multiplies Equation 8 by β and adding it to Equation 7. By subtracting out the τ term, Equation 9 provides the mapping of the general statistical model back to the simplistic budget-curve model of Section 2, which did not rely on the reduced-form parameter τ.

E.1 Measurement Error (Statistical Noise)

The parameter β still represents the relative price of wage and amenity. In this model, if we observed η then we could learn β from the data. All unexplained heterogeneity between workers of similar skill could be attributed to preferences, and their differing choices would trace out a line of slope β.

However, there may be other sources of heterogeneity in wage and amenities that are also unobserved. In Section 3, I address unobserved skill heterogeneity in the model of Section 2 by decomposing η_i into an observed part called h_i and an unobserved part called ε_i. But a variety of competing models – including discreteness of job choices, optimization frictions, and multi-dimensional skills – suggest that not all unobserved heterogeneity in job choices can be modeled as unobserved inward-outward shifts of the same-sloping offer curve. To discuss these likely complications to the simple model, I introduce into the statistical model idiosyncratic perturbations in a worker's wage or amenity, rather than common shocks to both. I denote ν_i^w and ν_i^z as independent idiosyncratic components of w_i and z_i. Below, Equations 10 and 11 reproduce Equations 7 and 8

[33] The definition that τ_i and η_i are independent in no way limits the model to a particular class of utility functions. It is simply a definition of τ_i relative to similarly skilled workers. Thus, there is not intended to be a direct correspondence between τ_i and any parameter of the utility function.

with these additional sources of idiosyncrasy.

$$w_i = \gamma^w \eta_i + \beta \tau_i + \nu_i^w \tag{10}$$

$$z_i = \gamma^z \eta_i - \tau_i + \nu_i^z \tag{11}$$

With the additional level of unobserved heterogeneity, it is not obvious what can be learned from the data. In the strict model of Section 2, any worker with an average level of measured skills (h) seen with a high wage and average level of amenities would be interpreted as having a high level of unobserved skill (ε) combined with a strong preference to translate that skill into wage (τ) at the prevailing price, β. With the additional sources of heterogeneity, it's equally likely this is a worker of average skill and average preferences, who simply received an inexplicably high wage offer, represented by a high ν^w.

In order to learn from data that is thought to have been generated by the general statistical model of Equations 10 and 11, it is necessary to place an assumption of well-behavedness on these error terms, given by Equation 12 below. If this well-behaved condition is satisfied, then the same estimator that consistently estimates the parameter of the model of Section 2 will also consistently estimate β from Equation 11.

$$\frac{\sigma^2_{\nu^w}}{\sigma^2_{\nu^z}} = -\beta \frac{\gamma^w}{\gamma^z} \tag{12}$$

Equation 12 states that the variance of the idiosyncratic term in wage relative to the geometric mean of variance in wage due to preferences and skills is equal to the analogous ratio for the amenity. Figure A.6 shows a graphical interpretation of this condition by means of a simulation. If the condition does not hold, it is still possible to discuss the direction of the bias. As the variance of ν^w gets large, the estimate of $\hat{\beta}$ will approach infinity, whereas the estimate of $\hat{\beta}^{GMM}$ tends to 0 as the variance of ν^z increases. The figure also depicts the tilting of $\hat{\beta}$ relative to β in each of these scenarios, as deviations from the baseline simulation.

F $\hat{\beta}^{Linear}$ is equivalent to a ratio of coefficients

The interpretation of the estimator as a ratio of OLS coefficients is particularly useful in adapting the model to multiple amenities.

Consider the linear regression of $h$on w and z:

$$E^*[h_i | w_i, z_i] = \hat{\delta_w} w_i + \hat{\delta_z} z_i. \tag{13}$$

The estimator is algebraically equivalent to the ratio of these two coefficients, $\hat{\beta}^{GMM} = \frac{\hat{\delta_z}}{\hat{\delta_w}}$.

The usefulness of the regression of Equation 13 is somewhat surprising given that the error term,

which is thought to contain unmeasured skill, would be correlated with the regressors. Endogeneity of the error term typically biases coefficients. Although there is not an obvious definition of bias in this case as Equation 13 does not correspond to any sensible structural equation, it is true that we expect the coefficients $\hat{\delta}_w$ and $\hat{\delta}_z$ to change depending on the precision with which skill is observed. If we suppose workers are on linear frontiers of Section 2 such that $w_i+\beta z_i = \eta_i = h_i+\varepsilon_i$ (i.e., wage and amenity coefficients of 1 and β, respectively), then the probability limits of the coefficients are $\hat{\delta}_w^* = \frac{1}{1+\lambda}$ and $\hat{\delta}_z^* = \frac{\beta}{1+\lambda}$, where the constant λ is a scaling factor for the relative variances of h and ε: $\lambda = \frac{cov(h_i,z_i|w_i)}{cov(\varepsilon_i,z_i|w_i)} = \frac{cov(h_i,w_i|z_i)}{cov(\varepsilon_i,w_i|z_i)}$ (the conditional covariances should be read as linear projections). That this equality holds such that the λ scaling is the same for both probability limits hinges on the linearized version of Condition 2 holding:

$$cov^*(w,h|\eta) = cov^*(z,h|\eta) = 0 \tag{14}$$

F.1 Proof

Consider the linear regression of h on w and z, reproducing Equation 13:

$$E^*[h_i|w_i,z_i] = \hat{\delta_w}w_i + \hat{\delta}_z z_i.$$

Proposition: The proposed estimator is algebraically equivalent to the ratio of these two coefficients, $\hat{\beta}^{GMM} = \frac{\hat{\delta_z}}{\hat{\delta_w}}$.

Proof of Proposition

Lemma 1: If Condition 14 holds, then $\frac{cov(w,\varepsilon|z)}{cov(w,h|z)} = \frac{cov(z,\varepsilon|w)}{cov(z,h|w)} \equiv \lambda$ for some number λ.

Proof of Lemma 1

Condition 14 states that $cov(w,h) - \frac{cov(w,\eta)cov(h,\eta)}{var(\eta)} = 0$ and $cov(z,h) - \frac{cov(z,\eta)cov(h,\eta)}{var(\eta)} = 0$, so $\frac{cov(w,h)}{cov(z,h)} = \frac{cov(w,\eta)}{cov(z,\eta)}$. Since $\eta = h + \varepsilon$, a condition similar to that of Condition 14 must hold, but replacing h with ε, so we also have $\frac{cov(w,\varepsilon)}{cov(z,\varepsilon)} = \frac{cov(w,\eta)}{cov(z,\eta)}$. Equating the two steps and re-arranging, we can write $\frac{cov(w,h)}{cov(w,\varepsilon)} = \frac{cov(z,h)}{cov(z,\varepsilon)} \equiv K$ for some number K. By a creative multiplication with 1, K can be rewritten $\frac{Kcov(z,\varepsilon) - K\frac{cov(w,\varepsilon)cov(z,w)}{var(w)}}{cov(z,\varepsilon) - \frac{cov(w,\varepsilon)cov(z,w)}{var(w)}}$. Since our identification assumption yielded $cov(w,h) = \frac{cov(w,\eta)cov(h,\eta)}{var(\eta)}$ and $cov(w,\varepsilon) = \frac{cov(w,\eta)cov(\varepsilon,\eta)}{var(\eta)}$, we have $cov(w,h) = Kcov(w,\varepsilon)$ and by similar steps $cov(z,h) = Kcov(z,\varepsilon)$. These substitutions allow us to rewrite K further as $\frac{cov(z,h) - \frac{cov(w,h)cov(z,w)}{var(w)}}{cov(z,\varepsilon) - \frac{cov(w,\varepsilon)cov(z,w)}{var(w)}}$. Applying the definition of conditional covariance, $K = \frac{cov^*(z,h|w)}{cov^*(z,\varepsilon|w)}$, and analogous steps can be followed to show that K is also equal to $\frac{cov(w,h|z)}{cov(w,\varepsilon|z)}$. By combining these two expressions of K, $\frac{cov^*(z,h|w)}{cov^*(z,\varepsilon|w)} = \frac{cov^*(w,h|z)}{cov^*(w,\varepsilon|z)}$, which we will call attenuation parameter λ.

End Proof of Lemma 1

In the model of Section 2, $w_i + \beta z_i = h_i + \varepsilon_i$. The probability limits of the coefficients are derived as follows, where all conditional variances and covariances are linear projections:

$$\begin{aligned}\hat{\delta}_w^* &= \frac{cov^*(h,w|z)}{var(w|z)} \\ &= 1 - \frac{cov^*(\varepsilon,w|z)}{var(w|z)} \\ &= 1 - \frac{cov^*(\varepsilon,w|z)}{var(w|z)}\frac{cov^*(h,w|z)}{cov^*(h,w|z)} \\ &= 1 - \hat{\delta}_w^*\lambda \\ &= \frac{1}{1+\lambda}\end{aligned} \qquad \begin{aligned}\hat{\delta}_z^* &= \frac{cov^*(h,z|w)}{var(z|w)} \\ &= \beta - \frac{cov^*(\varepsilon,z|w)}{var(z|w)} \\ &= \beta - \frac{cov^*(\varepsilon,z|w)}{var(z|w)}\frac{cov^*(h,z|w)}{cov^*(h,z|w)} \\ &= \beta - \hat{\delta}_z^*\lambda \\ &= \frac{\beta}{1+\lambda}\end{aligned}$$

Thus, the probability limit of $\frac{\hat{\delta}_z}{\hat{\delta}_w}$ is β when the model of Section 2 is applied.

In the more general statistical model of Section E, the probability limits of the coefficients can be derived in a similar way. As a function of primitives, the outcome-generating equations are $w_i = \gamma^w(h_i + \varepsilon_i) + \beta\tau_i + \nu_i^w$ and $z_i = \gamma^z(h_i + \varepsilon_i) - \tau_i + \nu_i^z$. Solving for h, we have the slightly more complicated frontier $h = \frac{w+\beta z - \nu^w - \beta\nu^z}{\gamma^w + \beta\gamma^z} - \varepsilon$.

$$\begin{aligned}\hat{\delta}_w^* &= \frac{cov^*(h,w|z)}{var(w|z)} \\ &= \frac{var(w|z) - \sigma^2_{\nu^w}}{var(w|z)(\gamma^w + \beta\gamma^z)} - \frac{cov^*(\varepsilon,w|z)}{var(w|z)} \\ &= \frac{var(w|z) - \sigma^2_{\nu^w}}{var(w|z)(\gamma^w + \beta\gamma^z)} - \frac{cov^*(\varepsilon,w|z)}{var(w|z)}\frac{cov^*(h,w|z)}{cov^*(h,w|z)} \\ &= \frac{var(w|z) - \sigma^2_{\nu^w}}{var(w|z)(\gamma^w + \beta\gamma^z)} - \hat{\delta}_w^*\lambda \\ &= \frac{var(w|z) - \sigma^2_{\nu^w}}{var(w|z)(\gamma^w + \beta\gamma^z)(1+\lambda)}\end{aligned}$$

$$\begin{aligned}\hat{\delta}_z^* &= \frac{cov(h,z|w)}{var(z|w)} \\ &= \beta\frac{var(z|w) - \sigma^2_{v^z}}{var(z|w)(\gamma^w + \beta\gamma^z)} - \frac{cov^*(\varepsilon,z|w)}{var(z|w)} \\ &= \beta\frac{var(z|w) - \sigma^2_{v^z}}{var(z|w)(\gamma^w + \beta\gamma^z)} - \frac{cov^*(\varepsilon,z|w)}{var(z|w)}\frac{cov^*(h,z|w)}{cov^*(h,z|w)} \\ &= \beta\frac{var(z|w) - \sigma^2_{v^z}}{var(z|w)(\gamma^w + \beta\gamma^z)} - \hat{\delta}_z^*\lambda \\ &= \beta\frac{var(z|w) - \sigma^2_{v^z}}{var(z|w)(\gamma^w + \beta\gamma^z)(1+\lambda)}\end{aligned}$$

The ratio is

$$\beta\frac{var(w|z)}{var(z|w)}\frac{var(z|w) - \sigma^2_{v^z}}{var(w|z) - \sigma^2_{\nu^w}} = \beta\frac{var(w|z)var(z|w) - var(w|z)\sigma^2_{v^z}}{var(w|z)var(z|w) - var(z|w)\sigma^2_{\nu^w}}$$

In the general statistical model, there are two sets of cases in which the probability limit of the ratio of coefficients is β. First, in the special case in which $\sigma^2_{v^w} = \sigma^2_{v^z} = 0$, we have a statistical model isomorphic to that of Section 2. Second, when the idiosyncratic components are well balanced such that $\frac{\sigma^2_{\nu^w}}{\sigma^2_{v^z}} = \frac{var(w|z)}{var(z|w)}$.

End Proof of Proposition

G Bias if Conditional Independence Does Not Hold

In the linear model, it can be shown with conditional independence that the estimator converges in probability to the following quantity (a full derivation is contained in the Working Paper):

$$\beta - \frac{var(\eta)}{var(z)} \frac{cov^*(h, z|\eta)}{cov^*(h, \eta|z)}$$

If $cov^*(h, z|\eta) = 0$, then bias is none.

Suppose the identification assumption does not hold. In particular, let our h variable be not just a proxy for η, but furthermore a positive correlate of z even after conditioning linearly on η, so at every η, $cov(z, h|\eta) > 0$ and $cov(w, h|\eta) < 0$. This could be the case, for instance, if h were a skill highly valued in high-amenity jobs or if workers with a high amenity preference sought to increase their h.

Of course the variances are positive, and we should typically think of the denominator $cov^*(h, \eta|z)$ as being positive, since at any level of the amenity, workers with more h are expected to have better compensation. Therefore, if it is the case that $cov^*(z, h|\eta) > 0$ then we should expect $\hat{\beta} < \beta$. If z is a "costly" amenity, then β will be negative. A further negative bias will cause the amenity to appear even more costly than it actually is.

H Further robustness checks

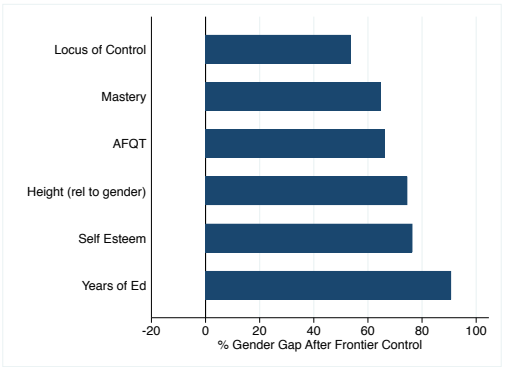
The first dimension of robustness I explore in the analysis of demographic pay gaps is the choice of h variable that proxies for the level of the total compensation frontier. Figure B.1 summarizes the main results for the gaps by gender and race for the six candidate h variables described earlier. I examine only the gaps by gender and race here, as such a wide set of h variables is not available in the dataset used to measure the gap by parent income. Between 50% and 75% of the gender pay gap in the main analysis sample persists within frontiers, depending on choice of h variable. In contrast, the results by race straddle 0 fairly closely. Reassuringly, the findings on demographic pay gaps are relatively similar regardless of which h variable is used.

The remaining robustness checks described below are also summarized in Table B.1, in the order in which they are discussed.

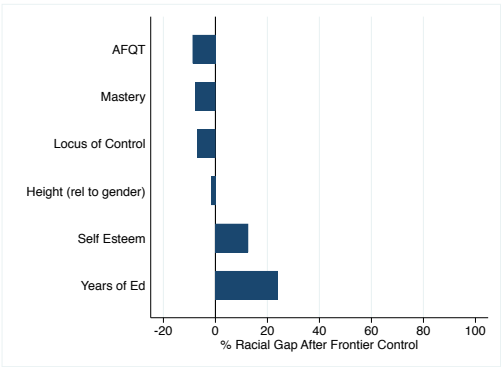
Much prior literature has emphasized the sensitivity of the gender pay gap to how one defines the sample. In particular, differences in hours worked seem to play important roles. For this reason, Panel A of Figure A.4 reproduces Figure 11 when the entire analysis is run only on the

Figure B.1: Pay Gap Accounting in the NLSY, by h variable

(a) Gender Gap



(b) Race Gap



Notes: This figure repeats parts of the analysis of Figure 11 with different h variables. Panel A repeats the result by gender, and Panel B repeats the result by race. A similar re-analysis is not performed for result by parent background because the same set of h variables is not available in the PSID. Additional sample details are found in Section 5.1.

Table B.1: Additional Specifications for Demographic Gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	baseline	full-time	own-sex prices	female prices	male prices	including prestige	individual prices
Gender	.662	.744	.815	.405	.756	.873	.844
Race	-.087	-.052	-.085	-.126	-.105	-.088	-.056
Parents	.065	.068	.157	-.001	.082	.296	.213

Notes: All columns report the share of the income gap remaining after Rosen frontiers are included as controls. Column 1 reports these numbers for the specification of Figure 11, and all other columns are perturbations of this exercise as follows. Column 2 restricts the sample to workers who report at least 40 hours of work per week. Column 3 estimates prices separately by gender and applies the relevant gender's prices. Columns 4 applies to the whole sample the prices estimated among only women, and Column 5 applies prices estimated among men. Column 6 additionally includes the amenity of prestige sourced from the sociology literature, as discussed in the text. Column 7 prices amenities without aggregating to the occupation level. Additional sample details are found in Section 5.1.

sub-sample of workers who report working at least 40 hours per week. As expected, the gender gap in the full-time sample is substantially smaller, at $24,273. However, the gap controlling for offer curves remains high at $18,049. Thus, amenity substitution accounts for 74% of the gender pay gap among full-time workers. It is also worth noting that whereas the gender gap falls considerably upon imposing the hours restriction, the race gap among full-time workers is substantially larger, although amenity substitution still plays little-to-no role for race.

Another concern particularly relevant to the gender pay gap is whether discrimination in the labor market may lead men and women to face different prices. Such a case seems unlikely if labor market discrimination functions in an absolute way to reduce women's total compensation. On the other hand, it may be that discrimination functions in a more relative way, in which different genders must pay a different wage cost to enter certain types of occupations. In such a model, the Rosen frontiers for men and women could have different slopes. To explore this possibility, I calculate amenity prices separately by gender, and use sex-specific price estimates to build Rosen frontiers. Within frontiers of sex-specific slopes, 82% of the gender gap persists due to amenity substitution. I also perform additional robustness applying female-specific and male-specific prices to the entire sample, which yield respective estimates of 41% and 76% of the gender gap remaining. Qualitatively, these results all show a large role for compensating differentials in causing the gender pay gap, regardless of which gender I estimate prices on.

More generally, a major caveat worth reiterating is that the results are likely to be sensitive to which amenities are included, as discussed in the Introduction. I do not have a systematic solution to this problem. However, in an attempt to broaden the scope of the amenities beyond the types of measures that have typically been examined by economists, I turned to the sociological literature, in which several studies relating to occupational prestige have been conducted dating back to Duncan's (1961) socioeconomic index and Blau and Duncan's (1967) application of the index to occupational mobility across generations. A key input into the original prestige rankings was surveyed opinion scores. Davis et al. (1992) provide survey measures of prestige rankings from the 1989 General Social Survey, but updated for the 1980 Census occupation titles. Occupational prestige was measured among a national sample of 1,500 respondents asked to rank each occupation's prestige on a scale from 1 to 9. Not every respondent had to rank every occupation, but sufficient overlap was included to normalize respondents' scores. Panel B of Figure A.4 modifies the analysis of Figure 11 to include the occupational prestige score as an amenity. In this case, the additional dimension of heterogeneity has magnified each result: 87% of the gender gap and 30% of the gap by parent income are now attributed to costly amenity substitution, and the race gap is 8% lower due to this channel.

Finally, although the main analysis has priced amenities at the occupation level, the fact that some amenities are observed from the NLSY at the worker level (regular schedule and non-profit) begs the question of whether the analysis would be much different if all amenities were just priced at the worker level. In this case, pricing amenities at the worker level happens magnifies the findings slightly by gender and parent background to 84% and 21% respectively while the race result is still relatively small at -6%. The fact that some results are made larger when amenities are priced at the individual level may indicate that there is meaningful variation to exploit in the relationship between income, amenities, and skills at a more granular level than at the occupation.

References

Altonji, J. G. and C. R. Pierret (2001): "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 116, 313–350.

Angrist, J. D. and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press, 1 edition ed.

Bell, A., S. B. Billings, S. Calder-Wang, and S. Zhong (2025): "An Anti-IV Approach for Pricing Residential Amenities: Applications to Flood Risk *," .

Blau, P. M., O. D. Duncan, and A. Tyree (1967): *The American occupational structure*, New York, NY: Wiley, oCLC: 232432.

Davis, J. A., T. W. Smith, R. W. Hodge, K. Nakao, and J. Treas (1992): "Occupational Prestige Ratings from the 1989 General Social Survey: Version 1," Tech. rep.

Duncan, O. D. (1961): "A socioeconomic index for all occupations," *Class: Critical Concepts*, 1.

Farber, H. S. and R. Gibbons (1996): "Learning and Wage Dynamics," *The Quarterly Journal of Economics*, 111, 1007–1047.

Goldberger, A. S. (1984): "Reverse Regression and Salary Discrimination," *The Journal of Human Resources*, 19, 293–318.

Goldin, C. and L. F. Katz (2011): "The Cost of Workplace Flexibility for High-Powered Professionals," *Annals of the American Academy of Political and Social Science*, 638, 45–67.

Han, S. and S. Yamaguchi (2015): "Compensating wage differentials in stable job matching equilibrium," *Journal of Economic Behavior & Organization*, 114, 36–45.

Hwang, H., W. R. Reed, and C. Hubbard (1992): "Compensating Wage Differentials and Unobserved Productivity," *Journal of Political Economy*, 100, 835–858.

Mortensen, D. (2003): *Wage dispersion: why are similar workers paid differently*, Zeuthen lecture book series, Cambridge, Mass: MIT Press.

Neal, D. A. and W. R. Johnson (1996): "The role of premarket factors in black-white wage differences," *Journal of Political Economy*, 104, 869.

Rosen, S. (1986): "Chapter 12 The theory of equalizing differences," in *Handbook of Labor Economics*, Elsevier, vol. 1, 641–692.