Three Essays on Estimation with Unpriced Amenities

by

Ryan D. Nunn

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Public Policy and Economics)
in The University of Michigan
2012

Doctoral Committee:

Professor Miles S. Kimball, Co-chair
Associate Professor Daniel S. Silverman, Co-chair
Professor Brian P. McCall
Associate Professor Melvin Stephens Jr.

To my parents

ACKNOWLEDGEMENTS

First, thanks to my committee: Miles Kimball, Daniel Silverman, Brian McCall, and Mel Stephens. They have been a constant source of great ideas and encouragement. I have enjoyed working with and learning from them during my time at Michigan.

The support of my family - Doug, Lisa, Kyle, Tyler, Olivia, Grace, and Dylan - has meant a great deal to me. My friends, including especially Adam, Aditya, Bert, Brendan, Collin, Dmitry, Josh, Mary Anne, Matt, Michael, Minayo, Moira, Ozan, and Seb, have made the last six years much more pleasant than they would otherwise have been. I am especially indebted to Laura Kawano for all her job market advice and comments on many paper drafts.

For her considerable patience, and in recognition of the spiritual toll exacted by years of "conversation" related to match quality heterogeneity, I thank Stephanie Guthrie.

Finally, I couldn't have done it without Gabriel Ehrlich, David Ratner, and David Foster Wallace.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

This dissertation is a study of how estimation and optimal policy can be incomplete when important labor market and hedonic variables are unobservable. Methods for inferring the relevant unobserved values and correcting public policy are presented in the dissertation. The first chapter is an examination of labor match quality and match amenities, neither of which are directly observable. The second chapter is an analysis of dynamics in subjective well-being data, with methods developed to facilitate comparison of happiness-relevant events that are otherwise not comparable. The final chapter is a more detailed exploration of the empirical and policy consequences of omitting labor market match amenities.

In the first chapter, "Match Quality with Unpriced Amenities", I examine the role that variation in job match quality plays in determining workers' search decisions. Typically, monetary productivity is assumed to be the sole determinant of the quality of a worker-firm match. I develop a structural search model that allows job match quality to depend additionally on unpriced job amenities, permitting match quality estimation that is robust to both unobserved amenities and selection. I estimate the model with tenure data using the simulated generalized method of moments. The chapter demonstrates that previous estimates relying principally on wage data, rather than duration data, are incomplete in certain respects. The standard deviation of job amenities is found to be nearly as large as that of monetary productivity in data from the 1979 NLSY. I then use the model to investigate the welfare consequences of wage taxation and unemployment insurance. Traditional estimates of deadweight loss from wage taxation are increasingly overstated as job amenity dispersion rises.

In “Accounting for Adaptation in the Economics of Happiness”, coauthored with Miles Kimball and Dan Silverman, we attempt to take the dynamics of happiness seriously. Happiness economics is one potentially useful way to approach unpriced and non-market goods, since subjective well-being responds to such goods in a predictable way. Previous work generally estimates fixed, time-invariant effects on happiness, which can lead to various biases. Our contribution is an econometric approach, based on nonlinear least squares estimation, that is robust to individual fixed effects, imprecisely-dated data, bias from imperfect recall, and permanent event consequences. Most importantly, though, it incorporates differences in the time path of happiness after events, permitting comparison of events with dissimilar time paths and mean reversion. Our method is used to analyze a variety of events in the Health and Retirement Study panel. Many of the variables studied have substantial consequences for subjective well-being - consequences that differ greatly in their time profiles.

In the final chapter, “Taxation, Match Quality and Social Welfare”, Brendan Epstein and I extend the analysis conducted in the first chapter. The costly, endogenous supply of job amenities by firms has a number of interesting implications for labor markets and public policy. In particular, results from the present paper reinforce a conclusion from the first chapter: the public finance literature concerning deadweight loss estimation should take amenities into account. Previous work has made the case for utilizing data on taxable income elasticities, rather than labor supply elasticities, for the purpose of calculating social welfare consequences of taxation. We show that it is necessary to estimate the heterogeneity and parameters of amenity supply in order to correctly infer deadweight loss from taxation. Deadweight loss is generally overestimated in research that omits explicit consideration of amenities, and this overestimation is proportional to the quantitative significance of heterogeneity in amenities across job matches. The endogenous supply of amenities, rather than exogenous endowment, is shown to exacerbate this overestimation. Examination of a friction-less economy reveals that this result is not necessarily an artifact of search. Finally, we examine the dynamic response of taxable income and social welfare to unexpected tax changes.

CHAPTER II

Match Quality with Unpriced Amenities

2.1 Introduction

A chief function of labor markets is to match firms and workers in such a way as to exploit gains from specialization and differences in ability and preferences. Some of what matters for the optimal allocation of labor is related to monetary productivity. However, many relevant variables are unobservable and unpriced. The monetary and non-monetary total surplus generated by a particular worker-firm pairing ("match quality") will vary across possible matches, creating a ranking of potential matches by their surplus. The heterogeneity of match quality is strongly suggested by various data, most notably wage dispersion (e.g., Mortensen, 2003, Woodcock, 2007, Bowlus, 1995). The existence of this heterogeneity has important implications for worker search behavior and its associated social welfare consequences. If labor market comparative advantage is substantial, there is tremendous surplus generated by the optimal coordination of workers and jobs. If comparative advantage is minimal, there is little need to be concerned with policies that affect match quality.

Some of match quality is entirely monetary: a worker may possess an unusual talent for a particular job task, for instance, and this will be reflected in increased output and wages. Another important sort of worker-firm interaction, however, consists of non-monetary factors. For example, some jobs entail a considerable amount of psychologically costly work. Individual workers will vary in their tolerance for this, and will possess different rankings of jobs by level of unpleasantness. Many jobs even include some pleasant aspects, which will be more or less valuable to different individuals. These non-monetary considerations are difficult to observe and cannot typically be inferred from wage variation.

To see why total match quality variation is not obtainable from wage data in the presence of an unpriced amenity, consider the following example. Imagine an employed worker who chooses to accept a new job at a higher overall match quality, where the increase in quality is equally divided between higher productivity and higher amenity. For simplicity, let the worker's compensation (wage plus amenity) be half the surplus to the match. Though the worker's full compensation will be higher after the switch, her observed wage (compensation minus amenity) will be identical before and after the switch, because the amenity improvement balances out the productivity improvement. No wage dispersion is generated by this job transition, and yet match quality variation exists by stipulation.

Both on- and off-the-job search will allow workers to move into higher match quality jobs, which generates an accepted match quality distribution distinct from the match quality "offer" distribution. This paper develops a model that can identify both of these distributions. In contrast, direct wage regressions can recover only the accepted distribution, cannot distinguish productivity and amenity, and cannot recover total match quality without additional assumptions on the covariance of amenity and productivity. Attempts to infer match quality dispersion from wage residuals are also weakened by measurement error in wages, which is substantial and difficult to deal with in the familiar panel datasets.

Structural estimation using tenure data offers a credible solution to these problems. The model embeds match quality and amenity heterogeneity in a standard labor search context. The tenure distribution moves in response to changes in the dispersion of match quality, reflecting varying rates of separations and on-the-job transitions at different levels of match quality. This is the most important source of identification. Monetary and non-monetary match quality component distributions are separated by examining the correlation of tenure and wages. A higher correlation indicates a larger role for monetary match quality, since high quality matches (high tenure) are associated with high wages. A lower correlation indicates the opposite, with high quality matches exhibiting a weak or negative relationship with wages due to the large fraction of unobserved amenities in total worker compensation.

The estimated match quality and other parameters have important consequences for optimal wage tax policy and optimal unemployment insurance policy, among other things. Wage taxes generate a significant and quantifiable distortion by changing the endogenous productivity-amenity composition of match quality. Optimal unemployment benefits are calculated and shown to be increasing in the variance of match quality.

Section 2.2 is an outline of some simpler, closed-form models of labor supply that build intuition for the baseline empirical model. Restrictions are imposed, relative to the baseline model presented later, that allow for an analytic characterization of unemployment and wages in terms of the match quality distribution. Calculation of the deadweight loss from a wage tax requires knowledge of the "accepted" match quality distribution (i.e., the distribution observed in employed workers), which is an endogenous object and is solved for in terms of parameters of the model.

In Section 2.2.1, I discuss the limitations of some previous wage-based approaches and detail the considerations that motivate this paper's innovations: unpriced amenities and the unobserved match quality offer distribution (as contrasted with the accepted distribution). This section argues that accounting for these factors requires a different approach than previously employed. Job amenities are an important part of compensation that will bias approaches primarily reliant on wage data. Ignoring the offered/accepted distinction will lead to underestimates of offered match quality, which can bias policy implications.

Section 2.3 contains the baseline model. The economy consists of ex ante identical workers and firms with wages endogenously set each period by Nash bargaining. In each period, workers may receive a job offer, occurring with a probability that is conditional on employment status. Importantly, each worker-firm meeting is characterized by a match quality composed of both a monetary productivity term and a non-monetary job amenity. Each component has an exogenous normal distribution with a variance that will be separately identified.

Section 2.4 is a description of the data, which comes from the 1979 National Longitudinal Survey of Youth, including tenure, wage, and demographic information. Section 2.4.1 explains the identification strategy. Most importantly, identification of the match quality parameters is from the shape of the tenure distribution. Section 2.4.2 is a description of the estimation procedure. Section 2.5 contains the results of the baseline model. Amenities are found to be about four-fifths as large as monetary productivity, in terms of standard deviation, and a standard deviation of overall match quality is approximately \$12.19 per hour, equivalent to nearly 40 percent of average flow output. Job switching costs are estimated to be about four months of the average observed wage.

In Section 2.5.1 and Section 2.5.2, I conduct applications of the model, including an analysis of the distortion from preferential tax treatment of unpriced amenities and the contribution of match quality dispersion to an understanding of optimal unemployment insurance. Wage taxation encourages workers to find work in higher-amenity, lower-wage (and lower monetary productivity) matches, which leads to lower social welfare. The distortion quickly becomes large at plausible tax rates, and I show that match quality heterogeneity makes it impossible to infer the deadweight loss from the taxable income elasticity alone.

As match quality variance rises, the optimal UI benefit also rises, reflecting the increased social benefit to job search (imperfectly internalized by the worker). I find that the optimal UI benefit as a fraction of average flow surplus rises by about 14 percentage points when the standard deviation of match quality is doubled. The extent of match quality variation should be an input to a more general theory of optimal UI.

Section 2.5.3 is a comparison of the model's results with results from an entirely wage-based approach. Section 2.6 is an examination of the relevant literature. Section 2.7 discusses future research. First, estimates of match quality variation are a necessary input to an understanding of cyclical productivity dynamics. Significant work has gone into characterizing the behavior of average and marginal labor productivity over the cycle, as they have important implications for the welfare cost of recessions. Second, a more general treatment of social welfare and match quality remains to be conducted. I explore this in a separate paper with a coauthor.

2.2 Theory

Before proceeding to the structural model used for estimation, I consider a simpler setting that develops some of the relevant intuition. It retains search frictions, but uses a simpler wage-setting rule and a few other model restrictions that allow for closed-form expressions.

Consider an economy with a measure of workers and a measure of profit-maximizing firms. A match between a worker and firm generates surplus $m > b$, where b is the flow benefit to unemployment (i.e., the worker's outside option). m is specific to a particular match between the worker and a firm, but does not vary over time within the match. New job offers, with new draws of match quality, are distributed according to a cdf M_{offer} and arrive with probability α whether the worker is employed or unemployed. Workers make take-it-or-leave-it wage offers to firms (note that Section 2.3 will instead assume a Nash bargain over the match surplus). Assume further that workers transition exogenously into unemployment with probability s, and that this separation shock either occurs or does not occur prior to the "new job" shock. Then the worker's discrete value function will be

$$W(m) = m + \beta sU + \beta(1-s)(1-\alpha)W(m) + \beta(1-s)\alpha E_{m'}[max\{W(m'), W(m)\}],$$

where β is the discount factor and U is the unemployment value function, given by

$$U = b + \beta\alpha E_{m'}[max\{W(m'), U\}].$$

Note that the wage is simply m, as firms are competitive in the labor market. I do not model the demand side of the labor market in this section. The reservation wage will be m_r such that $W(m_r) = U$. It is easy to see that $w_r = m_r = b$. Since the on- and off-the-job arrival probabilities are equal, workers will always prefer employment to unemployment when they receive at least the flow benefit to unemployment. In the baseline model presented later, this will no longer be the case, because differences in arrival probabilities will lead workers to prefer search from either the employed or unemployed state.

In this simple economy, it is possible to characterize the steady state cumulative distribution function of wages χ (which is just the distribution of actual match quality) and the unemployment rate. To accomplish this, some steady state identities are required. First, inflow into employment below or equal to a particular match quality must equal outflow from below or equal to that match quality. Second, inflow into unemployment will equal outflow from unemployment. For simplicity, set the lower bound of the M_{offer} support to be b. This will cause M_{offer} and χ to have identical supports, albeit different values throughout the support. The equations below characterize u and $\chi(m)$.

$$\alpha u(M_{offer}(m) - M_{offer}(m_r)) \tag{2.1}$$

$$= s(1-u)\chi(m) + (1-s)\alpha(1-u)\chi(m)(1 - M_{offer}(m)) \;\; \forall m \in \chi^{-1} \tag{2.2}$$

$$s(1-u) = \alpha u(1 - M_{offer}(m_r)), \tag{2.3}$$

where u is the fraction of workers who are unemployed. Note that workers always switch to jobs with higher wages. M_{offer} is exogenously given, which leaves the function χ and the scalar u to be endogenously determined. The second equation yields an unemployment rate $u = \frac{s}{s+\alpha}$, recalling that $M_{offer}(m_r) = 0$ by stipulation. χ is pinned down by u and equation 1, as shown below

$$\chi(m) = \frac{sM_{offer}(m)}{s + (1-s)\alpha(1 - M_{offer}(m))}.$$

Now separate match quality m into two components: productivity π and non-monetary amenity q. The amenity is randomly endowed and not constructed by a firm. Wages are now a function of π rather than m. Let the two components of match quality be uncorrelated and perfectly substitutable in consumption, and let productivity π be taxed at rate τ. Amenities remain untaxed. After-tax wages are now $(1-\tau)\pi$, and

$$W(\pi,q) = (1-\tau)\pi + q + \beta sU + \beta(1-s)(1-\alpha)W(\pi,q) + \beta(1-s)\alpha E_{\pi',q'}[max\{W(\pi',q'),W(\pi,q)\}].$$

U is altered similarly. The reservation wage is set such that $b = (1-\tau)w_r + q \;\rightarrow\; w_r = \frac{b-q}{1-\tau}$. Intuitively, the taxation of income will introduce a distortion, with workers choosing jobs that have relatively high amenity value and low wages. Since utility is linear in match quality, it is possible to quantify the social welfare loss by calculating the integral

$$\int_{\underline{\pi}}^{\overline{\pi}} \int_{\underline{q}}^{\overline{q}} (1-u)(\pi+q)m_a(\pi,q)dq\,d\pi,$$

where m_a is the density associated with the cdf $M_{accepted}$, the fraction of employed workers in jobs at or below productivity π and amenity q, and $\{\overline{\pi},\underline{\pi}\}$, $\{\overline{q},\underline{q}\}$ are the upper and lower bounds of π and q, respectively.

First, however, it is necessary to calculate $M_{accepted}$ and m_a. The new equations describing their behavior are

$$\begin{aligned}
&\alpha u(M_{offer}(\pi,q) - M_{offer}(\pi_R(q),q)) = s(1-u)M_{accepted}(\pi,q)\\
&+ (1-s)\alpha(1-u)M_{accepted}(\pi,q)(1-M_{offer}(\pi_r(q'|q,\pi),q)) \;\; \forall \pi,q \in M^{-1}_{accepted}\\
&s(1-u) = \alpha u(1-M_{offer}(\pi_R(q),q)).
\end{aligned}$$

I relax the assumption that all wage offers are sufficient to induce movement out of unemployment, because one of the mechanisms by which higher taxes induce a distortion is the unemployment they generate. The resulting expressions are thus somewhat messier.

Unemployment is now $u = \frac{s}{s+\alpha-\alpha M_{offer}(\frac{b-q}{1-\tau},q)}$. The accepted match quality distribution is

$$M_{accepted}(\pi, q) = \frac{\alpha s(M_{offer}(\pi, q) - M_{offer}(\frac{b-q}{1-\tau}, q))}{\alpha(1-s)(\alpha - \alpha M_{offer}(\frac{b-q}{1-\tau}, q))(1 - M_{offer}(\pi + \frac{q-q'}{1-\tau}, q')) + s(\alpha - \alpha M_{offer}(\frac{b-q}{1-\tau}, q))}.$$

Given a functional form assumption on M_{offer}, m_a and the size of the distortion can be calculated. The baseline model presented in Section 2.3 will aim to estimate M_{offer} in a more complicated context, permitting the calculation of $M_{accepted}$ and the DWL.

2.2.1 Comparative Advantage, Job Amenities, and the Offered/Accepted Distinction

The discussion of comparative advantage in the labor market has been limited in at least one important way. Abowd et al. (2009) describe the literature that treats correlation between worker and firm productivity types as fully constituting what is meant by "comparative advantage": "The estimated correlation between worker and firm effects from the earnings decomposition is close to zero, a finding that is often interpreted as evidence that there is no sorting by comparative advantage in the labor market."(Abowd et al., 2009, abstract) For example, consider the following earnings decomposition by Abowd et al on matched firm-worker data: $\log(w_{it}) = x_{it}\beta + \theta_i + \psi_{J(i,t)} + \epsilon_{it}$, where w_{it} is the person-year specific wage, x is a set of demographic and labor market variables, θ is a person-specific fixed effect and ψ is a firm-specific fixed effect. This earnings decomposition is not dispositive with respect to labor market comparative advantage, however. Comparative advantage should be understood to refer to the relative superiority of *any* conceivable labor market match. The usual interpretation allows comparative advantage to operate only across worker and firm type; i.e., workers with low average wages may be more productive when associated with low productivity firms, but no allowance is made for the possibility that some low-wage workers may be profitably associated with *particular* low or high productivity firms but not others. Put more formally, a finite number of possible worker-firm matches are ordered by their productivity, and there is no requirement that matches by workers and firms of a particular $\{\overline{\theta}, \overline{\psi}\}$ combination all have the same productivity.

While interesting and informative about the production function, the worker and firm fixed effects correlation does not end the discussion of labor market comparative advantage. However, other wage regression specifications are potentially more informative about match quality. Consider a regression similar to the decomposition shown above.

$$\log(w_{it}) = x_{it}\beta + \theta_i + \psi_{J(i,t)} + \mu_{i,j} + \epsilon_{it},$$

where the only additional variable is μ_{ij}, a worker-firm match fixed effect. With no amenities and for the accepted match quality distribution, the distribution of μ_{ij} summarizes match quality heterogeneity.

Interestingly, Woodcock (2007) finds that the data reject the Abowd et al specification in favor of the same specification but with match effects included. Woodcock (2008) further develops the econometric model with match effects, discusses the nature of biases in the various specifications, and applies the match effects model to inter-industry and inter-sex wage differentials.

The approach taken in this paper is quite different than that of the previously-discussed literature. In addition to explicitly dealing with idiosyncratic firm-worker match quality, I account for unpriced job amenities and the necessity of recovering the unobserved offer distribution of match quality. This section demonstrates the importance of the two concerns for a proper understanding of the labor market.

In the absence of these issues, this paper would provide an alternative but not obviously preferable picture of the match quality distribution, relative to a wage-based approach. The first issue is the distinction between the component of job surplus that accrues directly to employers (termed "productivity") and the component that enters directly into the worker's utility ("job amenities"). The latter is not directly observable and has traditionally been inferred from wage variation in a model that presumes equality of utility across jobs, conditional on observable characteristics of the firm and worker.[1] Alternatively, one can assume that amenities do not exist, and interpret wage dispersion as match quality dispersion. This approach leads one away from classical labor markets and towards a search framework, since classical workers would not be allocated to jobs in which match quality is less than ideal.

In a search context, the assumption of equal utility across jobs is unnecessary and unwarranted. Since it is not possible for a worker to instantly examine all offers, we should expect job offers to vary in quality and workers to often choose employment that would be non-ideal in a classical market. This is a natural setting for examining match quality and comparative advantage, though amenities and compensating differentials require more structure or data to identify (as in the classical "equal utility" setup).

Workers in a search model receive compensation that is a function of job surplus (productivity plus amenities). Note that "compensation" here includes the amenity, and so is distinct from the observed wage. Neither this compensation nor the observed wage is equal to productivity, as in a classical labor market, but rather is set by a Nash bargain over the surplus. Without making an assumption about how productivity and the amenity endogenously co-vary, it is impossible to infer total match quality from observed wages. Under any assumption about the covariance, amenities and productivity are not separable.

Productivity and amenities are mechanically related in the following way. First, note that $\sigma_m^2 = \sigma_\pi^2 + \sigma_q^2 + 2Cov(\pi, q)$ by construction, since $m = \pi + q$. If productivity and amenities are orthogonal, the variance of match quality will be higher than that of productivity. One way to concretize the estimation problem is to imagine a worker switching from a job with match quality m_1 to a job with match quality m_2, $m_2 > m_1$. Suppose that $m_2 - m_1$ is equally divided between increases in π and q, the monetary and non-monetary components, and that workers and firms split the surplus in half. The worker's initial compensation (in utility) is $w_1 + q_1 = \frac{\pi_1 + q_1 + b}{2}$, which implies that $w_1 = \frac{\pi_1 - q_1 + b}{2}$. The new wage, then, will be

[1] Rosen 1986.

$w_2 = \frac{\pi_2 - q_2 + b}{2} = \frac{\pi_1 - q_1 + b}{2}$, since by stipulation $\pi_2 - \pi_1 = q_2 - q_1$ and surplus is shared evenly. The econometrician observes no change in the wage, yet match quality variation has been generated by a move to a superior match. Information about the true match effect, inclusive of non-pecuniary characteristics, can fortunately be extracted from the worker's decision to stay at or exit from her job.

The second distinction that motivates this paper's analysis is between the accepted and offered match quality distributions. The offer distribution, which governs the range of possible match quality, will generate many proposed matches that are not acceptable to workers and firms, while the accepted distribution includes only actual, realized matches. Necessarily, wages can only be directly informative about the accepted distribution, which first order stochastically dominates the offered distribution. This is true with or without on-the-job search, as long as workers reject at least a fraction of wage offers. With on-the-job search, it is true regardless. Changes in the accepted distribution are an important part of cyclical dynamics, and the steady state accepted distribution is an object of interest in its own right. However, the offer distribution is the relevant object for search considerations, and by implication optimal unemployment insurance and the deadweight loss of wage taxation. Workers are interested in the set of possible jobs and set their reservation wages accordingly; the returns to search do not depend on the distribution of already-accepted job offers.

Further, if the offer and accepted match quality distributions are assumed to be identical, dispersion of the offer distribution will be underestimated. Workers cluster in the right tail of the distribution, initially because they reject a fraction of all offers, and then over time as on-the-job search increases the quality of their matches. For both reasons, a fraction of offered low quality matches do not appear in the accepted distribution, reducing the variance of the accepted distribution relative to the offer distribution.

2.3 Model

The model builds on those developed by Mortensen and Pissarides (1994) and Shimer (2006). I account for on-the-job search, match quality heterogeneity, endogenous job destruction stemming from both job-to-job transitions and changes in the idiosyncratic productivity of a match, and job switching costs. Search frictions are such that workers and firms meet each other only occasionally. Matches produce a flow surplus that depends on idiosyncratic (time-varying) productivity x, a time-invariant match-specific monetary productivity π, and an amenity q that is produced endogenously by firms. Because matching opportunities are scarce, surplus is generated by successful matches, and wages are set by bargaining over this surplus. A worker's values of being unemployed or employed, as well as firms' value of employment, are represented as Bellman equations. Since the resulting functions are both monotonic and discounted, these equations are contraction mappings.[2]

Search models have the virtue of rationalizing many stylized labor market facts, like involuntary unemployment and the behavior of gross job flows. The canonical versions are carefully constructed so as to permit analytical solutions and clear intuition, but the particular assumptions required will in some cases render the model less satisfactory both as a realistic portrayal of a given labor market and as a device for explaining some features of the data. In this paper, I forego the benefits of a closed-form model solution so as to more effectively deal with the problem of match quality heterogeneity.

The model is not solvable analytically, so I solve it numerically and simulate it in discrete-time, calibrated to a monthly frequency. The timing of the model is as follows. A period begins with any particular worker being either unemployed or employed. If employed, a worker-firm match is characterized by both a constant match productivity π and a constant amenity production parameter α.[3] Then, a time-varying idiosyncratic productivity shock x is drawn; this occurs every period. Employed individuals receive a wage and unemployed

[2]Blackwell (1965) and Sargent (1987).

[3]Though this paper is concerned with both the distribution of match quality draws and the actual distribution of accepted draws, "match quality distribution" refers to the former, unless otherwise specified.

individuals receive exogenous unemployment flow benefits. Next, an exogenous separation shock s may occur. If it does, unemployment results and the worker does not receive an employment offer until at least the subsequent period. If a separation does not occur, or if the worker was already unemployed, then a job-finding shock may occur. Firms and workers meet each other with a probability that depends only on the worker's employment status: α_0 for an unemployed worker and α_1 for an employed worker. The idiosyncratic productivity draw x occurs simultaneously with the match shock α, allowing workers to choose between unemployment and employment (the former being chosen if the productivity draw is such that the value of unemployment is higher), switching to a new job (occurring, for employed workers, if the surplus of the new match exceeds that of the present match, in which case switching costs are paid immediately), and remaining in the old job (occurring if an individual is already employed and the continuation value of the match exceeds that of all other alternatives). All job-finding shocks are characterized by a match quality draw and a new idiosyncratic productivity draw on which wages (fully flexible and instantly renegotiated) are based. Following this, the economy moves into the new state and the period ends. Note that the switching cost is considered to be a sunk cost for wage-setting purposes. This is consistent with the usual practice in most of the hiring cost literature, which can be thought of as analogous to the switching cost considered here.

The model timing is depicted by the following graph:

$\boxed{\text{wage or } b \text{ earned}} \Longrightarrow \boxed{s \text{ shock}} \Longrightarrow \boxed{\alpha \text{ and } x \text{ shock}} \Longrightarrow \boxed{\text{worker takes action}} \Longrightarrow \boxed{\text{period ends}}$

Idiosyncratic productivity draws are match-specific and time-varying, so it is possible for workers to switch to lower match quality jobs with sufficiently high productivity draws. x is drawn from a lognormal distribution with a persistence ρ_x, following the process $\ln(x') = \rho_x \ln(x) + \epsilon_x$, where $\epsilon_x \sim N(0, \sigma_x^2)$.[4]

[4] The x grid and discrete transition matrix are formed according to the Tauchen (1986) procedure.

The worker value function is defined by the following equation

$$W(m,x) = w(m,x,q) + q + \beta s U + \underbrace{\beta(1-s)(1-\alpha_1)Prob((U > W(m,x'))|x)\,U}_{\text{bad } x \text{ shock: separate}}$$
$$+ \underbrace{\beta(1-s)(1-\alpha_1)E[\mathbb{1}(W(m,x') \geq U)\,W(m,x')]}_{\text{no job offer: stay in job}}$$
$$+ \underbrace{\beta(1-s)\alpha_1 E[max\{W(m',x'_{nj}) - c, W(m,x'), U\}]}_{\text{continuation value conditional on new job offer}},$$

where w is the observed wage, x is the idiosyncratic shock to the current match, x'_{nj} is the idiosyncratic shock associated with a new job offer, and β is the discount factor. The latter enters the firm's value function linearly. Employed workers separate endogenously from their jobs, but they also suffer exogenous separation from employment with probability s. α_1 and α_0 are the on- and off-the job arrival probabilities, respectively. Workers do not accept all offers, whether they are initially unemployed or employed. Unemployed workers receive a flow benefit b. Firms and workers encounter one another at probabilities that depend only on employment status and are constant over time. Match quality m consists of two components: a monetary productivity term π that accrues to the firm and a non-monetary benefits term q that is entirely consumed by the worker. Further, q is not produced by firms; rather, it is endowed when a worker meets a firm. π draws are distributed according to a normal cdf Π with mean zero and standard deviation σ_π; q draws are distributed according to a normal cdf Q with mean zero and standard deviation σ_q. The two are assumed uncorrelated. The sum of these draws m then obeys a normal cdf M with mean zero and standard deviation $\sigma_m = \sqrt{\sigma_\pi^2 + \sigma_q^2}$. A one-time switching cost c is incurred by employed workers who accept new job offers.

The value of unemployment is similarly defined as

$$U = b + \beta(1-\alpha_0)U + \beta\alpha_0 E_{x',m'}[W(m',x'),U].$$

Note that job separations due to bad idiosyncratic draws are bilateral in the sense that joint surplus from the job is extinguished. This allows the job value function to be written largely in terms of the worker value function, because the employer sees fit to end a relationship under the same circumstances that motivate a worker to end the job. The job value function is given below.

$$J(m,x) = x + (m-q) - w(m,x,q) + \underbrace{\beta(1-s)(1-\alpha_1)E[\mathbb{1}(W(m,x') \geq U)J(m,x')]}_{\text{no job offer: worker stays in job}}$$
$$+ \underbrace{\beta(1-s)\alpha_1 E[\mathbb{1}((W(m',x'_{nj}) - c < W(m,x')) \cap (W(m,x') \geq U))J(m,x')]}_{\text{worker receives bad offer: stays in job}}.$$

Nash bargaining over the surplus yields the usual wage equation:

$$(1-\gamma)(W(m,x) - U) = \gamma J(m,x),$$

where γ is the fraction of job surplus going to the worker. The symmetric Nash equilibrium with $\gamma = 0.5$ is solved for throughout. While workers and firms do not care about the particular composition of m, and hence W and J take m rather than $\{\pi, q\}$ as an argument, the observed wage is a function of q.

Note that not all offers are accepted, which allows a distinction to be made between the offer arrival probability and the unemployment-employment transition probability. Without this distinction, endogeneity in the acceptance of offers would potentially bias estimation of the match quality offer distribution, as discussed previously.

I solve the model through value function iteration on a discrete grid, then simulate a panel of job spells. Except under very particular assumptions (as in Shimer 2003, for instance) that would make identification difficult or impossible, on-the-job search and the switching cost make it impossible to find a closed-form solution to the model.

The baseline model as described above makes use of the fact that π and q, the productivity and amenity components of match quality, are perfect substitutes. This allows workers and firms to care only about the sum $m = \pi + q$, which substantially simplifies the numerical solution of the model. During simulation, however, workers receive distinct draws of π and q (their distributions conforming to the distribution of m assumed by the model). Though workers and firms are indifferent between drawing $\{\pi_{low}, q_{high}\}$ or $\{\pi_{high}, q_{low}\}$, where $\pi_{low} + q_{high} = \pi_{high} + q_{low}$, the two draws do generate distinct observed wages. As will be discussed in the identification section, the correlation between tenure and observed wage allows for separate identification of the π and q distributions. In Section 2.5.1, the baseline model will have to be relaxed to take separate account of the two match quality components, because workers will now prefer to take compensation in the form of untaxed amenities.

2.4 Data

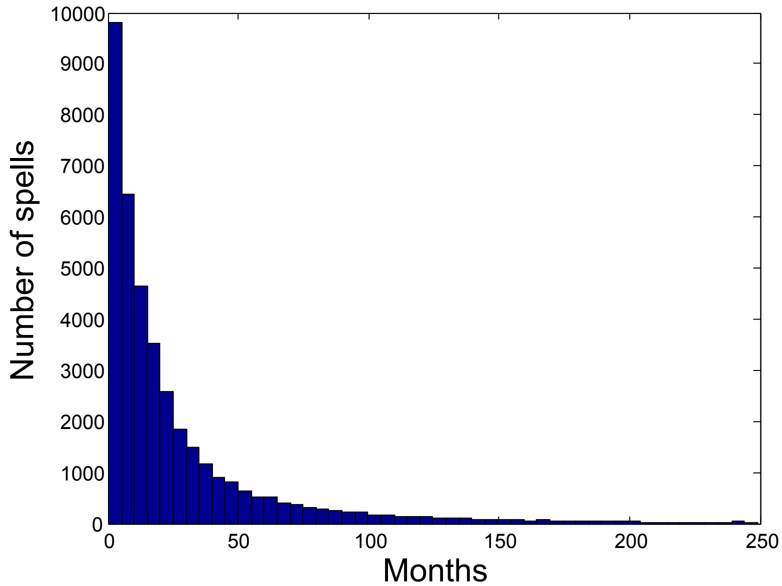
Data is taken from the 1979 National Longitudinal Survey of Youth. The NLSY79 is a nationally-representative panel of more than ten thousand young people at inception in 1979, with periodic successive surveys through the present. The detailed employment, demographic, and job spell data is necessary for this study, and the panel nature of the data permits comparison of this paper's results with those obtained from a more conventional wage regression approach, which requires person and match fixed effects.

I drop the military and supplemental subsamples. I also drop currently-enrolled high-school and college students from the sample. This leaves me with about 30,000 primary job spells. Wages are adjusted by the CPI-U to 2010 dollars. Table 2.1 gives summary statistics for relevant unweighted NLSY79 variables. Figure 2.1 shows the unweighted empirical tenure distribution.

Particularly notable is the fact that a surprising fraction of job-to-job wage changes are negative; fully one third of such wage changes are decreases, though measurement error is almost certainly exaggerating this figure.[5]

[5] Due to concerns about outliers, I use only the middle 95 percentiles of the wage data, making the same adjustment in the model's simulations so that only the trimmed data is generated. This mitigates the effect of error-ridden outliers, though it does not address measurement error more generally.

Figure 2.1: Unweighted empirical tenure distribution



Data are weighted to yield a nationally-representative sample before use in the model. I use the NLSY79's tenure variables corresponding to the primary jobs of respondents as well as various demographic variables that allow construction of a conditional tenure distribution. The tenure variables are constructed by the NLSY using worker-reported job start and stop dates, and are connected across waves of the survey by means of employer identification numbers. Demographic and other variables are dated to the end of each completed job spell. In addition, I set the simulated panel length equal to the average time a worker is present in my sample. This ensures that any truncation in the data (due to the requirement that spells terminate before the sample ends) is matched by truncation of the simulated spells.

I do not use the weighted empirical moments directly. Certain observable variables - age, education, etc. - explain some variation in job duration. Ex ante variation in worker characteristics is not part of the baseline model, so I prefer to adjust the data to more closely approximate the model's assumption of ex ante identical agents. To construct the empirical tenure variables, I first run the regression

$$\tau_{it} \;=\; X_{it}\beta + \epsilon_{it},$$

where τ_{it} is the tenure for a particular person's completed job spell beginning in year t, X_{it} is demographic information associated with a person in year t including age, education, sex, and race. I experimented with different specifications of industry and occupation dummies, but these made little difference to the results after the demographic variables were included. The time-varying variables are taken at the end of a job spell. Since the structural model is one of homogeneous agents who differ only ex post, I then construct tenure spells that are purged of observable variation due to age and other variables. The new tenure variable is given by $\hat{\tau_{it}} \;=\; \hat{\epsilon_{it}} + \overline{X}\hat{\beta}$, where $\overline{X}$ is the vector of population means corresponding to the demographic variables. Wages are adjusted in precisely the same way.

2.4.1 Identification

It is well-known that, in expectation, a ranking of jobs by duration corresponds to an ordering of jobs by match quality (Jovanovic, 1979, Hagedorn and Manovskii, forthcoming). However, the match quality ranking thus derived is ordinal and not cardinal, and so does not allow for an examination of match quality variation in relation to any other market quantity (though it does permit interesting cross-worker comparisons, as in Hagedorn and Manovskii (forthcoming)). The first contribution of this paper is a model in which the match quality distribution is identified by duration data, and parameters of the distribution are relatable to other market quantities. Intuitively, the shape of the tenure distribution identifies the model. If all job-worker pairings entailed the same match quality (and assuming no idiosyncratic shocks, for simplicity), workers would never switch jobs. Job endings would come only from exogenous separations and the induced tenure distribution would be exponential. If jobs were heterogeneous but switching costs were zero, workers would switch at every higher-match quality opportunity.

To better understand this, consider raising the variance of match quality while holding the mean of the offer distribution constant. One effect is an increase in the incentive to switch jobs. Since job switching is costly, the higher returns to switching "compress the tenure distribution". Put another way, a larger fraction of job offers will be sufficiently attractive to justify the hassle of moving to new employment, which will cause jobs to end sooner at all levels of match quality. Second, if the off-the-job arrival probability is higher than the on-the-job probability, the prospect of even better employment at the high end of match quality raises the reservation wage of the unemployed. In other words, why take a mediocre job when the payoff to waiting is much higher? Finally, a more dispersed match quality offer distribution will lead to a higher steady state average level of match quality. This last effect (higher dispersion implies higher average match quality) will be more pronounced when exogenous separations are unlikely and unemployment is generated primarily by negative idiosyncratic productivity shocks. In simulations at the estimated parameter vector, this effect is important in the following specific way. All else equal, *higher* match quality dispersion means that initial jobs (i.e., after a spell of unemployment) are likely to be farther from the reservation match quality. Fewer jobs of very short duration will be induced and destroyed by time-varying x shocks. At the estimated values of the parameters, this effect dominates for the mean of tenure. All these effects operate differently at various parts of the tenure distribution, which makes it essential to use multiple moments of the distribution.

It is also important to note that the unemployment-to-employment transition probability is required for identification of the match quality distribution. An economy with many workers briefly experiencing unemployment will tend to have a more compressed tenure distribution relative to an economy with a few workers experiencing long unemployment spells. The unemployment-to-employment transition probability, in conjunction with the tenure moments, pins down the match quality distribution that generates the observed pattern of job spells.

It is not immediately obvious why it it necessary to include job switching costs in the model. After all, simulated tenure distributions will still move with changes in match quality variance even when the switching cost is set to zero: though the "incentive to switch" effect disappears, workers with higher off-the-job arrival probabilities than on-the-job will still set their reservation wages as a function of the match quality distribution. This implies that tenure distributions will indeed vary with changes in the parameter of interest. However, the nature and size of this effect depends entirely on the values of α_0 and α_1, which are already pinned down by the empirical unemployment-to-employment and job-to-job transition probabilities. A stripped-down model with the switching cost restricted to be zero is not capable of generating the observed variation in tenure spells. The switching cost, in addition to being required for explanation of the data, has the added virtue of being consistent with observation.

The correlation between job duration and wage determines the fraction of match quality that is due to amenities q (and consequently the fraction due to productivity π). Imagine that match quality was entirely composed of π. In this case, the highest quality matches would be associated with the highest observed wages (ignoring time-varying idiosyncratic shocks for the moment), and the correlation between tenure and wage would be perfect. If, on the other hand, match quality consisted entirely of amenities q, then the observed wage would be a negative function of match quality (recall the examples presented in Section 2.2). The correlation between tenure and wages would then be perfectly negative. This does assume, perhaps counterfactually, that π and q are drawn independently.

Moments of the accepted wage distribution are useful for separately identifying the exogenous and endogenous components of the separation process. Recall that there exists an exogenous probability of leaving one's job, s, and an endogenous process of time-varying idiosyncratic shocks with standard deviation σ_x. The former is unrelated to the wage, while the latter will induce substantial wage variation. Intuitively, both s and σ_x affect the unemployment rate. Unemployment is clearly increasing in s, since higher inflows to unemployment will increase the stock, all else equal. Unemployment is also increasing in σ_x, because the probability of receiving a very negative x shock rises with σ_x. By itself, however, the unemployment rate cannot separately identify the two parameters. Since variance of wages is rising in σ_x but not in s, the inclusion of this moment provides for separate identification. Because I do not want to match the average wage in the data, but rather the unit-less dispersion of wages, I construct a moment equal to the standard deviation of wages divided by the average wage (i.e., the coefficient of variation).

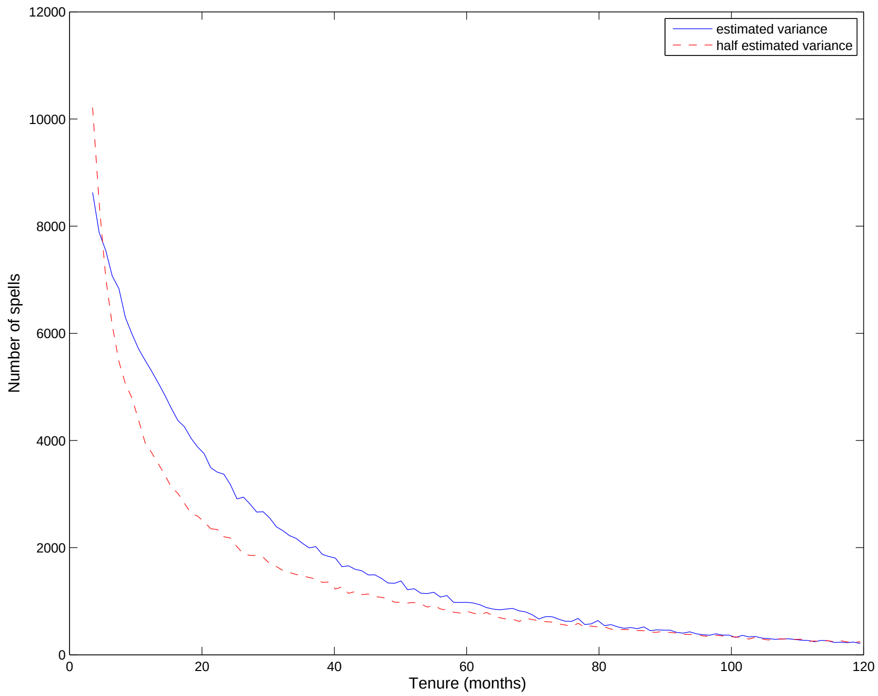
Due to the complexity of identification in this model, I have generated a number of artificial datasets and proceeded with estimation using the structural model with the intent of demonstrating that estimation would indeed recover the true parameters. This exercise also yielded intuition about the relative importance of the effects mentioned previously. For parameter values near those implied by my data, the "incentive to switch" effect dominates for the variance of tenure and the "higher dispersion implies higher average match quality" effect dominates for the mean of tenure. Figure 2.2 reflects this result.

2.4.2 Estimation

I estimate the model using simulated generalized method of moments (SGMM). This has the virtue of being robust to alternative specifications of the errors, unlike maximum likelihood estimation, though at some computational cost. Another advantage is that SGMM is consistent for a finite number of simulations, while simulated maximum likelihood is not.[6] As mentioned previously, it is necessary to use the simulated estimator due to some features of the model, most notably the switching cost and arrival probabilities that vary by employment status, that render it impossible to find analytic representations of the value functions and impossible to construct a closed-form relationship between the model's moments and the estimated parameters. The parameters $\{\sigma_\pi, \sigma_q, c, b, \alpha_0, \alpha_1, s, \sigma_x\}$ are estimated. At a monthly

[6] Ackerberg 2001.

Figure 2.2: Effect of match quality dispersion



frequency, ρ_x and β are calibrated to be 0.916 and 0.996, respectively (Fujita and Nakajima, 2009). σ_π and σ_q are the parameters of interest, and it is the NLSY duration data that pin down these parameters. The flow benefit to unemployment, b, is a controversial parameter. Some studies have put it as high as 68 percent of the mean flow output of a job, while others put it below 40 percent.[7] I set the $\frac{b}{\text{flow output}}$ moment to 50 percent of the average flow job output. The on- and off-the-job arrival probabilities are pinned down in steady state by the empirical unemployment-to-employment and job-to-job flows recorded in the data. The correlation of tenure and wage as well as the wage coefficient of variation are obtained from NLSY data, both using only wages associated with the last period of a job spell. Table 2.2 gives the values of moments, from the NLSY79 and other datasets, that were used during estimation, along with the values of the simulated moments generated by the model.

Figures 2.5 and 2.6 depict the sensitivity of particular moment conditions to deviations in the structural parameters from their estimated values. The bottom horizontal axis label

[7]Menzio and Shi (2008) and Shimer (2005), respectively.

gives the model parameter being manipulated while all other parameters are held constant at their estimated values, while the top label shows the simulated moment graphed on the vertical axis. The dashed vertical line is drawn at the estimated value of the parameter.[8] Moment-parameter pairs are chosen so to illustrate the source of identification for various parameters. In particular, the figures show that moments of the tenure distribution are primarily and jointly affected by the key model parameters of match quality variance and switching cost.

The simulated method of moments estimator is

$$\hat{\theta}_{S,A}(W) = \arg\min_{\theta} \left[\sum_{a=1}^{A} \left(\mu(x_a) - \frac{1}{S} \sum_{s=1}^{S} \mu(x(u_a^s, \theta)) \right) \right]' W^{-1} \left[\sum_{a=1}^{A} \left(\mu(x_a) - \frac{1}{S} \sum_{s=1}^{S} \mu(x(u_a^s, \theta)) \right) \right].$$

W is the weighting matrix, S is the number of simulations, A is the number of simulated agents, $\mu(x_a)$ is the vector of empirical moments, and $\mu(x(u_a^s, \theta))$ is the vector of simulated moments for a particular agent and for a given draw of the simulated errors. The choice of W is irrelevant to consistency of the estimator, though it has some effect on the efficiency of the estimator; I set W equal to the identity matrix. Because of the computational resources required and the negligible benefits, I do not implement efficient SGMM. Since the SGMM problem is typically characterized by some discontinuity, a Nelder-Mead simplex method is employed rather than a gradient-based approach for finding the minimum.

The simulated estimator converges in probability to the true θ as the number of agents approaches infinity:

$$\sqrt{A}(\hat{\theta}_{S,A} - \theta_0) \to N(0, Q_S(W)).$$

The covariance matrix of the vector of moment conditions is given by[9]

$$Q_S(W) = (1 + \frac{1}{S}) \left[E_0 \frac{\partial \mu'}{\partial \theta} W^{-1} \frac{\partial \mu}{\partial \theta'} \right]^{-1} E_0 \frac{\partial \mu'}{\partial \theta} W^{-1} \Sigma(\theta_0) W^{-1} \frac{\partial \mu}{\partial \theta'} \left[E_0 \frac{\partial \mu'}{\partial \theta} W^{-1} \frac{\partial \mu}{\partial \theta'} \right]^{-1},$$

[8] Simulation error is responsible for discrepancies between the simulated moment values shown here and in the tables.

[9] Gourieroux and Monfort (1996).

where

$$\Sigma(\theta_0) = E_0 \left[\left(\mu(x_a) - E_0\mu(x_a^s(\theta))\right) \left(\mu(x_a) - E_0\mu(x_a^s(\theta))\right)' \right].$$

As S approaches infinity, $Q_S(W)$ limits to the standard GMM variance-covariance matrix.

2.5 Results

The estimated values of the model's parameters are in Table 2.3.

The first thing to note from the results is that the standard deviation of the amenity component of match quality is about four fifths of the standard deviation of the productivity component. As discussed previously, this reflects the fact that tenure and wages are positively correlated. The standard deviation of overall match quality $\pi + q$ is 6.22, but this and the other parameter values are in terms of "model currency" and must be converted to current US dollars. Specifically, I multiply the figures by $\frac{\overline{w}_{emp}}{\overline{w}_{sim}}$, where $\overline{w}_{emp}$ is the average wage from the data and $\overline{w}_{sim}$ is the average simulated wage. Since I assume the correctness of the underlying model, the average simulated wage is equal to the average empirical wage. The standard deviation of overall match quality is then 12.19 dollars per hour (2010 dollars). This number is large, but keep in mind that it reflects the entire offer distribution of matches, a distribution with support over a wide range of matches that will never be chosen. However, the large magnitude does indicate that the returns to search are quite high, and that the typical employed worker would produce vastly disparate social surplus in different randomly chosen jobs. Unpriced and unobserved (by the econometrician) job amenities are quite large. Intuitively, the existence of these amenities seems particularly notable when one remembers that they are match and not job-specific. Large amenity estimates imply large social welfare gains from efficient coordination of workers and jobs.

I also normalize the standard deviation of accepted match quality to make it comparable with the data. Since match quality enters the job value function linearly, this allows for an interpretation of $\sigma_m^{accepted}$ in terms of flow surplus: a one-standard deviation increase in

$\sigma_m^{accepted}$ increases flow surplus by 8.05 dollars per hour, or 25 percent of average flow surplus. Note that estimates of the switching cost, while not a focus of this study, are plausible at nearly four times average monthly income.

Simulated and empirical moments are presented in Table 2.2. The model does not closely match two particular moments: the unemployment rate and the job-to-job transition probability. At 0.069, the observed NLSY job-to-job transition probability is quite high: a multiple of the monthly rate generally thought to obtain in the US labor market (see, for example, the 0.028 figure calculated using SIPP data by Nagypal, 2008). Our model predicts a substantially smaller quantity. This is essentially due to the large inframarginal rents implied by substantial match quality variation: many workers quickly move to matches with high match quality and then receive comparatively few offers that induce them to quit for either another job or unemployment. It is perhaps the case that a more flexible framework for match quality determination - amenity depreciation, for example, or separate offer distributions for unemployed and employed workers - would more effectively match both the observed tenure moments and job-to-job transition data.

2.5.1 Taxation and Match Quality

The division of match quality into productivity and amenity suggests an important consequence of the previous analysis: since job amenities cannot be taxed, wage taxation will impel workers to select into jobs with lower wages and more amenities. Estimation of the distortion generated on this margin is possible using a version of the model presented. It will now be necessary to add a state variable and track the distinct components of match quality.

The social welfare consequences are interesting and somewhat subtle. First, recall that a relatively simple model of labor supply implies that the taxable income elasticity is a sufficient statistic for the deadweight loss of an income tax (Feldstein, 1999). Workers adjust their hours and participation in response to a tax, but they also evade the tax, alter their consumption of deductible/tax-excluded consumption, and so forth: all of which are captured by changes in taxable income but not always by labor supply changes. Consider the following decision problem, taken with slight modifications from Feldstein (1999).

Let workers maximize the utility function $U(L, C, Q, E)$ subject to the constraint $C = (1 - \tau)[w(1 - L) - Q - E]$.

L is leisure, C is "ordinary" or taxable consumption, Q is the value of job amenities, E is all non-taxed consumption aside from amenities, w is the pre-tax wage, and τ is the rate of wage tax. Feldstein's insight was that a wage tax, in this setting, causes an increase in the price of ordinary consumption relative to all other goods, but no change in any other relative prices (e.g., leisure and tax-excluded consumption). This implies that the deadweight loss from the tax is a function only of the elasticity of taxable income, assuming a few other conditions not relevant to this paper, like the absence of fiscal and classic externalities.

Labor search and heterogeneity in match quality create another exception to the original result. Intuitively, available matches differ in their relative prices of leisure and amenity. The ability to choose a different, more amenity-intensive match in response to a tax means that a wage tax changes all relative prices, not just the price of ordinary consumption. Leisure actually becomes more expensive relative to the amenity, for instance. In the short run, a new tax τ will only change the relative price of C; as in Feldstein's setup, initial deadweight loss will be proportional only to the taxable income elasticity. In the long run, when workers are able to move to new jobs, the size of the distortion will depend on the match quality offer distribution and especially the amenity offer distribution.

For this reason, if standard measures of deadweight loss are predicated upon short-run estimates of taxable income elasticities, deadweight loss will be calculated to be lower than if long-run taxable income elasticities were employed.[10]

The search setting considered in this paper is different than the Feldstein (1999) problem,

[10]This is discussed in Saez et al. (2009), but the difference here is heterogeneity in match quality. This creates another margin on which distortion occurs only belatedly, exaggerating the short-run/long-run distinction.

which is competitive and has fixed wages. Though this paper's model does not include an hours margin, one can think of the unemployment-to-employment reservation wage as being analogous to the intensive margin in the modified Feldstein problem. The exercises conducted in this section, informed by estimates from the baseline model, help to illuminate the social welfare effects of taxes. Details of the search process modify the connection between the taxable income elasticity and the elasticity of social welfare.

Relative to an (otherwise-identical) economy with no match quality variation, the deadweight loss of a tax is higher in my baseline model. In the no variation economy, the only channel through which wage taxes reduce welfare is the unemployed workers' reservation wage (and consequently the unemployment rate). With match variation, workers are impelled by the tax to select higher-amenity, lower-wage jobs, which generates its own distortion.[11]

Inclusion of wage taxation requires an elaboration of the baseline model presented above. Assuming that the tax is paid by workers, the value functions are now

$$\begin{aligned} W(\pi,q,x) = (1-\tau)w(\pi,q,x) + q + \beta s U + &\underbrace{\beta(1-s)(1-\alpha_1)Prob((U > W(\pi,q,x'))|x)\,U}_{\text{bad } x \text{ shock: separate}} \\ &+ \underbrace{\beta(1-s)(1-\alpha_1)E[\mathbb{1}(W(\pi,q,x') \geq U)\,W(\pi,q,x')]}_{\text{no job offer: stay in job}} \\ &+ \underbrace{\beta(1-s)\alpha_1 E[max\{W(\pi',q',x'_{nj}) - c, W(\pi,q,x'), U\}]}_{\text{continuation value conditional on new job offer}}, \end{aligned}$$

where τ is the wage tax rate and other parameters are defined as before: w is the observed wage, x is the idiosyncratic shock to the current match, x'_{nj} is the idiosyncratic shock associated with a new job offer, and β is the discount factor. s is the exogenous separation probability, and α_1 and α_0 are the on- and off-the job arrival probabilities, respectively. Un-

[11] The Nash bargaining assumption of this paper's model will enhance the taxable income elasticity relative to the simpler model discussed in Section 2.2. In the simpler model, workers were paid their monetary productivities. In the baseline model, workers and firms bargain over job surplus, and both know that this includes an amenity. Observed wages in jobs with higher amenity levels will be lower, which increases the incentive for a worker to bypass income tax through selection of higher-amenity jobs.

employed workers receive a flow benefit b. Firms and workers encounter one another with probabilities that depend only on employment status and are constant over time. Match quality m consists of two components: a monetary productivity term π that accrues to the firm and a non-monetary benefits term q that is entirely consumed by the worker. Further, q is not produced by firms; rather, it is endowed when a worker meets a firm. π draws are distributed according to a normal cdf Π with mean zero and standard deviation σ_π; q draws are distributed according to a normal cdf Q with mean zero and standard deviation σ_q. The two are assumed uncorrelated. The sum of these draws m then obeys a normal cdf M with mean zero and standard deviation $\sigma_m = \sqrt{\sigma_\pi^2 + \sigma_q^2}$. A one-time switching cost c is incurred by employed workers who accept new job offers.

The unemployment value function is

$$U = b + \beta(1-\alpha_0)U + \beta\alpha_0 E_{x',\pi',q'}[W(\pi', q', x'), U].$$

The job value function is now

$$\begin{aligned} J(\pi, q, x) =& x + \pi - w(\pi, q, x) + \beta(1-s)(1-\alpha_1)E[\mathbb{1}(W(\pi, q, x') \geq U)J(\pi, q, x')] \\ &+ \beta(1-s)\alpha_1 E[\mathbb{1}((W(\pi', q', x'_{nj}) - c < W(\pi, q, x')) \cap (W(\pi, q, x') \geq U))J(\pi, q, x')]. \end{aligned}$$

Social welfare is the steady state flow surplus $x+m$ associated with a job, summed over all agents, less the switching costs paid. Taxable income is the sum of pre-tax wages received in a particular period: $\sum_i^I (\mathbb{1}(employed_{it} = 1) \cdot (w_{it}))$. Both social welfare and taxable income are normalized to 100 in the absence of taxation. Note that job switching costs are not considered to be tax-deductible. A substantial portion of switching costs are non-monetary, and many workers in the data do not itemize their deductions in any event. Social welfare, taxable income, and corresponding tax elasticities are given in Table 2.4 and illustrated in Figure 2.3.

The tax distortion is highly nonlinear in the level of the tax. While a ten percentage point tax increase leaves welfare almost unchanged when starting from an untaxed economy, a ten percentage point tax increase from 40 to 50 percentage points causes a reduction in social welfare of about 2.5 percentage points. The deadweight loss at a 40 percent tax rate is 3.0 percentage points of social welfare.

Recall, however, that the original Feldstein result suggests that the response of taxable income to the tax rate is sufficient to understand social welfare consequences. It is clear that this result is violated in the model considered here, but perhaps not clear in which direction the bias runs: are we likely to be overstating or understating the deadweight loss of a tax in the presence of match-specific amenity variation?

To answer this question, I conduct the following experiment. Divide the standard deviation of offered amenity variation, σ_q, by two and increase the standard deviation of offered productivity variation, σ_π, by enough to hold σ_m constant. This allows us to isolate the effect of a change in the extent of amenity variation. Now I generate the same taxable income and social welfare curves shown previously in Table 2.4. If the econometrician underestimates amenity variation in this manner, Table 2.5 and Figure 2.4 show that social welfare and taxable income will be thought to be approximately identical. In reality, however, taxable income is increasingly *overstating* the social welfare decline as σ_q rises.

This may seem counterintuitive, since the introduction of amenities leads to an additional distortion: the endogenous mix of amenity and monetary productivity responds to taxation, moving away from optimal levels as the tax increases. However, taxable income will respond more dramatically. Imagine an arbitrarily small tax $d\tau$ applied to an economy that is currently implementing a socially optimal equilibrium. It is clear that this tax generates no deadweight loss, since the no-tax allocation is efficient. In terms of social welfare, any reduction in wage income is matched by an increase in utility from the amenity. By the same token, however, the tax will generate a reduction in taxable income. This is indeed what we observe in Figure 2.3.

2.5.2 Application to Optimal Unemployment Insurance

Unemployment insurance is one of the most important policies for which estimates of match quality and returns to search are directly relevant. In an economy with no labor market comparative advantage (no match quality heterogeneity) and risk neutral workers, unemployment insurance would serve only to tax labor and reduce employment below the optimal level. Various considerations complicate the calculation of optimal UI, not least of which is the heterogeneity of match quality. This section contains an exercise that is informative about the relationship between match quality and the optimal unemployment insurance benefit.

Acemoglu and Shimer (1999) show that unemployment insurance can induce higher productivity by allowing workers to search for higher-quality jobs. Several authors have conducted reduced-form analyses of changes in unemployment insurance law, with mixed evidence of an effect of UI (unemployment insurance) generosity on subsequent job duration. Ours and Vodopivec (2008) examine a change in Slovenian law and see no significant effect on subsequent tenure. Belzil looks specifically at UI benefit duration, rather than the level of the benefit, and finds that "increasing the maximum benefit duration by one week will raise expected unemployment duration by 1.0 to 1.5 days but expected job duration by 0.5 to 0.9 days only." (p.635, Belzil (2001)) Mario Centeno, in a number of papers, finds substantial post-unemployment job duration effects (2004, 2006, and 2009).

The baseline model unambiguously predicts that more generous unemployment insurance, implemented via a higher flow benefit to unemployment b, induces higher match quality and longer job spells, albeit at the cost of lower aggregate employment. The advantage of the structural approach is that it is possible to speak precisely about UI welfare effects. The model shows the effect a change in b has on the entire accepted distribution of match quality, which in conjunction with a social welfare function is sufficient to find the full welfare consequences. Note that amenities are unimportant in this experiment, in contrast to the tax application. The match quality offer distribution is the object of interest.

It is possible, however, that match quality variation is sufficiently negligible as to allow it to be ignored in the setting of unemployment insurance policy. The exercise shown below is evidence against this possibility.

Social welfare in a steady state period is given by

$$\sum_{i}^{I}(\mathbb{1}(employed_{it}=1)(x_{it}+m_{it}))-\sum_{i}^{I}(\mathbb{1}(jobswitch_{it}=1)c\,.$$

Socially optimal b values are calculated by assuming all the other estimated parameters, then maximizing (through simulation) the expression above with respect to b. The assumption of risk neutrality is unrealistic, though consistent with the permanent income hypothesis. Accordingly, the "optimal" UI benefits calculated here should be understood primarily as a useful way to isolate the effects of match quality dispersion. Introducing an insurance motive to the model would complicate this.

The UI benefit chosen by policymakers operates through the reservation wage (for movement from unemployment) and the threshold wage required for a job switch. Higher benefits induce both higher unemployment and higher average match quality, since workers optimally search longer when the cost is reduced.

This paper conducts an exercise in which estimated variation in match quality is halved and socially optimal UI benefits compared between the two cases. The primary tradeoff in this economy is between reduced output from added unemployment (associated with higher b) and worse job mismatch (associated with lower b). Workers do not internalize all the benefits of search because firms receive a portion of the rents from employment. Results are shown in Table 2.6.

The socially optimal flow benefit to unemployment more than doubles when the standard deviation of match quality doubles to the estimated value. Perhaps more informatively, the ratio of the optimal UI benefit over average flow surplus rises by 14 percentage points. This gives an indication of the importance of accurate match quality measurements to unemployment insurance policy.

2.5.3 Comparison with Conventional Match Quality Estimates Based on Wages

The estimates shown above come from a very different model than most that have been tasked to address match quality. They are likely to differ for many reasons that have already been discussed, but comparisons of the results may be instructive.

Discussion of labor market comparative advantage typically begins with wage variation. In an economy with no non-monetary job amenities, a fixed-effects analysis incorporating all relevant time-varying factors will be sufficient to reveal the fraction of wage dispersion due to match quality variation. However, since only accepted wages are observed, the match quality distribution implied by this analysis is the accepted match distribution and not the offered. This is unfortunate in that the underlying, unobserved offer distribution is the more important object for several purposes, including an understanding of the returns to search and optimal unemployment insurance. An unemployed worker will set a reservation wage that is a function of the wage offer distribution, not the accepted wage distribution.

The model presented in this paper is capable of recovering both the accepted and offer distributions, which permits a comparison between the results thus obtained and the results produced by a wage regression approach. In finite samples, however, the latter method will overestimate the extent of match quality heterogeneity, even assuming away non-monetary amenities. This is because persistent, time-varying idiosyncratic productivity shocks - incorporated in the model as the log AR(1) process $\ln(x') = \rho_x \ln(x) + \epsilon_x$, $\epsilon_x \sim N(0, \sigma_x^2)$ - generate wage variation that only limits to zero as the number of observations per match limits to infinity. Wages are increasing in both m and x, and a persistent, high x shock at the outset of one job will raise the average wage paid relative to an otherwise identical, low x shock job taken by the same person. With a typical persistence $\rho = 0.9$, the half-life of a shock, in logs, is about 6.5 months, so it is not the case that productivity shocks are decaying too quickly to meaningfully affect the average wage at a job. In simulated panels generated at the estimated parameters, I find a substantial small-sample positive bias to the wage regression estimates of match quality variation. This bias is strongly increasing in the persistence ρ.

Using the wage regression approach mentioned previously, I calculate the component of wage variation due to match quality variation from the same NLSY79 data. The wage regression is of the form:

$$w_{it} = X_{it}\beta_X + \theta_i + \mu_{ij} + \epsilon_{it},$$

where w is the observed wage, X_{it} is a vector of demographic characteristics, θ_i is an individual fixed effect, μ_{ij} is a match-specific fixed effect, and ϵ_{it} is a time-varying error term.

In addition to the previously-discussed weaknesses of the wage-based approach, there is one more practical problem with extracting a match quality estimate from wage information. Wages are highly skewed even after controlling for observables and individual fixed effects, in part due to what appear to be coding errors in the NLSY79. Measurement error in wages creates difficulties for any approach to estimating match quality that relies principally on wage data.

As described above, I calculate the standard deviation of μ_{ij}, with wages trimmed at the cutoffs[12] of \$0.1 and \$1000, to be \$14.18 per hour. However, with more aggressive trimming that keeps only the middle 95 percentiles, the same figure is \$6.53 per hour. This suggests that the treatment of measurement error and outliers in a wage decomposition is quite important, and potentially makes the results sensitive to the assumptions made. For comparison, recall that the standard deviation of accepted match quality was estimated to be \$8.05 per hour. Note, however, that the estimates are not strictly comparable, in that the structural model produces estimates of match quality proper, while the wage regression produces estimates of the (wage) variation ascribable to match quality. Under certain assumptions these may be identical, but not under the assumptions of the model.

[12] Barlevy (2008).

2.6 Related Literature

This project is built on two literatures and two seminal papers: labor search (Mortensen and Pissarides, 1994) and job duration as match quality (Jovanovic, 1979). Mortensen and Pissarides provide the basic framework for the random search model employed in this paper, albeit without on-the-job search and a few other modifications. Working within the Mortensen and Pissarides class of model, Shimer (2006) adds on-the-job search in a partial equilibrium context, with job offers arriving at an exogenous rate constrained to be the same on and off the job.

Jovanovic (1979) and related papers like McCall (1990) provide a theoretical basis for the assertion that tenure is informative about job match quality, embedding match quality variation in an equilibrium model of job turnover. A large, mostly empirical literature has developed that takes tenure as a proxy for match quality and examines, for example, job mismatch over the business cycle (Bowlus, 1995) and the effects of changes in unemployment insurance law (Ours and Vodopivec, 2008). Papers in the latter category are discussed briefly in Section 2.5.2. Some authors have pursued related questions with identification or calibration strategies making use of duration data. Nagypal (2007), for instance, distinguishes accumulation of human capital from learning about match quality using, in part, matched firm-worker data including tenure information. Becker (2009) focuses on job amenities and finds that they are quantitatively substantial.

Paul Sullivan and Ted To assess the relative importance of non-wage and wage job utility in a 2011 working paper. They estimate the offer distributions of wages and non-wage utility in a search context, and use the fraction of job switches associated with wage declines to identify the distribution of non-wage utility. Although they are not concerned with estimating the distribution of match quality, like Becker, Sullivan and To find non-wage utility to be substantial.

Though this paper is not explicitly about compensating differentials, separate estimation of productivity and amenities, along with the particular wage bargain assumed, implicitly involve tradeoffs between wages and amenities. As is intuitively the case, workers in jobs with low amenities will (holding match quality constant) receive higher wages. The baseline model makes use of wages to help separate productivity and amenities, which puts it in debt of papers like Rosen (1974) and Friedman and Kuznets (1954). Unlike most compensating differentials papers, however, this paper does not assume equality of utility across options and the focus is on parameters of the aggregate match quality distributions rather than any cross-sectional tradeoffs.

A large theoretical and empirical literature has developed around the explanation and decomposition of wage dispersion in an on-the-job search context. Postel-Vinay and Robin (2002) construct a variant of the Burdett and Mortensen model that they use to decompose wage variation into person, firm and "market friction" contributions. The latter does not include match quality variation, as it is simply the wage dispersion endogenously generated by the Burdett-Mortensen structure. Mortensen (2003) discusses similar explanations for wage dispersion. Hagedorn and Manovskii (2010) take a less structural approach, estimating the variance of match quality from wage data in a multi-step procedure that subtracts the contributions of tenure, experience, fixed effects, and within-job wage shocks, obtaining match quality variation as a residual.

On-the-job search models typically predict that all job changes will be associated with increases in observed wages. However, this is decidedly not the case in the data[13], and various solutions have been proposed. Of course, some or all of wage declines associated with job-to-job transitions can be interpreted as measurement error. In the NLSY79, for instance, that some reported wage declines are spurious is explicitly noted by the survey administrators.[14] Wolpin (1987) and related literature explicitly incorporate measurement error into their models. It is difficult to explain the extent of job-to-job transitions that

[13]See Sullivan and To (2011) for data from the NLSY97. Postel-Vinay and Robin (2002) find a similar result in French data, with a third to a half of workers reporting wage decreases as they change jobs.

[14]From NLSY79 documentation at `http://www.nlsinfo.org/nlsy79/docs/79html/79text/wages.htm`: "Note that: the calculation procedure, which factors in each respondent's usual wage, time unit of pay, and usual hours worked per day/per week produces, at times, extremely low and extremely high pay rate values; no editing of values reported by a respondent occurs even if the value is extreme, such as $25,000 per hour..."

involve wage declines with measurement error alone, considering, for instance, that a full third of job-to-job wage changes in the NLSY79 are reductions. Another approach is to construct some aspect of the model that leads workers to (occasionally) optimally switch to lower-wage jobs. As an example, Postel-Vinay and Robin (2002) and Cahuc et al. (2006) use a bargaining arrangement that allows workers to choose lower-wage but higher-productivity firms that will in the future be able to better match outside offers. The present paper, though not principally motivated by the concern of wage reductions in job-to-job transitions, provides a different answer to this question. When unpriced amenities are accounted for, optimal wage choice is such that workers will frequently choose lower-wage jobs that are nonetheless preferred to previous jobs due to their superior amenities.

2.7 Future Work

A large theoretical and empirical literature has developed around the question of cyclical variation in productivity and labor match quality. The model developed above could be used to evaluate the competing theoretical models describing the cyclical evolution of match quality. The degree of labor market specialization is expected to vary cyclically; this variation is informative about the size and timing of the welfare costs of recessions. Using the match quality offer distribution estimated previously, I can simulate the effect of a recession on average and marginal realized match quality at various lags, tracing out the so-called "cleansing" and "sullying" effects of recessions (Barlevy, 2002). The cleansing effect is relatively well-known, as it refers to the Schumpeterian process of creative destruction operating in the labor market. Recessions lower the joint surplus to employment across the board, which destroys the least productive jobs and impels workers and firms to search for better matches. In this way, a recession may immediately raise average labor productivity. The "sullying" effect is a more recent coinage that refers to the ongoing and gradual *reduction* in average labor productivity generated by the reduced search effort expended during a recession. As workers move up the match quality ladder more slowly during a period of low aggregate productivity, average labor productivity gradually diminishes.

Another avenue for future research is accounting for ex ante individual heterogeneity. I plan to add ex ante effects to the model. This will serve the dual purposes of making explicit the robustness of σ_m estimation to individual fixed effects, as well as generating an endogenous positive correlation between π and q, which is more in keeping with observation. Identification of these ex ante differences will require a more extensive use of NLSY79 wage moments.

In the final chapter of this dissertation, coauthored with Brendan Epstein, we elaborate on the theory of the wage tax distortion, the relationship between match quality and social welfare with endogenous amenity supply, and implications of the dynamic aspect of the wage tax distortion, among other things. Search is not required to generate the distortion on the amenity margin, which exists in more general settings. We plan to examine a model with rigid wages but flexible amenity supply, which may generate interesting results consistent with stylized facts of the business cycle.

2.8 Appendix

The model's objects are recovered from a simulated steady state in which all the objects are stationary. I simulated 40,000 agents over 724 periods, the latter being selected to match the average duration of a subject's participation in the NLSY79 sample. As with the data, the simulated panel does select on job spells that end prior to the last month generated. Note that exogenous separations are helpful for speeding the transition to steady state in models with on-the-job search; workers with higher match quality are less likely to receive a sufficiently negative idiosyncratic shock to induce unemployment, and exogenous shocks keep the job quality ladder from becoming too top-heavy.

Two key grids were constructed: one for time-invariant job-specific match quality, and one for time-varying idiosyncratic productivity. The latter was defined on a grid with $N_x = 9$, with the grid and the transition matrix generated according to the Tauchen (1986) algorithm, but modified to provide a grid of equiprobable (not equidistant) points.[15] This modification is somewhat more efficient, as it does not waste time with grid points that are unlikely to be reached through the log AR(1) process. The second important grid is that of match quality, with $N_m = 12$. I interpolate during simulation rather than dramatically increasing the grid size. The matrices used to form expectations over possible jobs, for purposes of deriving the value function, are in general several-dimensional arrays with size increasing exponentially in N_x and N_m, so the computational problem quickly becomes intractable as the grids become finer.

As previously mentioned, value function iteration is employed and the workers simulated to follow the associated policy functions. Value functions $W(m, x)\, U$,and $J(m, x)$ along with a wage grid $wgrid(m, x)$ are initialized for all values of the support, with $wgrid$ simply a deterministic function of the value functions, assuming Nash bargaining. This is possible because the wage enters linearly into both the worker and firm value functions, allowing the $i - 1$ iteration of the wage to be subtracted (added) from the worker (firm) value function

[15] Ryan Michaels, correspondence.

before calculating the ith iteration: $wgrid' = \gamma(J + W - U) - (W - U) + wgrid$. Value functions are repeatedly generated, assuming the correctness of the previous iteration of value functions, until the discrepancy across iterations becomes small.

During simulation, records are made of every relevant variable. Random shocks are drawn before simulation for all of the stochastic variables and held constant as model parameters change. Very few simulations are required to substantially reduce error relative to the GMM estimator.[16] For computational reasons, I generate only two simulated panels for each value of the parameter vector, which is enough to obtain precise estimates when the panel is sufficiently large. I tried several different weighting matrices; results were very similar under each.

[16] Ackerberg (2001).

Table 2.1: Summary statistics

	Mean[1]
Age	29.5 years
Sex	49.7% female
Race	7.6% Hispanic, 13.1% black, 79.3% non-black,non-Hispanic
Education	12.6 years
Tenure	30.2 months
Wage	$15.8/hour

1. Data are from the NLSY79.

Table 2.2: Simulated and empirical moments

Moments	**Simulated**	**Target**
mean tenure	30.3	30.5
variance of tenure	1419	1419
skewness of tenure	2.34	2.29
b as a fraction of flow output	0.417	0.500
unemployment rate	0.043	0.065
job-to-job probability	0.012	0.069
unemployment-to-employment probability	0.267	0.193
wage tenure correlation	0.140	0.225
wage coefficient of variation	0.461	0.489

Table 2.3: Model parameters

	Estimated parameter
σ_π	4.84 (0.1138)
σ_q	3.93 (0.0294)
c	32.86 (0.0004)
b	6.82 (0.0912)
α_1	0.80 (0.0108)
α_0	0.88 (0.0075)
s	0.01 ($2.3*10^{-7}$)
σ_x	0.54 (0.0001)

Table 2.4: Effects of wage taxation

Wage tax	Social welfare	Taxable income	Social welfare elasticity[1]	Taxable income elasticity
0 percent	100.0	100.0		
10 percent	99.6	92.2	0.04	0.77
20 percent	98.9	89.1	0.05	0.29
30 percent	98.1	86.7	0.06	0.20
40 percent	97.0	83.4	0.08	0.26
50 percent	94.5	75.4	0.14	0.55
60 percent	91.9	67.8	0.12	0.47
70 percent	87.8	56.5	0.16	0.63
80 percent	82.2	41.9	0.16	0.74
90 percent	75.9	26.1	0.12	0.68

1. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.
2. Elasticities are with respect to the net-of-tax rate.

Table 2.5: Effects of wage taxation with reduced amenity variation

Wage tax[1]	**Social welfare**[2]	**Taxable income**	**Social welfare elasticity**[3]	**Taxable income elasticity**
0 percent	100.0	100.0		
10 percent	99.3	98.3	0.06	0.16
20 percent	98.1	96.7	0.10	0.14
30 percent	96.7	96.2	0.11	0.04
40 percent	94.8	95.4	0.13	0.05
50 percent	91.9	92.7	0.17	0.15
60 percent	86.1	86.5	0.29	0.31
70 percent	76.9	76.1	0.39	0.45
80 percent	55.1	48.9	0.82	1.09
90 percent	46.0	39.4	0.26	0.31

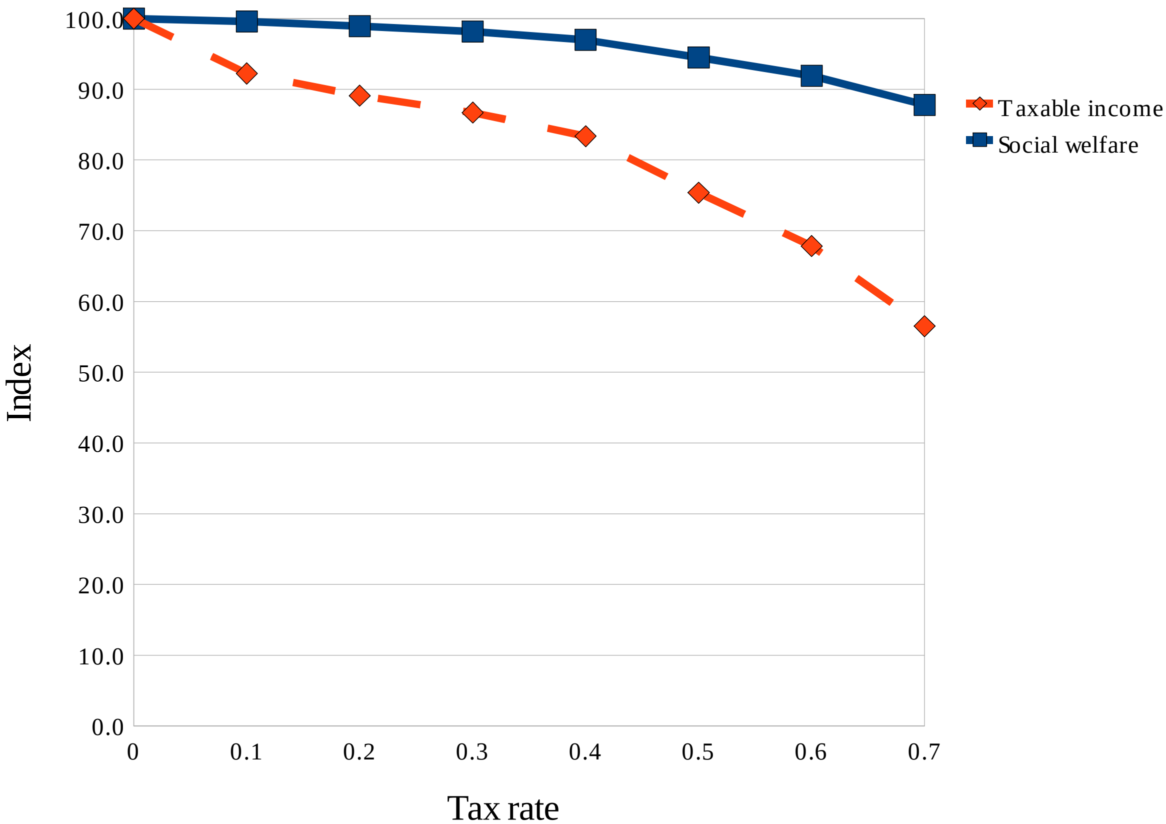
1. Standard deviation of amenity variation is halved relative to baseline.
2. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.
3. Elasticities are with respect to the net-of-tax rate.

Table 2.6: Optimal unemployment insurance

	Estimated match quality	**Half estimated match quality**
Optimal UI benefit	7.5	3.0
UI benefit ratio[1]	45.7%	31.5%

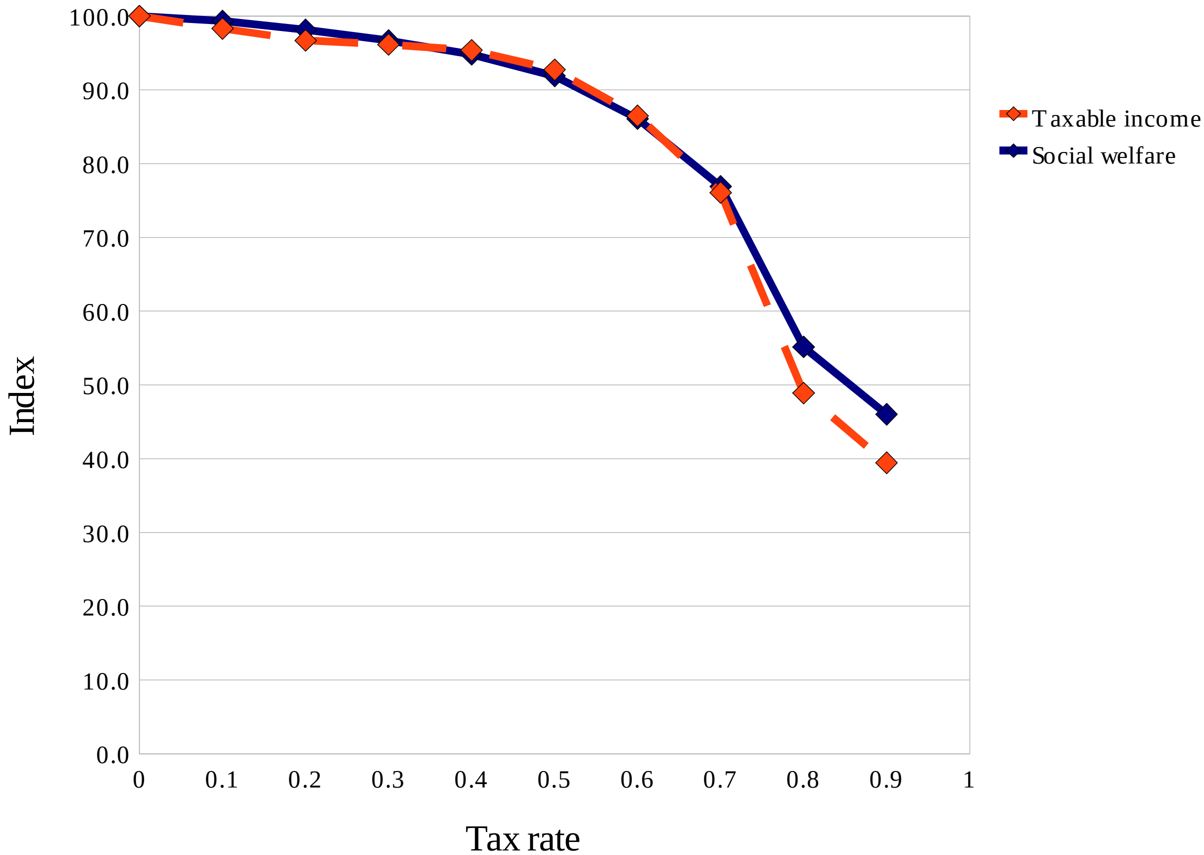
1. The benefit ratio is the UI benefit divided by the average flow output.

Figure 2.3: Effects of wage taxation



1. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.

Figure 2.4: Effects of wage taxation with reduced amenity variation



1. Standard deviation of amenity variation is halved relative to baseline.
2. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.

Figure 2.5: Sensitivity of simulated moments to parameters

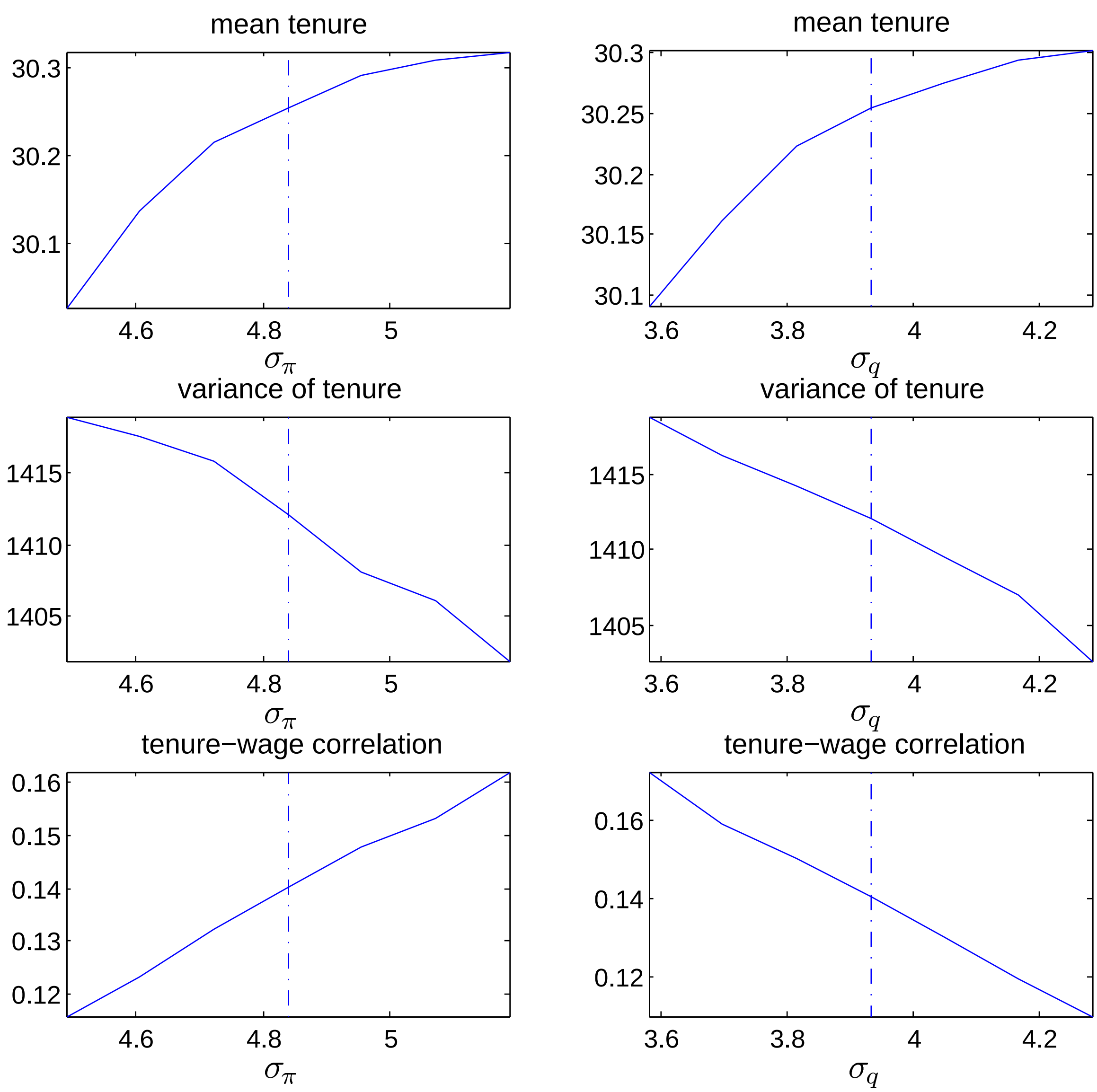
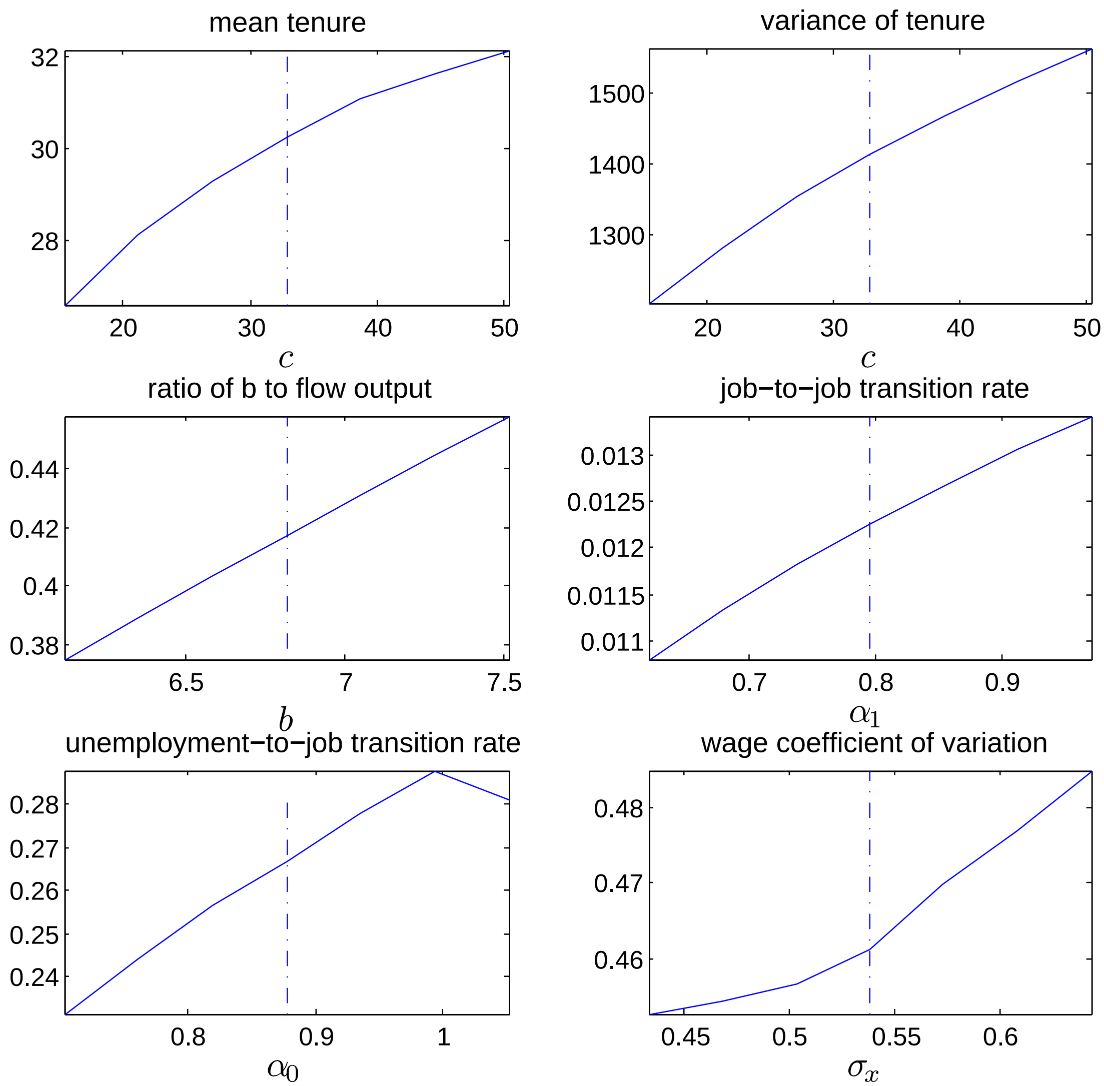


Figure 2.6: Sensitivity of simulated moments to parameters



CHAPTER III

Accounting for Adaptation in the Economics of Happiness

3.1 Introduction

In recent years, researchers have increasingly turned to survey data on subjective well-being (happiness) to investigate a wide range of economic questions. Happiness data have been used to study preferences over inflation and unemployment (Di Tella et al., 2001), the consequences of excise taxes (Gruber and Mullainathan, 2005), and the progress of women (Stevenson and Wolfers, 2009), to take just a few examples. The interpretation of responses to survey questions about happiness varies with the study. Some have taken these measures of recent mood or life-satisfaction as indicators of utility (e.g., Oswald and Powdthavee, 2008). Others have viewed happiness measures as outcomes of intrinsic interest analagous to health or income (e.g., Deaton et al., 2008).

While, in practice, the use of happiness data is often controversial, the idea of measuring an individual's recent mood or life-satisfaction has important potential. Taken simply as an outcome, it is natural to view an individual's emotional state or level of satisfaction as, at least, an important aspect of her well-being. Taken as an indicator of utility, happiness data provide an opportunity to infer an individual's preferences under circumstances, such as the presence of externalities, non-market goods, or cognitive biases, where choices and prices alone will typically be inadequate. The potential value of these measures motivates the adaptation and development of tools for rigorous economic analysis of happiness data.

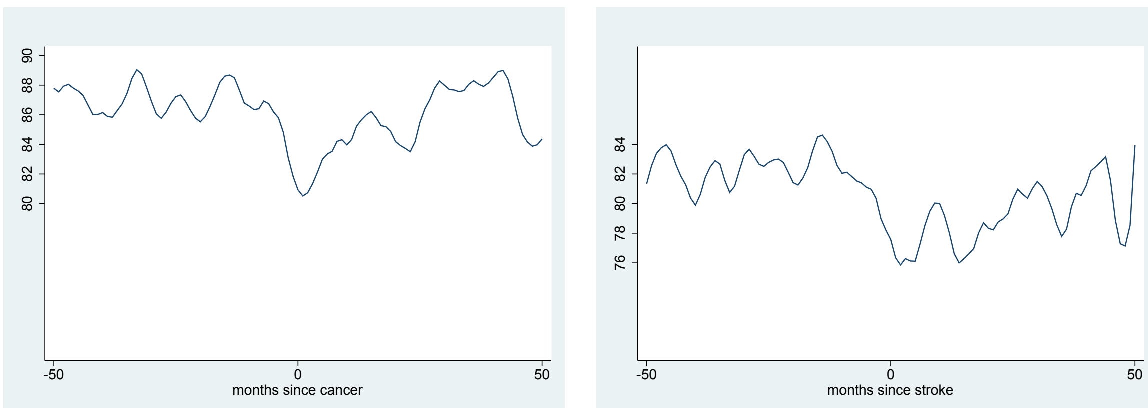
This paper is concerned with a feature of happiness data that has not, as yet, been thoroughly accomodated by economic analysis: "hedonic adaptation." Hedonic adapation refers to a dynamic response of happiness to changes in life-circumstances; the magnitude of the response decreases as the change fades into the past. In a variety of studies, using a variety of methods, evidence often indicates that happiness responds in expected ways to the arrival of both good and bad events, but individuals return to their prior levels of mood with surprising thoroughness and speed.[1] In a canonical example, people are much less happy upon the arrival of a serious health problem (paraplegia, renal failure) but eventually appear to adapt and reveal measured happiness at or near normal levels (Riis et al., 2005). The idea that happiness quickly adapts to changes in income has been central to investigations of the Easterlin Paradox, but hedonic adaptation has not otherwise been thoroughly incorporated into economic analysis of happiness data.

Much of economists' work in happiness economics has been at a macro level, exploring topics like unemployment/inflation tradeoffs (Di Tella et al., 2001) and making use of country-level data like the Gallup World Poll (Deaton et al., 2008). The same methods, when applied to individual-level data, become more controversial. Even if happiness self-reports are a useful proxy for utility, regressions making use of them are subject to the usual endogeneity concerns. The panel nature of our data helps us to address two concerns of this kind. First, one might imagine that news about subjective well-being-relevant events is likely to arrive prior to an event's occurrence. The econometric strategy developed here is able to accommodate this possibility. In addition, our analysis incorporates person-specific fixed effects. Cross-sectional individual-level data does not permit this, and selection will likely be important. Furthermore, this paper focuses on high-frequency SWB variation around events that are plausibly unexpected by survey respondents. This all gives us more confidence that

[1]Frederick and Loewenstein (1999) offer a review of this evidence. Diener et al. (2006) review evidence on the limits of hedonic adaptation.

happiness consequences are properly identified.

The innovation of this paper is in the treatment of individual adaptation to SWB-relevant events. Previous work has largely ignored adaptation or dealt with it in an insufficiently general way. This paper demonstrates that events have widely differing hedonic consequences, both in magnitude and in composition (e.g., temporary vs. permanent effects). Analyses that assume all hedonic consequences to be permanent will necessarily err in the comparison of events that differ in the composition of their effects.



(a) cancer **(b)** stroke

Figure 3.1: Two health events in the HRS

Figure 3.1 illustrates some of the differences in hedonic adaptation that require explicit consideration. The graphs depict smoothed SWB measurements (ranging $0-100$), averaged across all respondents, at varying temporal distance from event occurrence. First, note that reported well-being drops prior to the month in which illness occurs (anticipation). Second, while the overall drops in happiness are roughly comparable across the two events, respondents report persistent happiness reductions in the wake of a stroke, while most of the initial happiness drop after cancer diagnosis is regained in later months. To the extent that the hedonic response to an event contains a temporary component, the rate of recovery may also be of interest.

The paper proceeds as follows. First, the related literature is discussed in Section 3.1.1.

The econometric approach taken in previous work is described in Section 3.1.2, and limitations of the identification are discussed. The model is then presented in Section 3.2, along with multiple variants of the baseline specification, and in Section 3.2.2 a procedure for comparing our results to previous work is developed. In Section 3.3 and Section 3.4, data and results are then presented, with a discussion and conclusion following.

3.1.1 Related Literature

The subjective well-being literature was advanced by early papers like Andrews and Withey (1976), Bradburn (1969), Campbell et al. (1976) that explored the potential usefulness of survey questions on happiness. The use of SWB data has been controversial, however. Schwarz and Strack (1999) and Schwarz (1987) describe some of the difficulties with interpretation of reported happiness data, including specifically the temporary and context-dependent nature of the reports. It is also the case that SWB is not always best conceived of as a straightforward proxy for utility; Benjamin et al. (2010) show evidence for systematic discrepancies between SWB and revealed preference utility. Kimball and Willis (2010) develop a theory that reconciles the two.

Hedonic adaptation has been the focus of substantial recent work[2]. If adaptation and relative income are important in the determination of SWB, a number of implications follow for cross-sectional individual and national data, as well as optimal taxation, consumption, and other topics (Clark et al., 2008b). Concerns about "set points" (SWB levels to which individuals eventually return after a disturbance) and the measurement of happiness are explored in Diener et al. (2006). Clark et al. (2008a) looks for permanent SWB responses to various events, generally not finding significant effects. Bottan and Truglia (2011) present evidence indicating that happiness itself is positively autocorrelated. Oswald and Powdthavee (2008) and Powdthavee (2009) investigate adaptation in the wake of disability and widowing. They acknowledge the importance of hedonic adaptation but emphasize the mutability and

[2]For a review of some of the literature on the topic, see Frederick and Loewenstein (1999).

heterogeneity of set points.

The choice of SWB measure may matter for estimates of hedonic adaptation and the decomposition of SWB effects into temporary and permanent components. Life satisfaction in the German Socioeconomic Panel, as used by Headey et al. (2010) and Lucas et al. (2004), occasionally behaves differently than happiness variables we observe in the HRS.

3.1.2 Previous work

Several authors estimate the impact of life events on happiness using a cross-sectional regression that includes indicator variables for various events. In particular, Deaton et al. (2008) use the following specification with individual-level data from the Gallup World Poll:

$$H_{it} = \alpha_c + X_{it}\beta_X + Y_{it}\beta_Y + \mathbf{1}_{it}\beta_{\mathbf{1}} + \epsilon_{it}, \tag{3.1}$$

where H is a SWB measure, α_c is a vector of country fixed effects, X is a vector of demographic variables, Y is log income, and $\mathbf{1}$ is a vector of dummies for the death of a family member from a given disease in the last twelve months. Without the possibility of individual fixed effects, the authors rely on the assumption that "baseline mood" (i.e., the SWB level reported before an event) is not systematically different across respondents in a way that is endogenous to the specification.

Other papers exploit country-level data. As mentioned previously, Di Tella et al. (2001) run regressions of the following form.

$$LS_{ct} = \alpha_c + U_{ct}\beta_X + \Pi_{ct}\beta_I + \epsilon_{ct},$$

where LS is a modified life satisfaction measure, U is the unemployment rate, Π is the inflation rate, and α_c is a vector of country fixed effects. Finkelstein et al. (2008) utilize individual panel data and estimate the following equation, among others:

$$H_{it} = \alpha_i + X_{it}\beta_X + Y_{it}\beta_Y + \mathbf{1}_{it}\beta_{\mathbf{1}} + (\mathbf{1}_{it} * Y_{it})\beta_{\mathbf{int}} + \epsilon_{it}, \tag{3.2}$$

where H is a SWB measure, α_c is a vector of country fixed effects, X is a vector of demographic variables, Y is log income, and $\mathbf{1}$ is a vector of dummies for whether a respondent has ever had a particular disease. The effect of the interaction between income and health events is given by $\beta_{\mathbf{int}}$. However, this panel approach still fails to distinguish immediate consequences from enduring effects. Because there is evidence that adaptation occurs, we prefer a method that allows for dynamic happiness responses. Further, we find that life events vary significantly in the ratio of temporary to permanent happiness effects and dynamic methods are necessary to properly compare them.

3.2 Model

The primary motivation for our baseline specification is a desire to address hedonic adaptation in a flexible and tractable fashion. A nonparametric approach would make excessive demands on the (often sparse and incomplete) data, but a pooled cross-sectional or fixed effects regression fails to recognize the substantial and nonlinear adaptation characteristic of the data. Our specification is nonlinear but quick to estimate. Furthermore, as will be demonstrated below, the method generally returns reasonable results.

3.2.1 Baseline specification

Our dynamic approach can be interpreted in an event-study framework. News contained in the occurrence of important life events will have an immediate effect on self-reported happiness, one component of which will be permanent and the other temporary. The temporary component is assumed to decay exponentially at a rate that is specific to the type of event (i.e., different decay rates are estimated for heart attacks and strokes).

The simplest form of our dynamic equation is given below. In essence, it decomposes the cumulative happiness response into an immediate, temporary effect that decays exponentially, and a permanent effect that persists indefinitely. The temporary effect is assumed to vanish exponentially at rate δ.[3] Since many life events have important consequences for income or are related to income levels, we include the log level of household income as a covariate. The estimating equation is given by

$$H_{it} = \underbrace{\alpha_i}_{\text{fixed effect}} + \underbrace{Y_{it}\beta_Y}_{\text{income effect}} + \chi_i \mathbf{1}(t >= t_0)[\underbrace{\beta_P}_{\text{permanent effect}} + \underbrace{\beta_T e^{-\delta(t-t_0)}}_{\text{temporary effect}}] + \epsilon_{it}, \tag{3.3}$$

where H is a SWB measure ranging $0 - 100$, t is the time that happiness is observed, t_0 is the time the event occurs, β_Y is the income effect, β_P is the permanent effect, β_T is the temporary effect, α_i is the person fixed effect, and δ is the rate of decay of the shock. Since δ enters the equation nonlinearly, we use nonlinear least squares (NLS) estimation. The NLS estimator is given by

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} [y_i - f(x_i;\theta)]^2,$$

where $f(x_t;\theta)$ is the nonlinear model, y is the endogenous variable, N is the number of observations, and θ is the parameter vector.

With some SWB-relevant events, it may be the case that the likelihood of occurrence is related to baseline mood. This may be the case even if the event is unanticipated by the respondent. For instance, health problems such as heart attacks may be induced by stress, which could itself imply lower baseline mood. Since the healthy respondents report higher SWB, the effect of a heart attack will be over-estimated with a specification that fails to account for the already-lower baseline mood of respondents who are about to have heart attacks. To account for these differences in baseline mood, we include individual-specific fixed effects.

[3] We experimented with less restrictive specifications but found no evidence of non-exponential decay.

3.2.2 Comparability with previous estimates

For some purposes, it may be interesting to consider the cumulative impact of an event. In particular, some of the estimates in previous work (along the lines of equations 3.1 and 3.2) have an interpretation as the cumulative SWB consequence of an event. For comparability with these results, we adapt our baseline specification, equation 3.3, by integrating to find the total effect. For an individual with d annual mortality risk and interest rate r, the "area under the curve" will have the following form:

$$\int_{t_0}^{\infty} (\beta_P e^{-(d+r)(s-t_0)} + \beta_T e^{-(\delta+d+r)(s-t_0)})ds = \frac{\beta_P}{d+r} + \frac{\beta_T}{d+r+\delta} \tag{3.4}$$

This formulation gives a single statistic that can be used to rank events by their hedonic importance. It also allows for comparison of our dynamic results with the previous literature's static estimates, since both are measures of a cumulative hedonic effect. Following previous work, the fixed effects regression

$$H_{it} = \alpha_i + Y_{it}\beta_Y + \mathbf{1}_{it}\beta_{\mathbf{1}} + \epsilon_{it}, \tag{3.5}$$

is conducted to make this comparison of our results with the usual linear specification. $\mathbf{1}_{it}$ is an "absorbing state" indicator, which means that it is set to 1 for all observations after the initial event occurrence. This is consistent with some of the previous SWB literature (e.g., Finkelstein et al., 2008) and aims to capture a cumulative SWB effect. Were $\mathbf{1}_{it}$ to equal 1 only in the initial event occurrence observation, $\beta_{\mathbf{1}}$ would capture only a portion of the temporary SWB consequences, and would not be comparable to our cumulative results.

The chief virtue of our baseline specification is its ability to separately identify temporary and permanent SWB effects. While important in its own right, mean reversion also complicates estimation of the cumulative happiness effect. In equation 3.5, $\beta_{\mathbf{1}}$ is only identical to β_P in equation 3.3 if $\beta_T = 0$. For $\beta_T > 0$, equation 3.5 may imply a different estimate of the cumulative SWB effect than equation 3.3. Consider the ordinal ranking of SWB-relevant events by $\beta_{\mathbf{1}}$. Since the specification of equation 3.5 makes no use of time since occurrence, this parameter will depend on the probability that a respondent subsequently leaves the sample, which may differ across events. For example, an event that is substantially mean-reverting may be associated with subsequent SWB observations over many years. Another event, with the same balance of temporary and permanent effects, may have relatively few subsequent SWB observations. Assume, for specificity, that both temporary and permanent effects are negative in sign. For the event with few post-occurrence observations, $\beta_{\mathbf{1}}$ will be estimated to be larger in magnitude, because the temporarily depressed SWB values just after event occurrence will dominate the data. By contrast, β_T, β_P, and by implication $\beta_{cumulative} = \frac{\beta_T}{d+r+\delta} + \frac{\beta_P}{d+r}$ do not depend on these factors, but are estimated consistently if the underlying model is correct. Rankings of SWB-relevant events based on the latter expression will then be different, in general, than a ranking based on $\beta_{\mathbf{1}}$.

3.2.3 Allows for various sorts of decompositions

One interesting extension of this method involves subjective life expectancy. Some life events, in particular health events like heart attacks and strokes, will in general have important consequences for life expectancy. The original dynamic specification, equation 3.3, can be modified to examine the role of subjective life expectancy changes in creating the observed hedonic effects. A particular implementation is given below:

$$H_{it} = \alpha_i + Y_{it}\beta_Y + \chi_i \mathbf{1}(t >= t_0)[\beta_P + (\beta_T + \eta \Delta SLE)e^{-\delta(t-t_0)}] + \epsilon_{it}, \tag{3.6}$$

where SLE is a constructed measure of subjective life expectancy, denominated in years, and η is the temporary effect of changes in subjective life expectancy.

3.2.4 Anticipation

In order for our interpretation of the baseline specification to be correct, it must be the case that the "news" component of event happiness consequences does not occur prior to the event itself. In other words, respondents must not learn of and (in terms of subjective well-being) react to an event that has yet to occur. Since this assumption is likely violated in a number of cases, we include as a control a dummy for SWB in the six-month period prior to an event. This is done for all the paper's results. If news about the event generates a change in SWB prior to the event date, our modified specification will correctly capture the consequences of the event itself (as opposed to news of the event) in the parameters β_P, β_T, δ.

3.2.5 Sparse data

Another advantage of our method is that it can be easily modified to handle infrequently-measured data. For instance, with all the HRS data, we posit an underlying continuous-time data generating process, with the happiness data only observed periodically (every two years in the core of the HRS). Some of the SWB-relevant events are dated precisely to the month in which they occur, others are known only to the calendar year in which they occurred (with extra information for same-year events coming from the fact that they must be before the survey date), while still others can only be dated as occurring sometime between waves of survey data. To deal with this, we assume a uniform distribution of the logically possible interval of time in which an event could have occurred given the data. We time-aggregate the equations for the continuous-time data generating process to obtain a nonlinear estimating equation. These equations posit that an event will have a permanent effect on happiness (which we interpret as a permanent change in baseline mood) and a transitory effect on happiness (which we interpret as the dynamics of elation). The key identifying assumption is that there are no important transitory movements in baseline mood after an event.

For instance, events that are dated only to the year are estimated by the following equation.

$$H_{it} = \alpha_i + Y_{it}\beta_Y + \chi_i \mathbf{1}(y >= y_0)[\beta_P + \beta_T e^{-\delta(t-t_0)} * (\mathbf{1}(y > y_0) * \frac{e^{\delta - 1}}{\delta} + \mathbf{1}(y = y_0) * \frac{e^{\delta * \frac{m}{12} - 1}}{\delta})] + \epsilon_{it}, \tag{3.7}$$

where y is the year happiness is measured, y_0 is the year an event occurred, and m is the interview month in which happiness is measured. For events that are dated only to the wave,

$$H_{it} = \alpha_i + Y_{it}\beta_Y + \chi_i \mathbf{1}(y > y_0)[\beta_P + \beta_T e^{-\delta(t-t_0)} * \frac{e^{(t_{w_1} - t_{w_0})\delta - 1}}{\delta}] + \epsilon_{it}, \tag{3.8}$$

where t_{w_1} is the time of the interview directly after an event occurred and t_{w_0} is the time of the interview directly before.

3.2.6 Recall bias

With some life events, respondents might suffer from imprecise recollection of the event's date. A reasonable assumption is that this recall bias is an increasing function of the time since event occurrence. Given this assumption, a simple correction consists of multiplying β_T by the term $e^{\gamma(t_r - t_0)}$, which (for negative values of γ) is a decreasing function of the time elapsed between event and first recall. This correction yields the adjusted estimating equation:

$$H_{it} = \alpha_i + Y_{it}\beta_Y + \chi_i \mathbf{1}(t >= t_0)[\beta_P + \beta_T e^{-\delta(t-t_0)} e^{\gamma(t_r - t_0)}] + \epsilon_{it}, \tag{3.9}$$

where t_r is the time the event is first recalled.

3.3 Data

We use data from the Health and Retirement Study (HRS), which conducts a biennial representative survey of Americans over the age of 50. The resulting panel data spanning the years 1992 through 2010 includes detailed happiness reports and information on a rich set of important life events. Although there are important panel data sets for subjective well-being for other countries (most notably the German Socioeconomic Panel and the British Household Panel Study), for the U.S., the HRS is the only survey with a long panel of repeated observations on subjective well-being. In addition to the core HRS waves, we use the Asset and Health Dynamics Among the Oldest Old (AHEAD), the Children of the Depression Age (CODA), and the War Babies cohorts[4]. For some of the variables, we use a version of this data provided by the RAND Center for the Study of Aging that includes some additional imputations.

Each wave of the HRS asks respondents the following questions: "Now think about the past week and the feelings you have experienced. Please tell me if each of the following was true for you much of the time this past week: a) You felt you were happy b) You felt sad c) You enjoyed life d) You felt depressed." We treat the binary variables "happy" and "enjoylife" based on these data, along with the reverse-coded "notsad" and "notdepressed" as four indicators of the underlying latent value of happiness at the time of the interview. Thus, we treat the probability of answering in the positive direction for each of these variables as an increasing function of latent happiness.

HRS respondents are questioned about many health and other important life events in each wave of the survey. Some of these events are precisely dated to the month but some are only known to have occurred at some point between waves. Examples of the latter include episodes of incontinence, congestive heart failure, hip fractures, cataract surgery, births of grandchildren, and changes in social isolation. Widowing, heart attacks, strokes, cancer, retirement, unemployment, and entry into nursing homes are dated to the precise month. The Psychosocial Leave-Behind (PLB) component of the HRS provides retrospective information on a number of other events with dating only to the year: death of a child, natural disaster, combat, drug and alcohol addiction of family members, physical assault, serious illness, serious illness of a family member, labor and financial market discrimination,

[4]The AHEAD cohort was initially part of a distinct study and includes respondents born before 1924. CODA and the War Babies cohorts were added in 1998 and includes respondents born 1924-1930, and 1942-1947, respectively.

police discrimination, firing, job search, changes in neighborhood safety, robbery, and others.

A measure of household income is available for all waves, which we include as a control throughout. The HRS includes life insurance status, allowing us to partition the hedonic response to widowing. Interestingly, there are also questions about life expectancy for many of the respondents. We use these to construct a measure of subjective life expectancy, then decompose the temporary response to an event into a component related to changes in subjective life expectancy and a residual, which becomes the standard β_T coefficient.

The subjective life expectancy (SLE) measure is constructed in the following way. For certain waves, respondents are asked what probability they assign to their living to a particular age, where said age depends on the current age of the respondent. We linearly interpolate the survival probability for all future years, then calculate SLE as the expectation of years remaining.

3.4 Results

Because the HRS is a panel, we are able to conduct fixed effects estimation. Without individual fixed effects, any heterogeneity in baseline happiness would bias estimation of the temporary effect and the decay rate.

In preliminary work, we compared multiple approaches to the use of HRS subjective well-being variables. As discussed in Section 3.3, the HRS provides four closely-related variables pertaining to a respondent's mood at the time of interview. Our preferred approach uses the sum of the (appropriately-coded) variables as the dependent variable in all our specifications. We obtained broadly similar results when conducting probit estimation with individual SWB variables.

Table 3.1 gives estimates for β_P, β_T, and δ using our dynamic method for events dated to the month (widowing, heart attack, stroke, cancer, unemployment, nursing home entry, and retirement). Standard errors are all heteroskedasticity-robust.

Cancer, heart attacks, and strokes all follow a somewhat different pattern. We estimate a significant negative permanent effect in all three cases, with somewhat larger temporary than permanent effects. Interestingly, strokes have both the largest permanent effect, relative to temporary, as well as the slowest decay of the temporary setback. The temporary effect of cancer is about one-third as large as the temporary effect of widowing for those without life insurance. A relatively small but statistically significant negative permanent effect is found with cancer as with the other two major health events, heart attacks and strokes. Strokes induce roughly double the permanent effect found for heart attacks and cancer, which is plausibly consistent with the enduring disability often suffered by stroke victims.

Regressions concerning the other events for which we have precise timing information - unemployment, nursing home entry, and retirement - yield reasonable estimates. Unemployment has a negative temporary effect in line with the magnitude of the aforementioned health problems, but no significant permanent effect. Entry into a nursing home produces a similar temporary effect, but also a large negative permanent reduction in SWB. Retirement produces almost no SWB effect. Both nursing home entry and retirement are perhaps less plausibly exogenous than the other events, and it may be that correlated factors are driving the estimates.

The estimated depreciation rates are themselves of interest. For all the precisely-dated events, with the exception of widowing (without life insurance), heart attacks, and cancer, estimated depreciation rates are such that the half-life of an event is between about 9 and 13 months. That is, half of the temporary shock associated with an event has disappeared after this time. Widowing (without life insurance) has a half-life of 5 months, while heart attacks and cancer have half-lives of 4 and 3 months, respectively.

Figure 3.2a-3.2e illustrates all of this graphically, showing the predicted values of both a non-parametric regression and an "impulse response" corresponding to our baseline specification for events dated to the month. The impulse responses are constructed by assuming the estimated β_T, β_P, δ and applying them to population means of all variables. In other words,

$$\hat{H}_t = \overline{H}_b + \chi_i \mathbf{1}(t >= 0)[\beta_P + \beta_T e^{-\delta(t)}], \tag{3.10}$$

where $\overline{H}_b$ is the mean pre-event SWB level and $\hat{H}_t$ is the predicted SWB value. This yields a graphical representation of the temporary and permanent effects at work in the regressions.

Widowing is of particular interest, as widows experience an unusually large reduction in SWB, all of which appears to be temporary. Individuals with life insurance suffer roughly half as large a drop in SWB when compared to those without insurance, though the latter are estimated to recover more quickly from their (larger) fall. This is depicted, both parametrically and non-parametrically, in Figure 3.3. In both cases, we identify a very large effect that is almost entirely temporary.

Table 3.2 compares results from equation 3.5 and our dynamic method. The first three columns show estimates (identical to Table 3.1) for $\beta_{T,}$ $\beta_{P,}$ and δ, respectively. The fourth, fifth, and sixth columns give the cumulative temporary, permanent, and total effects, whose construction was described previously. The mortality rate d is estimated off of the unconditional mean mortality rate in our sample and is approximately 0.02, while the interest rate assumed is 0.05.

The "Lost Area" columns, 4, 5, and 6 (so-called because they show the "area under the curve" associated with the hedonic response to an event), condense the information provided in the baseline specification, giving measures of cumulative SWB consequences. High depreciation rates render the cumulative temporary effects of heart attacks and stroke relatively small in magnitude, while magnifying the permanent effects by comparison. Total SWB loss is substantially larger for all health events than it is for widowing, largely because of small but enduring permanent positive effects estimated for widowing.

We also implement equation 3.5, the linear regression described in Section 3.1.2. The final column displays $\beta_{\mathbf{1}}$ from that regression. These results are generally signed consistently with the Total Lost Area quantities, but relative magnitudes differ. Recall that **1** is an indicator that toggles on permanently after an event. $\beta_{\mathbf{1}}$ registers larger magnitude effects for widowing than for health events, with the most negative result coming from nursing home entry at -7.5.

Table 3.3 presents results that utilize data on subjective life expectancy for the same life events. This is an example of the flexibility of our approach, which facilitates any decomposition of SWB effects permitted by the data. The first four columns provide the usual parameter estimates plus η, the effect of a one-year increase in subjective life expectancy on the temporary hedonic effect. The final three columns give estimates from the previous specification for comparison.

In all cases where we find a statistically significant result, the effect of η is as expected: an event that increases life expectancy will increase SWB through this channel. The magnitude of these life expectancy effects is small, however, and the available data on life expectancy are crude.

Table 3.4 gives estimates of β_P, β_T, and δ for events that are dated relatively imprecisely: either to the year or to the wave. In the former case, these are retrospective data from the Psychosocial Leave-Behind survey. Because these estimates are based on retrospective data, there are sometimes long gaps between event occurrence and recollection. For this reason, we implement the recall-bias adjustment described previously. The data available for these events is often quite poor, which is reflected in the occasional inability of the NLS procedure to identify a δ significantly above zero. In these cases, the data does not permit the separation of temporary and permanent SWB effects, and consequently β_T and β_P values should likely be disregarded.

3.5 Discussion and Conclusion

In Section 3.2.2, conditions were described that may produce discrepancies in the SWB ranking (i.e., the ordering by magnitude of the "Total" cumulative effect) based on our baseline specification, and on a regression that does not incorporate SWB mean reversion. Table 3.2 suggests some concrete examples. Interestingly, within the three major health events we consider (heart attacks, strokes, and cancer), our cumulative SWB ranking mirrors the ranking generated by equation 3.5. However, when comparing widowing with any of the major health events, we see a different pattern. Using our approach, the health events all have larger cumulative SWB consequences than widowing, though significantly smaller initial (temporary) effects. Regression 3.5, on the other hand, yields the opposite result: widowing, with or without life insurance, has an effect of greater magnitude.

This illustrates that the temporary/permanent decomposition of SWB data is sometimes a first-order concern. The overall utility consequence of an event, good, or experience is quite sensitive to this decomposition, since temporary effects are generally fairly quick to decay. Without arbitrarily-frequent and indefinitely-extended panels of SWB measurements, a regression specification that is not sensitive to hedonic dynamics will typically generate errors, particularly when comparing events of dissimilar dynamic profiles. If one of the aims of happiness economics is to inform public policy about relative valuations of events and non-market goods, insensitivity to dynamic effects will compromise that project.

The approach presented here will be useful in a variety of ways. Work that aims at pricing non-market goods, many of which take on characteristics of durable goods and produce a changing flow of utility, will benefit from the explicit treatment of dynamic effects. When data is infrequently-collected, retrospective, or otherwise limited, our approach will facilitate the extraction of usable information. Since SWB variables have only recently been added to some datasets, the ability to handle retrospective data is likely to be useful.

Table 3.1: Baseline results

Parameter Estimates

	b_T	b_P	δ	N
Widowing (w/ insurance)	-21.96***	1.533**	0.763***	3042
	(2.552)	(0.648)	(0.158)	
Widowing (w/o insurance)	-40.41***	1.319	1.657***	1417
	(6.223)	(1.007)	(0.408)	
Cancer	-7.242***	-1.302***	2.598***	23205
	(1.335)	(0.327)	(0.802)	
Heart attack	-5.154***	-1.178***	2.059*	13168
	(1.526)	(0.426)	(1.060)	
Stroke	-3.374***	-2.296***	0.888	13304
	(1.120)	(0.556)	(0.583)	
Unemployment	-6.304***	0.861*	0.631***	12232
	(0.911)	(0.508)	(0.194)	
Entered nursing home	-7.525***	-4.665***	0.876*	8871
	(1.818)	(1.189)	(0.477)	
Retired	-0.602	0.559***	0.787	79359
	(0.512)	(0.179)	(1.094)	

1. Dependent variable is the (0-100) index of happiness equal to 25(sum of the 4 indicators of recent mood). See the text for a description of the indicators.*

2. δ are annual rates of recovery.

3. Standard errors are in parentheses.

4. Events are all dated to the month.

Table 3.2: Comparison of event study results with pooled results

	Parameter Estimates			Lost Area			
	b_T	b_P	δ	Temp	Perm	Total	Pooled coefficient
Widowing (w/ insurance)	-21.96***	1.533**	0.763***	-26.37***	22.06*	-4.303	-3.515***
	(2.552)	(0.648)	(0.158)	(3.245)	(11.77)		(1.004)
Widowing (w/o insurance)	-40.41***	1.319	1.657***	-23.40***	18.99	-4.415	-4.943***
	(6.223)	(1.007)	(0.408)	(4.043)	(17.06)		(1.758)
Cancer	-7.242***	-1.302***	2.598***	-2.715***	-18.73***	-21.45	-2.045***
	(1.335)	(0.327)	(0.802)	(0.501)	(4.699)		(0.317)
Heart attack	-5.154***	-1.178***	2.059*	-2.421***	-16.95***	-19.37	-1.850***
	(1.526)	(0.426)	(1.060)	(0.714)	(6.153)		(0.411)
Stroke	-3.374***	-2.296***	0.888	-3.524***	-33.04***	-36.57	-3.416***
	(1.120)	(0.556)	(0.583)	(1.174)	(7.946)		(0.448)
Unemployment	-6.304***	0.861*	0.631***	-8.997***	12.39*	3.395	-1.763***
	(0.911)	(0.508)	(0.194)	(1.305)	(7.378)		(0.507)
Entered nursing home	-7.525***	-4.665***	0.876*	-7.962***	-67.15***	-75.11	-7.529***
	(1.818)	(1.189)	(0.477)	(1.921)	(17.24)		(0.756)
Retired	-0.602	0.559***	0.787	-0.703	8.040	7.337	1.145***
	(0.512)	(0.179)	(1.094)	(0.626)	(9.201)		(0.192)

1. Dependent variable is the (0-100) index of happiness equal to 25(sum of the 4 indicators of recent mood). See the text for a description of the indicators.*

2. δ are annual rates of recovery.

3. Standard errors are in parentheses.

4. Events are all dated to the month.

5. Pooled coefficients are from a fixed effects regression of happiness on log household income and an indicator for whether the event has ever occurred previous to or in the current wave.

6. Area measures are based on an interest rate of .05 and a constant mortality rate of .019, the latter of which is the unconditional annual death rate in the entire sample

Table 3.3: Results with and without subjective life expectancy

	With Life Expectancy				W/o Life Expectancy		
	b_T	b_P	δ	η	b_T	b_P	δ
Widowing (w/ insurance)	-22.47***	1.892	0.700***	-0.0594	-21.96***	1.533**	0.763***
	(4.086)	(1.359)	(0.243)	(0.0771)	(2.552)	(0.648)	(0.158)
Widowing (w/o insurance)	-39.51***	0.950	2.435**	0.136**	-40.41***	1.319	1.657***
	(10.89)	(1.875)	(1.137)	(0.0599)	(6.223)	(1.007)	(0.408)
Cancer	-7.096***	-0.440	1.955***	-0.0143	-7.242***	-1.302***	2.598***
	(1.576)	(0.417)	(0.755)	(0.0238)	(1.335)	(0.327)	(0.802)
Heart attack	-8.089***	-0.623	2.744**	0.103**	-5.154***	-1.178***	2.059*
	(2.645)	(0.591)	(1.370)	(0.0487)	(1.526)	(0.426)	(1.060)
Stroke	-5.300***	-2.027**	0.798*	0.0556*	-3.374***	-2.296***	0.888
	(1.826)	(0.970)	(0.485)	(0.0308)	(1.120)	(0.556)	(0.583)
Unemployment	-10.05***	1.696*	0.645***	0.0977***	-6.304***	0.861*	0.631***
	(1.549)	(0.967)	(0.211)	(0.0291)	(0.911)	(0.508)	(0.194)
Entered nursing home	-8.663**	-5.079**	0.759	0.361*	-7.525***	-4.665***	0.876*
	(3.911)	(2.539)	(0.611)	(0.192)	(1.818)	(1.189)	(0.477)
Retired	-5.260	6.191*	0.0204	0.0123***	-0.602	0.559***	0.787
	(3.693)	(3.655)	(0.0155)	(0.00373)	(0.512)	(0.179)	(1.094)

1. Dependent variable is the (0-100) index of happiness equal to 25(sum of the 4 indicators of recent mood). See the text for a description of the indicators.*

2. δ are annual rates of recovery.

3. Standard errors are in parentheses.

4. Events are all dated to the month.

5. η are the effect of an additional year of subjective life expectancy.

Table 3.4: Baseline results with imprecisely-dated events

	Parameter Estimates		
	b_T	b_P	δ
Events dated to the year			
Death of child	-6.003***	0.387	0.960***
	(1.938)	(0.558)	(0.256)
Family Illness	-10.42***	2.889***	0.186***
	(1.262)	(0.770)	(0.0453)
Family member addicted to drugs	-9.948*	9.024*	-0.00412
	(5.634)	(5.397)	(0.00484)
Family member unemployed more than 3 months (in last 5 years)	-3.026	1.060	0.461
	(2.435)	(2.121)	(0.303)
Fire, flood, earthquake, or other disaster	-3.893*	3.147*	-0.000400
	(2.239)	(1.851)	(0.00176)
Fired (in last 5 years)	-4.929	4.639	0.106
	(6.254)	(6.122)	(0.101)
Landlord/realtor discrimination	23.70**	1.656	-0.201
	(10.37)	(2.654)	(0.201)
Moved to worse nbhd (in last 5 years)	-30.60	25.88	0.0668
	(26.50)	(26.08)	(0.0700)
Respondent illness	-6.079***	3.808*	0.0362
	(2.137)	(2.155)	(0.0250)
Respondent unemployed more than 3 months (in last 5 years)	13.67**	-10.08*	0.0635
	(5.889)	(5.541)	(0.0704)
Robbed (in last 5 years)	-0	-0.977	38.50
	(0)	(1.184)	(115.1)
Serious physical assault	-15.04***	3.494	0.0691**
	(3.481)	(2.472)	(0.0327)
Unfair dismissal	-7.115***	1.336*	0.287**
	(1.717)	(0.713)	(0.121)
Unfairly denied loan	-22.16**	21.27***	0.00204
	(8.723)	(8.153)	(0.00412)
Unfairly denied promotion	-4.277	-0.321	2.023
	(9.789)	(0.849)	(5.435)
Unfairly not hired	-9.329*	0.938	0.349
	(5.113)	(1.093)	(0.360)

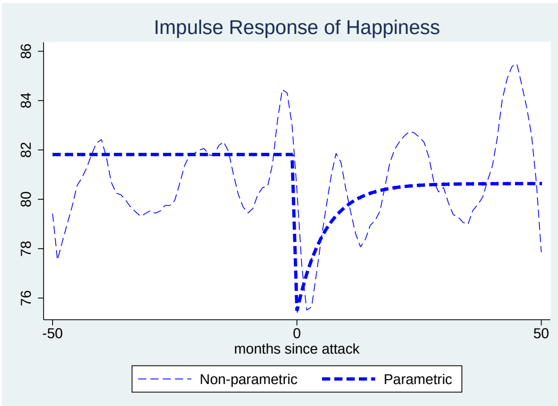
Baseline results with imprecisely-dated events

Unfairly treated by police	-5.42e-06	1.919	7.867
	(1.97e-05)	(1.197)	(6.871)
Combat experience	258.3	-243.4	0.000550
	(272.2)	(272.0)	(0.000652)
Events dated to the wave			
Cataract surgery	0.00156	25.75	-0.00831
	(0.00273)	(18.91)	(0.00559)
Child moved within 10 miles	-0.347**	32.16	0.00393
	(0.136)	(203.1)	(0.0248)
Congestive heart failure	-0.0154	352.2	0.000801
	(0.0279)	(491.1)	(0.00108)
Hip fracture	-10.58**	74.13	0.00625
	(5.066)	(78.30)	(0.00659)
Incontinence	-2.289***	0.860*	0.597*
	(0.515)	(0.488)	(0.317)
New grandchild	-0.134*	1061	4.64e-05
	(0.0762)	(2117)	(9.28e-05)

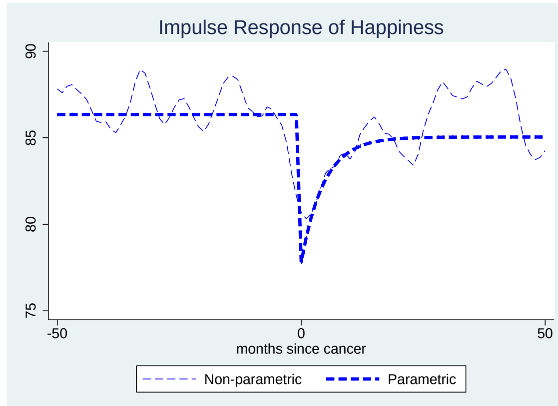
1. Dependent variable is the (0-100) index of happiness equal to 25(sum of the 4 indicators of recent mood). See the text for a description of the indicators.*

2. δ are annual rates of recovery.

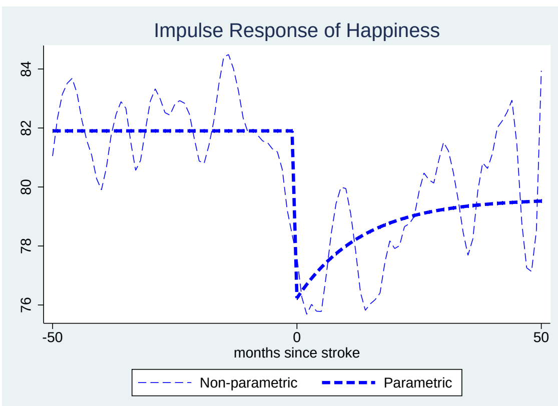
3. Standard errors are in parentheses.



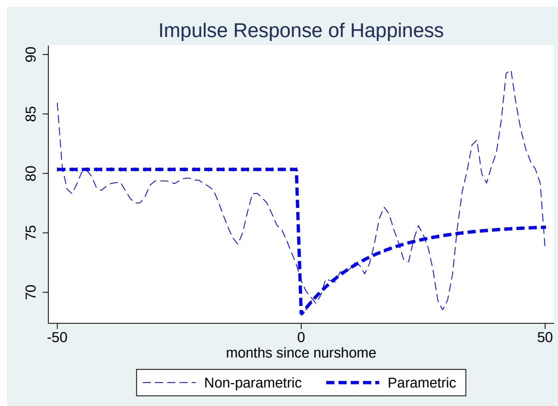
(a) Heart attack



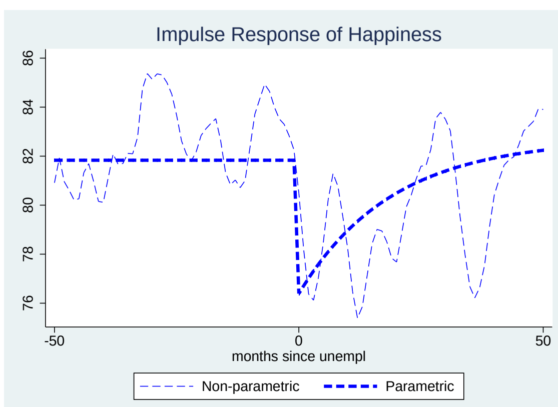
(b) Cancer



(c) Stroke

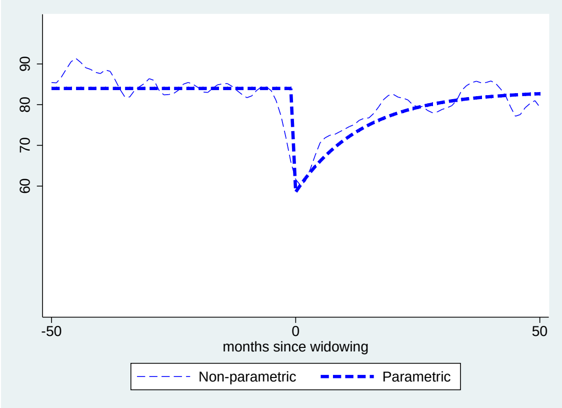


(d) Entered nursing home

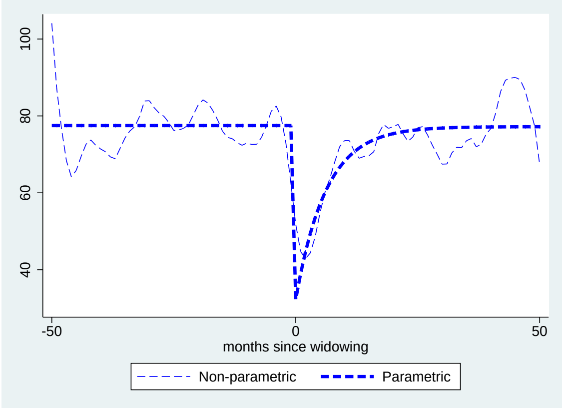


(e) Unemployment

Figure 3.2: Baseline and non-parametric results for precisely-dated events



(a) With life insurance



(b) Without life insurance

Figure 3.3: Widowing with and without life insurance

CHAPTER IV

Taxation, Match Quality and Social Welfare

4.1 Introduction

The relationship between taxation and social welfare is of central importance to public finance and the accurate appraisal of public policy. In a sufficiently simple model, the behavioral employment response to a tax entirely determines deadweight loss (DWL), and the labor suppy elasticity is sufficient for its calculation. However, as argued in Feldstein (1999), behavioral responses alone do not incorporate some important channels by which taxes reduce social welfare. The taxable income elasticity, rather, reflects information concerning tax evasion, shifts to deductible and tax-excluded consumption, and so forth, all of which affect deadweight loss. Recent literature has developed around the proposition that taxable income elasticities can generally serve as a sufficient statistic for the deadweight loss from taxation, assuming (as described in Chetty, 2008) the absence of fiscal and classical externalities as well as costless transfers between agents. In addition, Nunn (2012) suggests that social welfare inference from any observed elasticity is incorrect if quantitatively-significant job amenities are ignored by the econometrician. Nunn estimates a labor search model with exogenously-endowed match-specific amenities, shows that that non-wage amenities are an important component of worker compensation, and demonstrates that their existence induces the over-estimation of DWL from taxation.

The aim of the present paper is to develop a complementary and broader exploration of match amenities and their connection to social welfare. In particular, our focus is on the impact of the endogenous supply of amenities.

First, we take a step back and focus on the role of amenities in the absence of search frictions. We consider an economy in which match quality is fixed, but amenity provision is flexible and subject to convex production costs. Workers also choose hours, which are multiplied by match quality to yield the total surplus from a match. Within this context, adjustment to taxation occurs immediately, and deadweight loss is generated by inefficiently-high amenity consumption and a low supply of work hours. Interestingly, we find that, even in this static case, the introduction of amenities causes taxable income elasticities to overestimate the deadweight loss from taxation.

We then move on to considering an environment with search frictions, which allows for empirical estimation. The introduction of search frictions enables the coexistence of matches of varying quality and amenity productivity, as well as the opportunity for workers to move between these. We retain the assumption of fixed match-specific job quality and convex amenity production costs. As in Nunn (2012), we estimate the model with duration data from the 1979 National Longitudinal Survey of Youth (NLSY79), and compute steady state equilibria for various tax rates. The flexible-amenity model implies a more severe overestimation of the deadweight loss from taxation when amenities are mistakenly ignored. As such, the problem of reliance on taxable income elasticities as a proxy for the welfare effects of taxation is exacerbated. In addition, we examine the dynamic effect of a change in taxes. The amenity production function allows for both instantaneous and drawn-out responses to changes in market conditions, since firms can adjust amenity supply immediately, and workers can slowly move to matches with more efficient amenity production. Given this, we show that taxable income elasticities over-estimate deadweight loss by an even greater amount as they become more long-run in scope.

Three seminal papers make it possible to conduct our structural analysis: Mortensen and Pissarides (1994) (job search), Jovanovic (1979) (job duration as match quality), and Shimer

(2006) (on-the-job search). Mortensen and Pissarides provide the basic framework for the random search model employed in this paper, albeit without on-the-job search and a few other modifications. Working within the Mortensen and Pissarides class of model, Shimer (2006) adds on-the-job search in a partial equilibrium context, with job offers arriving at an exogenous rate constrained to be the same on and off the job.

By embedding match quality variation in an equilibrium model of job turnover, Jovanovic (1979) and related papers like McCall (1990) provide a theoretical basis for an assertion we adhere to, which is that tenure is informative about job match quality. Indeed, a large, mostly empirical literature has developed that takes tenure as a proxy for match quality and examines, for example, job mismatch over the business cycle (Bowlus, 1995, Kahn, 2008) and the effects of changes in unemployment insurance law (Ours and Vodopivec, 2008). Some authors have pursued related questions with identification or calibration strategies making use of duration data. Nagypal (2007), for instance, distinguishes accumulation of human capital from learning about match quality using, in part, matched firm-worker data including tenure information. Becker (2009) focuses on job amenities and finds that they are quantitatively substantial.

Sullivan and To (2011) assess the relative importance of wage and non-wage job utility. They estimate the offer distributions of wages and non-wage utility in a search context, and use the fraction of job switches associated with wage declines to identify the distribution of non-wage utility. Although they are not concerned with estimating the distribution of match quality, like Becker, Sullivan and To find non-wage utility to be substantial.

Though the paper is not explicitly about compensating differentials, separate estimation of productivity and amenities, along with the particular wage bargain assumed, implicitly involve tradeoffs between wages and amenities. As is intuitively the case, workers in jobs with low amenities will (holding match quality constant) receive higher wages. Our framework makes use of wages to help separate productivity and amenities, which puts it in debt of papers like Rosen (1974) and Friedman and Kuznets (1954). Unlike most compensating differentials papers, however, the present paper does not assume equality of utility across options and the focus is on parameters of the aggregate match quality distributions rather than cross-sectional tradeoffs.

On-the-job search models typically predict that all job changes will be associated with increases in observed wages. However, this is decidedly not the case in the data[1], and various solutions have been proposed. Of course, some or all of wage declines associated with jot-to-job transitions can be interpreted as measurement error. In the NLSY79, for instance, that some reported wage declines are spurious is explicitly noted by the survey administrators.[2] Wolpin (1987) and related literature explicitly incorporate measurement error into their models. It is difficult to explain the extent of job-to-job transitions that involve wage declines with measurement error alone, considering, for instance, that a full third of job-to-job wage changes in the NLSY79 are reductions. Another approach is to construct some aspect of the model that leads workers to (occasionally) optimally switch to lower-wage jobs. As an example, Postel-Vinay and Robin (2002) and Cahuc et al. (2006) use a bargaining arrangement that allows workers to choose lower-wage but higher-productivity firms that will in the future be able to better match outside offers. The present paper, though not principally motivated by the concern of wage reductions in job-to-job transitions, provides a different answer to this question. When unpriced amenities are accounted for, optimal wage choice is such that workers will frequently choose lower-wage jobs that are nonetheless preferred to previous jobs due to their superior amenities.

This paper proceeds as follows. Section 4.2 presents our analysis pertaining to the static economy, and Section 4.3 develops the model with search frictions. Section 4.4 presents details on the data we use for estimation, Section 4.5 presents results, and Section 4.6 concludes.

[1] See Sullivan and To (2011) for data from the NLSY97. Postel-Vinay and Robin (2002) find a similar result in French data, with a third to a half of workers reporting wage decreases as they change jobs.

[2] From NLSY79 documentation at `http://www.nlsinfo.org/nlsy79/docs/79html/79text/wages.htm`: "Note that: the calculation procedure, which factors in each respondent's usual wage, time unit of pay, and usual hours worked per day/per week produces, at times, extremely low and extremely high pay rate values; no editing of values reported by a respondent occurs even if the value is extreme, such as $25,000 per hour..."

4.2 Static economy with amenities

As shown in Feldstein (1999), under conditions outlined in that paper, the deadweight loss from taxation can be stated as a function of the elasticity of taxable income with respect to the net of tax rate ε. In particular, DWL is shown to be proportionally increasing in ε. Thus, knowledge of ε along with the proportionality parameter and any applicable intercept parameter is sufficient to infer DWL. In this section, we show that even in a static context (i.e., in the absence of search) when variables that matter for utility are not accounted for, DWL can be overestimated when relying on the aforementioned methodology.

Consider an economy inhabited by a representative worker and a profit-maximizing firm. Individuals' utility is an increasing function of after-tax consumption, the fraction of time spent on leisure $\ell \in (0,1)$, and on-the-job amenity q. Firms maximize

$$\pi (1-\ell) - W (1+\tau) - \phi (q),$$

where $\pi (1-\ell)$ is the worker's monetary productivity, W is the after-tax real wage, and $\phi (q)$ is the cost of producing the amenity level q. Equilibrium $\{W, \ell, q\}$ can be interpreted as the result of workers making firms take-it-or leave offers, or firms posting offers that workers choose, therefore inducing a zero-profit condition.

We focus on a worker's problem (the solution of which reveals equilibrium $\{W, q, \ell\}$) being the maximization of utility consistent with providing a firm at least zero profit:

$$\max_{W,q,\ell} u (W+T, \ell, q) \text{ such that } \pi (1-\ell) \geq W (1+\tau) + \phi (q).$$

All tax revenues are returned by the government to workers in transfers denoted T; T is not perceived by workers as being related to tax payments. Let λ be the Lagrange multiplier associated with the worker's problem. DWL is $-du/d\tau$, and $\varepsilon = d \ln ((1+\tau) W) / d \ln (1-\tau)$. Note that if the presence of the amenity is incorrectly omitted, then the worker's problem is assumed to be

$$\max_{W,\ell} u\left(W+T,\ell\right) \text{ such that } \pi\left(1-\ell\right) \geq W\left(1+\tau\right).$$

By making use of a simple utility and amenity production cost function, Proposition 1, which follows below, serves to illustrate that overestimation of DWL by use of taxable income elasticities is not necessarily an artifact of the search process itself. Intuitively, this occurs because omitting from consideration an untaxed choice variable that enters the utility function and over which individuals can adjust is equivalent to omitting a margin of adjustment that allows mitigation of the utility impact of taxation.

Proposition 1. *Suppose* $u = \alpha\log\left(W+T\right)+\beta\log\ell+\gamma\log q$ *and* $\phi = (1/A)\exp\left(q\right)$, *where* A *is a scaling constant. Then,*

1. *DWL can be stated as*

$$a\left(W,\ell,q,\cdot\right)+c\left(W,\ell,q,\cdot\right)dT/d\tau+b\left(W,\ell,q,\cdot\right)\left(1-\tau\right)^{-1}\varepsilon Y, \tag{4.1}$$

   where Y *is taxable income.*

2. *If all parameters and variables that enter equation* (4.1) *are observed, but utility is incorrectly assumed to be given by the sum of* $\alpha\log\left(W+T\right)$ *and* $\beta\log\left(\ell\right)$, *then, for* $T>0$ *and* $dT/d\tau \geq 0$, *DWL is overestimated.*

Proof. See appendix.

Corollary. *If all parameters and variables that enter equation* (4.1) *are observed and the amenity is not a choice variable, then even if utility is incorrectly assumed to be given by the sum of* $\alpha\log\left(W+T\right)$ *and* $\beta\log\left(\ell\right)$, *equation* (4.1) *yields the correct value for DWL.*

For simplicity, utility in search models is often assumed to be linear; this is something we assume ourselves later in the paper. Proposition 2 shows that the implications of Proposition 1 can also carry over to situations in which utility is linear.

Proposition 2. *Suppose $u = (W+T) + \ell + q$, $\phi = (1/A)\exp(q)$, and ℓ is inelastic at some level $\bar{\ell}$. Proposition 1 and its corollary hold, mutatis mutandis.*

Proposition 3. *The implications from Propositions 1 and 2 go through if instead the linear cost function $\phi = q$ is assumed.*

As developed above, our framework differs from that of Feldstein (1999) primarily by allowing for the possibility of a convex amenity cost function instead of an implicitly constant returns one. This, on its own, leads to changes in the price of leisure relative to the amenity after a tax is introduced, which makes the relationship between DWL and taxable income more complicated. Note that given convex costs, in the static case the amenity's relative price is increasing after a tax, whereas in the search setting the effective price may actually decline as workers find jobs with more readily available amenities.[3]

4.3 Model

The model builds on those developed in Mortensen and Pissarides (1994) and Shimer (2006). We account for on-the-job search, match quality heterogeneity, endogenous job destruction stemming from both job-to-job transitions and changes in the idiosyncratic productivity of a match, and job switching costs. Search frictions are such that workers and firms meet each other only occasionally. Matches produce a flow surplus that depends on idiosyncratic (time-varying) productivity x, a time-invariant match-specific monetary productivity π, and an amenity q that is produced endogenously by firms. The production function for q is $q(k) = \log(A \cdot k)$, where k is the capital input and A is the match-specific productivity of amenity supply. A is distributed lognormally with $\sigma = \sigma_A$ and $\mu = 0$. Because matching

[3] As shown in the appendix, if all parameters and variables that enter equation (4.1) are observed, but utility is incorrectly assumed to be given by the sum of $\alpha \log(w+T)$ and $\beta \log(\ell)$, then, for $T > 0$ and $dT/d\tau \geq 0$, DWL is also overestimated under the assumption of a linear cost function $\phi = q$. Also shown in the appendix is that, assuming that all parameters and variables that enter equation are (4.1) observed, but the cost of the amenity is incorrectly assumed to be q when in fact it is $(1/A)\exp(q)$, then use of equation (4.1) leads DWL to be underestimated; if the reverse occurs, then DWL is overestimated. Results applicable to the linear case are case specific as they can involve corner solutions; relevant details can be found in the appendix.

opportunities are scarce, surplus is generated by successful matches, and wages are set by bargaining over this surplus. A worker's values of being unemployed or employed, as well as firms' value of employment, are represented as Bellman equations. Since the resulting functions are both monotonic and discounted, these equations are contraction mappings.[4]

The model is not solvable analytically, so we solve it via value function iteration on a discrete grid and simulate it in discrete-time, calibrated to a monthly frequency[5]. The timing of the model is as follows. A period begins with any particular worker being either unemployed or employed. If employed, a worker-firm match is characterized by both a constant match productivity π and a constant amenity production parameter A[6]. These are drawn simultaneously when a prospective match forms, and do not change over the course of a match; additional details about draws of π, x, and A and consequent q production will be provided throughout the rest of this section. Then, a time-varying idiosyncratic productivity shock x is drawn; this occurs every period. Employed individuals receive a wage and unemployed individuals receive exogenous unemployment flow benefits. Next, an exogenous separation shock s may occur. If it does, unemployment results and the worker does not receive an employment offer until at least the subsequent period. If a separation does not occur, or if the worker was already unemployed, then a job-finding shock may occur. Firms and workers meet each other with a probability that depends only on the worker's employment status: α_0 for an unemployed worker and α_1 for an employed worker. The idiosyncratic productivity draw x occurs simultaneously with the match shock α, allowing workers to choose between unemployment and employment (the former being chosen if the productivity draw is such that the value of unemployment is higher), switching to a new job (occurring, for employed workers, if the surplus of the new match exceeds that of the present match, in which case switching costs are paid immediately), and remaining in the

[4]Blackwell (1965) and Sargent (1987).

[5]Except under very particular assumptions (e.g., Shimer, 2003) that would make identification difficult or impossible, on-the-job search and the switching cost make it impossible to find a closed-form solution to the model.

[6]Though this paper is concerned with both the distribution of match quality draws and the actual distribution of accepted draws, "match quality distribution" refers to the former, unless otherwise specified.

old job (occurring if an individual is already employed and the continuation value of the match exceeds that of all other alternatives). All job-finding shocks are characterized by a match quality draw and a new idiosyncratic productivity draw on which wages (fully flexible and instantly renegotiated) are based. Following this, the economy moves into the new state and the period ends. Note that the switching cost is considered to be a sunk cost for wage-setting purposes. This is consistent with the usual practice in most of the hiring cost literature, which can be thought of as analogous to the switching cost considered here.

The model timing is depicted by the following graph:

$$\boxed{\text{wage or } b \text{ earned}} \Longrightarrow \boxed{s \text{ shock}} \Longrightarrow \boxed{\alpha \text{ and } x \text{ shock}} \Longrightarrow \boxed{\text{worker takes action}} \Longrightarrow \boxed{\text{period ends}}$$

Idiosyncratic productivity draws are match-specific and time-varying, so it is possible for workers to switch to lower match quality jobs with sufficiently high productivity draws. x is drawn from a lognormal distribution with a persistence ρ_x, following the process $\log(x') = \rho_x \log(x) + \epsilon_x$, where $\epsilon_x \sim N(0, \sigma_x^2)$.[7]

Firms and workers encounter one another with probabilities that depend only on employment status and are constant over time. Match quality m consists of two components: a monetary productivity term π that accrues to the firm and a non-monetary benefits term q that is entirely consumed by the worker. The amenity q is produced endogenously by firms as described previously. π draws are distributed according to a normal cdf Π with mean zero and standard deviation σ_π. m is equal to the sum of π and q and obeys a normal cdf M with mean zero and standard deviation $\sigma_m = \sqrt{\sigma_\pi^2 + \sigma_q^2}$. All of these quantities are drawn simultaneously when a firm meets a worker. π and A (the amenity production parameter) remain constant for the duration of a match. A one-time switching cost c is incurred by employed workers who accept new job offers.

[7]The x grid and discrete transition matrix are formed according to the Tauchen (1986) procedure.

Denote a worker's value of employment by W, and her value of unemployment by U. Employed individuals receive as compensation the amenity level q, and also the pre-tax wage w (to which a linear tax rate τ is applied). All economic agents discount the future at rate β. With probability s, the match is destroyed and the worker is forced into unemployment. With probability α_1, the worker encounters an outside job opportunity. Thus, with probability $(1-s)\,\alpha_1$, an employed individual must decide whether to switch jobs (taking account of instantaneous job-changing costs c, which are time invariant and exogenous), remain in her current job, or transition into unemployment (voluntarily, if the value of doing so is greater than that of the alternatives). Similarly, with the product of the probabilities $(1-s)$ and $(1-\alpha_1)$, the worker must decide between remaining in her current job, or transitioning into unemployment. Therefore,

$$W(\pi,q,x) = (1-\tau)w(\pi,q,x) + q + \beta sU + \underbrace{\beta(1-s)(1-\alpha_1)Prob((U > W(\pi,q,x'))|x)\cdot U}_{\text{bad } x \text{ shock: separate}}$$
$$+ \underbrace{\beta(1-s)(1-\alpha_1)\mathbb{E}[\mathbb{1}(W(\pi,q,x') \geq U)\cdot W(\pi,q,x')]}_{\text{no job offer: stay in job}}$$
$$+ \underbrace{\beta(1-s)\alpha_1\mathbb{E}[max\{W(\pi',q',x'_{nj}) - c, W(\pi,q,x'), U\}]}_{\text{continuation value conditional on new job offer}},$$

where an apostrophe denotes the value of a variable in the following period, and therefore x'_{nj} is the idiosyncratic shock associated with a new job offer. Following similar considerations,

$$U = b + \beta(1-\alpha_0)U + \beta\alpha_0\mathbb{E}[W(\pi',q',x'),U],$$

where b is the constant net unemployment flow benefit and α_0 is the probability an unemployed individual meets a firm. As noted above, not all meetings between firms and workers lead to a match being formed. This allows, for instance, a distinction to be made between a worker's probability of receiving a job offer and the probability of a transition from unemployment to employment. In the absence of this distinction, endogeneity in the acceptance of offers would potentially bias estimation of the match quality offer distribution, as discussed previously.

Since job separations due to bad idiosyncratic draws are bilateral in the sense that an employer sees it fit to end a match under the same circumstances as the worker, the firm's value of a job, J, can largely be written in terms of worker value functions.

$$\begin{aligned} J(\pi, q, x) =& x + \pi - w(\pi, q, x) - \phi(q) + \beta(1-s)(1-\alpha_1)\mathbb{E}[\mathbb{1}(W(\pi, q, x') \geq U) \cdot J(\pi, q, x')] \\ &+ \beta(1-s)\alpha_1\mathbb{E}[\mathbb{1}((W(\pi', q', x'_{nj}) - c < W(\pi, q, x')) \cap (W(\pi, q, x') \geq U)) \cdot J(\pi, q, x')], \end{aligned}$$

where ϕ is the cost function associated with amenity supply. Note that although we do not model vacancies explicitly, following related literature, the firm's value of a job reflects free entry into vacancy creation, meaning that the value of a vacancy is zero.

The surplus from a match is given by

$$S(\pi, q, x) = W(\pi, q, x) - U + J(\pi, q, x).$$

Workers and firms do not care about the particular composition of $m = \pi + q$; therefore, W, J, and hence the surplus S can be though of as simply taking m rather than $\{\pi, q\}$ as an argument. Nash bargaining over the surplus implies that

$$\gamma S(m, x) = W(m, x) - U(m, x) \text{ and } (1-\gamma) S(m, x) = J(m, x),$$

where γ is the bargaining power of workers. The observed wage is a function of q. Throughout, we solve for the symmetric Nash bargaining equilibrium $\gamma = 0.5$.

The model just described is similar to Nunn (2012), with the following important exception: amenities are now generated according to a production function, the efficiency of which is heterogenous across matches. The log specification, $q(k) = \log(A \cdot k)$, is chosen for computational simplicity; it is appropriately concave in capital input, and as will be shown later, induces an optimal amenity level that is concave in A.

Amenity q is optimally chosen by firms to maximize the total surplus associated with a match. Note that q^* solves a subproblem separable from that solved by the Nash bargain. First, q is set to make a worker indifferent between having another dollar of output devoted to wage increase, and another dollar spent on the amenity, thereby maximizing both the surplus in any given period and the total surplus of a match.[8] Then, the worker and firm split the total match surplus according to a Nash bargain, conditional on the particular q^*. Another way to state this is as follows: when the amenity q is optimally chosen, with the marginal cost of amenity production equal to the (reciprocal of the) marginal benefit of the wage, an increase in the surplus fraction accruing to workers will simply increase the wage paid.

As before, assume that the amenity cost function is $\phi(q) = \frac{1}{A}\exp(q)$. Since q and the wage are perfect substitutes in consumption, and $\phi''(q) > 0$, there is a constant, match-specific optimal level of the amenity, $q^*(A)$, chosen by firms (assuming for the moment that there are no taxes). The condition defining this level is

$$\phi'(q) = \frac{\exp(q)}{A} = \frac{\partial u/\partial w}{\partial u/\partial q} = 1$$

$$\rightarrow q^* = \log(A).$$

Note that workers receive distinct draws of π and A. The latter draw implies a particular optimally-chosen q, and the distributions of π and q conform to the distribution of m assumed by the model. Though workers and firms are indifferent between drawing $\{\pi_{low}, q_{high}\}$ or $\{\pi_{high}, q_{low}\}$, where $\pi_{low} + q_{high} = \pi_{high} + q_{low}$, the two draws do generate distinct observed

[8]Without changes in tax rates, the amenity is optimally held constant during a match, with idiosyncratic productivity variation reflected only in wage variation. This is due to the assumption of linear utility, which substantially simplifies solution of the model by allowing q to be computed separately from the wage.

wages. The correlation between tenure and observed wage allows for separate identification of the π and q distributions.

A crucial aspect of this analysis is the effect of wage taxation on amenity supply and associated social welfare. Assume that w is the pre-tax wage. When a linear wage tax τ is implemented, the first order condition and optimal amenity level are given below. The optimal q is rising in the tax rate, as workers and firms shift compensation into the tax-preferred vehicle.

$$\phi'(q) = \frac{\exp(q)}{A} = \frac{\partial u/\partial w}{\partial u/\partial q} = \frac{1}{1-\tau}$$

$$\rightarrow q^* = \log(A) - \log(1-\tau)$$

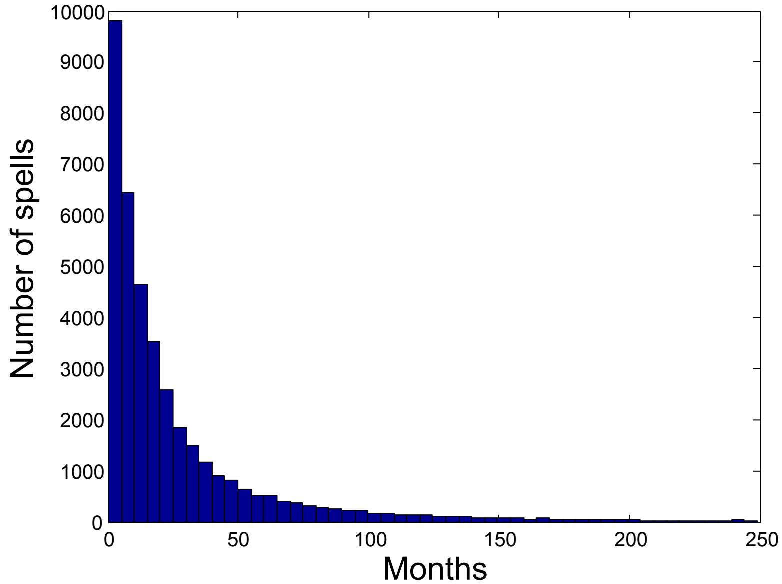
4.4 Data

Data is taken from the 1979 National Longitudinal Survey of Youth. The NLSY79 is a nationally-representative panel of more than ten thousand individuals at inception in 1979, with periodic successive surveys through the present. The detailed employment, demographic, and job spell data is necessary for this study, and the panel nature of the data permits comparison of this paper's results with those obtained from a more conventional wage regression approach, which requires person and match fixed effects.

We drop the military and supplemental subsamples. We also drop currently-enrolled high-school and college students from the sample. This leaves us with about 30,000 primary job spells. Table 4.1 gives summary statistics for relevant unweighted NLSY79 variables. Figure 4.1 shows the unweighted empirical tenure distribution.

Particularly notable is the fact that a surprising fraction of job-to-job wage changes are negative; fully one third of such wage changes are decreases, though measurement error is almost certainly exaggerating this figure. Due to concerns about measurement error, we use only the middle 95 percentiles of the wage data, making the same adjustment in the model's simulations so that only the trimmed data is generated. This mitigates the effect of

Figure 4.1: Unweighted empirical tenure distribution



error-ridden outliers. Wages are adjusted by the CPI-U to 2010 dollars.

Data are weighted to yield a nationally-representative sample before use in the model. We use the NLS79's tenure variables corresponding to the primary jobs of respondents as well as various demographic variables that allow construction of a conditional tenure distribution. The tenure variables are constructed by the NLSY using worker-reported job start and stop dates, and are connected across waves of the survey by means of employer identification numbers. Demographic and other variables are dated to the end of each completed job spell. In addition, We set the simulated panel length equal to the average time a worker is present in the sample. This ensures that any truncation in the data (due to the requirement that spells terminate before the sample ends) is matched by truncation of the simulated spells.

We do not use the weighted empirical moments directly. Certain observable variables - age, education, etc. - explain some variation in job duration. Ex ante variation in worker characteristics is not part of the baseline model, so we prefer to adjust the data to more closely approximate the model's assumption of ex ante identical agents. To construct the empirical tenure variables, we first run the regression:

$$T_{it} = X_{it}\beta + \epsilon_{it}$$

where T_{it} is the tenure for a particular person's completed job spell beginning in year t, X_{it} is demographic information associated with a person in year t including age, education, sex, and race. We experimented with different specifications of industry and occupation dummies, but these made little difference to the results after the demographic variables were included. The time-varying variables are taken at the end of a job spell. Since the structural model is one of homogeneous agents who differ only ex post, we then construct tenure spells that are purged of observable variation due to age and other variables. The new tenure variable is given by $\hat{T_{it}} = \hat{\epsilon_{it}} + \overline{X}\hat{\beta}$, where $\overline{X}$ is the vector of population means corresponding to the demographic variables. Wages are adjusted in precisely the same way.

4.5 Results

The moments and parameters identified are similar to those used in Nunn (2012). Simulated and empirical moments are presented in Table 4.2. The model does not closely match two particular moments: the unemployment rate and the job-to-job transition probability. At 0.069, the observed NLSY job-to-job transition probability is quite high: a multiple of the monthly rate generally thought to obtain in the US labor market (see, for example, the 0.028 figure calculated using SIPP data by Nagypal, 2008). Our model predicts a substantially smaller quantity than the NLSY79 figure. This is essentially due to the large inframarginal rents implied by substantial match quality variation: many workers quickly move to matches with high match quality and then receive comparatively few offers or productivity shocks that induce them to quit for either another job or unemployment, respectively. It is perhaps the case that a more flexible framework for match quality determination - amenity depreciation, for example, or separate offer distributions for unemployed and employed workers -

would more effectively match both the observed tenure moments and job-to-job transition data.

4.5.1 Endogenous amenity supply

In this section, we compare deadweight loss results from Nunn (2012) with the current analysis. The previous paper required amenities to be constant during a match and exogenously endowed at the moment a worker meets a firm. Amenities were therefore not permitted to adjust in response to tax changes. The current paper relaxes this assumption and allows for additional important margins of adjustment that may be relevant to the social welfare and taxable income responses to tax changes.

Social welfare in a period is calculated as the sum of total flow surplus minus the cost of amenities and switching costs

$$\sum_{i}^{I}(\mathbb{1}(employed_{it}=1)\cdot(x_{it}+\pi_{it}+q_{it}-\phi(q_{it}))) - \sum_{i}^{I}(\mathbb{1}(jobswitch_{it}=1))\cdot c,$$

where $\mathbb{1}(jobswitch_{it}$ is an indicator for whether a particular worker moves from one match to another in a given period.

Taxable income is the sum of pre-tax wages received in a particular period

$$\sum_{i}^{I}(\mathbb{1}(employed_{it}=1)\cdot(w_{it})).$$

Both social welfare and taxable income are normalized to 100 in the absence of taxation. Note that job switching costs are not considered to be tax-deductible. A substantial portion of switching costs are non-monetary, and many workers in the data do not itemize their deductions in any event.

The tables in the appendix give results under the two sets of assumptions. What is notable is that the paths of social welfare and taxable income, as the tax rate increases, are even more different in the case with endogenous amenity supply (as compared to the case with amenity endowment). For instance, between tax rates of 20% and 30%, the taxable income and social welfare elasticities are 0.20 and 0.06, respectively, in the case of exogenous amenity endowment. This is in contrast to the 0.39 and 0.08 elasticities for taxable income and social welfare, respectively, in the endogenous amenity supply case. In this latter case, taxable income falls very quickly as the tax rate rises, reflecting both the immediate ability of firms to reallocate compensation into amenities and out of wages, and also the long-run ability of workers to choose firms with higher amenity productivity A but lower monetary productivity π. Results are given in Table 4.3 and Figure 4.2. In the original case, shown in Table 4.4 and Figure 4.3, only the long-run movement to higher-amenity jobs was allowed to occur.

Note that social welfare and taxable income elasticities are generally higher in the endogenous amenity environment than in the exogenous environment. With endogenous amenities, there is now an additional margin on which firms can adjust. More specifically, consider the subset of workers in the exogenous amenity endowment case who are in good matches that happen to be characterized by high productivity and low amenity. A sufficiently small tax increase will not destroy these matches or induce these workers to quit for matches with higher amenity levels. However, in the endogenous amenity economy, even a small tax increase will immediately cause firms to offer higher amenity levels to everyone, including the aforementioned workers. This allows for a sharper response to tax changes.

To the extent that firms can adjust their supply of amenities valued by workers, then, the issue of deadweight loss over-estimation is more serious than in Nunn (2012). In Section 4.5.2, we will disaggregate the short- and long-run mechanisms of labor adjustment to tax changes.

4.5.2 Dynamic analysis

Section 4.5.1 develops a model that allows both workers and firms to flexibly respond to changes in economic conditions (namely, tax rates). Firms immediately adjust the mix of wage and amenity compensation, while workers have the additional ability to search for new matches with preferable amenity productivity. A tax reduction, for instance, will instantly cause a reduction in amenity supply. It will also induce workers to search for jobs with higher overall match quality but lower amenity productivity now that compensation in the form of wages is less penalized.

In a model with both immediate and frictional adjustment, it is now possible to examine the dynamics of the tax response. The differences between short- and long-run taxable income elasticities have been the subject of much research (Saez et al., 2009). Saez and coauthors discuss the importance of having the long-run elasticity for the purpose of estimating the total deadweight loss from an income tax, but recognize the substantial econometric difficulties implicit in calculating long-run consequences. Short-run elasticities necessarily do not reflect the costly labor market and tax adjustments that workers can make only over a substantial period of time. On the other hand, short-run elasticities overstate the true response to tax law changes when workers can shift income across time. [9] Chetty et al. (2011) develop a model in which labor supply elasticities are measured differently depending on whether the identification strategy incorporates worker search/adjustment costs. They show that macro elasticities do not get at these costs, while micro elasticities do, causing the macro estimates to be substantially larger.

The model presented in this paper allows for an examination of the tax response over time. We conduct an experiment in which the tax rate is initially assumed to be 24.8%, then unexpectedly increased by 10 percentage points[10]. As one would expect, taxable income falls sharply just after the unexpected tax increase, reflecting the shift of compensation to (costly) match amenities. Immediately after the tax increase, taxable falls by a bit more than one percentage point. Taxable income then falls slowly as workers gradually find new jobs with more efficient amenity supply, and associated lower wages.

[9] "...the tax reform of 1993 seems to have generated a temporary decline in top 1% incomes in 1993 and a temporary spike in 1992 as tax filers tried to pull reported taxable income from 1993 into 1992 to take advantage of the lower 1992 tax rate. As a result, the elasticity estimated using only the years 1992 and 1993 is large." (p.28, Saez et al. (2009))

[10] 24.8% is the 2010 ratio of tax revenues (at all levels of government) to gross domestic product as estimated in the OECD Tax Revenue Statistics.

Social welfare in the model also drops discontinuously in the aftermath of a new tax, as indicated in Figure 4.4 and Table 4.5. These graphs depict social welfare and taxable income just prior to and in the months after a 10 percentage point tax increase, with the pre-tax change baseline normalized to 100.

It is convenient to think about this counterfactual in terms of a social planner's problem. Intuitively, the social planner would choose a lower level of amenity and higher wage than workers and firms find it privately optimal to select in the aftermath of a tax increase, and this is reflected in the initial social welfare drop. This is simply the usual logic of a social planner preferring a muted behavioral response to taxation. Perhaps counter-intuitively, however, social welfare actually rises slightly, after the initial sharp decline, as workers make their way to matches with more efficient amenity production. While distorting in its own right, the migration of workers to these new jobs appears to mitigate the distortion associated with the initial sharp increase in amenity production. For this reason, taxable income elasticities actually over-estimate deadweight loss more as they become more long-run in scope. Nunn (2012) explains that the existence of unobserved amenities, in a search context, are already sufficient to bias upwards estimates of deadweight loss from taxable income elasticities. The current analysis indicates that this problem is compounded by a focus on long-run elasticities, simply because these larger long-run taxable income elasticities indicate an excessively large deadweight loss. However, work that does take amenities into account would ideally use long-run elasticities, since short-run responses fail to account for endogenous changes in amenity compensation.

4.6 Conclusion

The costly, endogenous supply of job amenities by firms has a number of interesting implications for labor markets and public policy. In particular, results from the present paper reinforce a conclusion from Nunn (2012): the public finance literature concerning deadweight loss (DWL) estimation should take amenities into account. Work by Martin

Feldstein and others has made the case for utilizing data on taxable income elasticities, rather than labor supply elasticities, for the purpose of calculating social welfare consequences of taxation. We show that it is necessary to estimate the heterogeneity and parameters of amenity supply in order to correctly infer deadweight loss from taxation. Deadweight loss is generally overestimated in work that omits explicit consideration of amenities, and this overestimation is proportional to the quantitative significance of heterogeneity in amenities across job matches. The endogenous supply of amenities, rather than exogenous endowment, is shown to exacerbate this overestimation. Examination of a friction-less economy reveals that this result is not necessarily an artifact of search. In addition, we argue that long-run taxable income elasticities, while often considered by the public finance literature to be preferable, can be even more misleading than short-run elasticities after taking into account the possibility of gradual migration of workers into more amenity-focused job matches in response to tax increases.

4.7 Appendix

4.7.1 Curvature in Utility

Convex amenity cost function: $\phi = (1/A)\exp(q)$. The Lagrangian is

$$\mathcal{L} = \alpha \log(W+T) + \beta \log \ell + \gamma \log q + \lambda\left(\pi(1-\ell) - W(1+\tau) - \phi(q)\right),$$

and the first-order conditions (FOCs) are

$$\frac{\alpha}{W+T} - \lambda(1+\tau) = 0 \implies \frac{\alpha}{W+T} = \lambda(1+\tau) \implies \frac{\alpha}{\underbrace{(W+T)(1+\tau)}_{g}} = \lambda \tag{4.2}$$

$$\frac{\beta}{\ell} - \lambda\pi = 0 \implies \frac{\beta}{\ell\pi} = \lambda \tag{4.3}$$

$$\frac{\gamma}{q} - \lambda\phi' \implies \frac{\gamma}{q\phi} = \lambda, \tag{4.4}$$

where the third FOC follows by the assumption that $\phi = (1/A)\exp(q)$. Combining the first and second FOCs, and the the first and third, yields

$$\frac{\alpha}{g} = \frac{\beta}{\ell\pi} \implies \underbrace{\frac{\alpha}{\beta}}_{x_1}\ell = \frac{g}{\pi} \tag{4.5}$$

and

$$\frac{\alpha}{g} = \frac{\gamma}{q\phi} \implies \underbrace{\frac{\alpha}{\gamma}}_{x_2}q\phi = g, \tag{4.6}$$

jointly implying that

$$x_1\ell\pi = x_2 q\phi = (W+T)(1+\tau) = g. \tag{4.7}$$

Total differentiation of equation (4.5) implies

$$
\begin{aligned}
x_1 \pi d\ell &= (1+\tau)\, dW + (1+\tau)\, dT + (W+T)\, d\tau \\
\implies \frac{d\ell}{d\tau} &= \frac{W+T}{x_1 \pi} + \frac{1+\tau}{x_1 \pi} \frac{dW}{d\tau} + \frac{1+\tau}{x_1 \pi} \frac{dT}{d\tau},
\end{aligned}
\tag{4.8}
$$

while total differentiation of equation (4.6) implies

$$
\begin{aligned}
x_2 \phi dq + q x_2 \phi' dq &= (W+T)\, d\tau + (1+\tau)\, dW + (1+\tau)\, dT \\
\implies \frac{dq}{d\tau} &= \frac{W+T}{(1+q)\, x_2 \phi} + \frac{1+\tau}{(1+q)\, x_2 \phi} \frac{dW}{d\tau} + \frac{1+\tau}{(1+q)\, x_2 \phi} \frac{dT}{d\tau},
\end{aligned}
\tag{4.9}
$$

where the second line follows from $\phi = \frac{1}{A} \exp(q)$. Futhermore, note that

$$
\varepsilon = \frac{d \log((1+\tau)\, W)}{d \log(1-\tau)} = -\frac{dW}{d\tau} \left(\frac{1-\tau}{W} \right) - \left(\frac{1-\tau}{1+\tau} \right) \tag{4.10}
$$

$$
\implies \frac{dW}{d\tau} = \frac{-W}{1+\tau} - (1-\tau)^{-1} W
$$

$$
\implies \frac{dW}{d\tau} = \frac{-W}{1+\tau} - \frac{(1-\tau)^{-1} \varepsilon Y}{1+\tau} \tag{4.11}
$$

where $Y = W\,(1+\tau)$ is taxable income.

Substitute equation (4.11) in (4.8), to obtain

$$
\begin{aligned}
\frac{d\ell}{d\tau} &= \frac{W+T}{x_1 \pi} + \frac{1+\tau}{x_1 \pi} \left(\frac{-W}{1+\tau} - \frac{(1-\tau)^{-1} \varepsilon Y}{1+\tau} \right) + \frac{1+\tau}{x_1 \pi} \frac{dT}{d\tau} \\
\implies \frac{d\ell}{d\tau} &= \frac{T}{x_1 \pi} - \frac{(1-\tau)^{-1} \varepsilon Y}{x_1 \pi} + \frac{1+\tau}{x_1 \pi} \frac{dT}{d\tau},
\end{aligned}
\tag{4.12}
$$

and in (4.9) to yield

$$
\begin{aligned}
\frac{dq}{d\tau} &= \frac{W+T}{(1+q)\, x_2 \phi} + \frac{1+\tau}{(1+q)\, x_2 \phi} \left(\frac{-W}{1+\tau} - \frac{(1-\tau)^{-1} \varepsilon Y}{1+\tau} \right) + \frac{1+\tau}{(1+q)\, x_2 \phi} \frac{dT}{d\tau} \\
\implies \frac{dq}{d\tau} &= \frac{T}{(1+q)\, x_2 \phi} - \frac{(1-\tau)^{-1} \varepsilon Y}{(1+q)\, x_2 \phi} + \frac{1+\tau}{(1+q)\, x_2 \phi} \frac{dT}{d\tau}.
\end{aligned}
\tag{4.13}
$$

Furthermore, differentiate the utility function and divide by $d\tau$ to obtain

$$\text{DWL} = -\left(\frac{\alpha}{W+T}\right)\frac{dW}{d\tau} - \left(\frac{\alpha}{W+T}\right)\frac{dT}{d\tau} - \frac{\beta}{\ell}\frac{d\ell}{d\tau} - \frac{\gamma}{q}\frac{dq}{d\tau}. \tag{4.14}$$

Substitute in for $d\ell/d\tau$ using equation (4.12) to obtain

$$\begin{aligned}
\text{DWL} &= -\left(\frac{\alpha}{W+T}\right)\frac{dW}{d\tau} - \left(\frac{\alpha}{W+T}\right)\frac{dT}{d\tau} \\
&\quad -\frac{\beta}{\ell}\frac{T}{x_1\pi} + \frac{\beta}{\ell}\frac{(1-\tau)^{-1}\varepsilon Y}{x_1\pi} - \frac{\beta}{\ell}\frac{1+\tau}{x_1\pi}\frac{dT}{d\tau} - \frac{\gamma}{q}\frac{dq}{d\tau} \\
&= -\left(\frac{\alpha}{W+T}\right)\frac{dW}{d\tau} - \left(\frac{\alpha}{W+T}\right)\frac{dT}{d\tau} \\
&\quad -\beta\frac{T}{g} + \beta\frac{(1-\tau)^{-1}\varepsilon Y}{g} - \beta\frac{1+\tau}{g}\frac{dT}{d\tau} - \frac{\gamma}{q}\frac{dq}{d\tau},
\end{aligned}$$

where the last line follows from definition of g in equation (4.7). Now, substitute in for $dq/d\tau$ using equation (4.13); this yields

$$\begin{aligned}
\text{DWL} &= -\left(\frac{\alpha}{W+T}\right)\frac{dW}{d\tau} - \left(\frac{\alpha}{W+T}\right)\frac{dT}{d\tau} \\
&\quad -\beta\frac{T}{g} + \beta\frac{(1-\tau)^{-1}\varepsilon Y}{g} - \beta\frac{1+\tau}{g}\frac{dT}{d\tau} \\
&\quad -\gamma\frac{T}{(1+q)\,g} + \gamma\frac{(1-\tau)^{-1}\varepsilon Y}{(1+q)\,g} - \gamma\frac{1+\tau}{(1+q)\,g}\frac{dT}{d\tau},
\end{aligned}$$

where the last line again follows from definition of g in equation (4.7). Finally, substitute in for $dW/d\tau$ using equation (4.11). This implies that

$$\begin{aligned}
\text{DWL} &= \frac{\alpha W}{g} + \frac{\alpha}{g}(1-\tau)^{-1}\varepsilon Y - \left(\frac{\alpha}{W+T}\right)\frac{dT}{d\tau} \\
&\quad -\beta\frac{T}{g} + \beta\frac{(1-\tau)^{-1}\varepsilon Y}{g} - \beta\frac{1+\tau}{g}\frac{dT}{d\tau} \\
&\quad -\gamma\frac{T}{(1+q)\,g} + \gamma\frac{(1-\tau)^{-1}\varepsilon Y}{(1+q)\,g} - \gamma\frac{1+\tau}{(1+q)\,g}\frac{dT}{d\tau},
\end{aligned} \tag{4.15}$$

where now the first line follows from definition of g in equation (4.7). Thus,

$$\text{DWL} = a(W, \ell, q, \cdot) + c(W, \ell, q, \cdot)\frac{dT}{d\tau} + b(W, \ell, q, \cdot)(1-\tau)^{-1}\varepsilon Y. \tag{4.16}$$

Of course, $a(W, \ell, q, \cdot)$, $b(W, \ell, q, \cdot)$, $c(W, \ell, q, \cdot)$, Y, and ε are endogenous and jointly determined.

Suppose that all parameters and variables that enter the DWL equation are observed, but utility is incorrectly assumed to be given by the sum of $\alpha \log(W + T)$ and $\beta \log(\ell)$. In this case, the last line of DWL in equation (4.15) is not taken into account, and DWL is therefore overestimated if and only if[11]

$$-\gamma\frac{T}{(1+q)g} + \gamma\frac{(1-\tau)^{-1}\varepsilon Y}{(1+q)g} - \gamma\frac{1+\tau}{(1+q)g}\frac{dT}{d\tau} < 0.$$

Using the definition of ε from equation (4.10), the previous holds if and only if

$$\begin{aligned}-\gamma\frac{T}{(1+q)g} + \gamma\frac{(1-\tau)^{-1}}{(1+q)g}W(1+\tau)\left(-\frac{dW}{d\tau}\left(\frac{1-\tau}{W}\right) - \left(\frac{1-\tau}{1+\tau}\right)\right) - \gamma\frac{1+\tau}{(1+q)g}\frac{dT}{d\tau} < 0 \\ \iff -(W+T) - (1+\tau)\frac{dW}{d\tau} - (1+\tau)\frac{dT}{d\tau} < 0. \end{aligned} \tag{4.17}$$

Now, note that total differentiation of the budget constraint yields

$$-\pi d\ell = (1+\tau)dW + Wd\tau + \phi dq,$$

where the last term makes use of $\phi = \frac{1}{A}\exp(q)$. Rearranging, this implies that

$$\frac{dW}{d\tau} = -\frac{W}{1+\tau} - \frac{\pi}{1+\tau}\frac{d\ell}{d\tau} - \frac{\phi}{1+\tau}\frac{dq}{d\tau}. \tag{4.18}$$

[11] It immediately follows from inspection of equation (4.15) that if the amenity is not a choice variable, then the third line of this equation is absent. Thus if all parameters and variables that enter equation (4.1) are observed and the amenity is not a choice variable, then even if utility is incorrectly assumed to be given by the sum of $\alpha \log(w + T)$ and $\beta \log(\ell)$, equation (4.1) yields the correct value for DWL.

Substituting in for $d\ell/d\tau$ from equation (4.8) and $dq/d\tau$ from equation (4.9) implies

$$\begin{aligned}\frac{dW}{d\tau} = -\frac{W}{1+\tau} - \frac{\pi}{1+\tau}\left(\frac{W+T}{x_1\pi} + \frac{1+\tau}{x_1\pi}\frac{dW}{d\tau} + \frac{1+\tau}{x_1\pi}\frac{dT}{d\tau}\right) \\ - \frac{\phi}{1+\tau}\left(\frac{W+T}{(1+q)\,x_2\phi} + \frac{1+\tau}{(1+q)\,x_2\phi}\frac{dW}{d\tau} + \frac{1+\tau}{(1+q)\,x_2\phi}\frac{dT}{d\tau}\right) \\ \implies (1+\tau)\frac{dW}{d\tau} = -W - \frac{W+T}{x_1} - \frac{1+\tau}{x_1}\frac{dW}{d\tau} - \frac{1+\tau}{x_1}\frac{dT}{d\tau} \\ - \frac{W+T}{(1+q)\,x_2} + -\frac{1+\tau}{(1+q)\,x_2}\frac{dW}{d\tau} + -\frac{1+\tau}{(1+q)\,x_2}\frac{dT}{d\tau}\end{aligned}$$

This implies that

$$\begin{aligned}\left(1+\tau+\frac{1+\tau}{x_1}+\frac{1+\tau}{(1+q)\,x_2}\right)\frac{dW}{d\tau} = -W - \frac{W+T}{x_1} - \frac{W+T}{(1+q)\,x_2} - \left(\frac{1+\tau}{x_1}+\frac{1+\tau}{(1+q)\,x_2}\right)\frac{dT}{d\tau} \\ \implies \left(\frac{(1+\tau)(1+q)\,x_1x_2 + (1+\tau)(1+q)\,x_2 + (1+\tau)\,x_1}{(1+q)\,x_1x_2}\right)\frac{dW}{d\tau} \\ = \frac{-W(1+q)\,x_1x_2 - (W+T)(1+q)\,x_2 - (W+T)\,x_1 - ((1+\tau)(1+q)\,x_2 + (1+\tau)\,x_1)\,dT/d\tau}{(1+q)\,x_1x_2},\end{aligned}$$

and therefore

$$\frac{dW}{d\tau} = \frac{-W(1+q)\,x_1x_2 - (W+T)(1+q)\,x_2 - (W+T)\,x_1 - (1+\tau)\left((1+q)\,x_2 + x_1\right)dT/d\tau}{(1+\tau)\left((1+q)\,x_1x_2 + (1+q)\,x_2 + x_1\right)}. \tag{4.19}$$

Substituting equation (4.19) into the inequality (4.17) therefore implies that the inequality holds if and only if

$$\begin{aligned}-(W+T) + \frac{W(1+q)\,x_1x_2 + (W+T)(1+q)\,x_2 + (W+T)\,x_1}{(1+q)\,x_1x_2 + (1+q)\,x_2 + x_1} \\ + \frac{(1+\tau)\left((1+q)\,x_2 + x_1\right)dT/d\tau}{(1+q)\,x_1x_2 + (1+q)\,x_2 + x_1} - (1+\tau)\frac{dT}{d\tau} < 0,\end{aligned}$$

that is, if and only if

$$
\begin{aligned}
-(W+T)(1+q)\,x_1x_2-(W+T)(1+q)\,x_2-(W+T)\,x_1 \\
+W(1+q)\,x_1x_2+(W+T)(1+q)\,x_2+(W+T)\,x_1 \\
+(1+\tau)\left((1+q)\,x_2+x_1\right)dT/d\tau-(1+\tau)\left((1+q)\,x_1x_2+(1+q)\,x_2+x_1\right)<0,
\end{aligned}
$$

which is true if and only if

$$
-T(1+q)\,x_1x_2-(1+\tau)(1+q)\,x_1x_2\frac{dT}{d\tau}<0.
$$

Of course, this is always true for $T, dT/d\tau > 0$, in which cases DWL is indeed overestimated by assuming away the existence of the amenity.

Linear amenity cost function: $\phi = q$. Assume instead that the cost function is linear so that the budget constraint satisfies

$$
\pi(1-\ell)=W(1+\tau)+q,
$$

the relevant FOCs yield

$$
\frac{\alpha}{g}=\frac{\beta}{\ell\pi}\implies\underbrace{\frac{\alpha}{\beta}}_{x_1}\ell=\frac{g}{\pi}
$$

and

$$
\frac{\alpha}{g}=\frac{\gamma}{q\phi}\implies\underbrace{\frac{\alpha}{\gamma}}_{x_2}q=g,
$$

which jointly imply that

$$
x_1\ell\pi=x_2q=(W+T)(1+\tau)=g.
$$

Using analogous procedures as detailed for the convex cost function case, it can then be shown that in the linear case

$$\frac{d\ell}{dq} = \frac{W+T}{x_1\pi} + \frac{(1+\tau)}{x_1\pi} + \frac{(1+\tau)}{x_2}\frac{dT}{d\tau},$$

$$\frac{dq}{d\tau} = \frac{W+T}{x_2} + \frac{1+\tau}{x_2}\frac{dW}{d\tau} + \frac{1+\tau}{x_2}\frac{dT}{d\tau},$$

and using the budget constraint

$$\frac{dW}{d\tau} = -\frac{\pi}{1+\tau}\frac{d\ell}{d\tau} - \frac{W}{1+\tau} - \frac{1}{1+\tau}\frac{dq}{d\tau}.$$

Using these three equations together implies that

$$\frac{dW}{d\tau} = \frac{-Wx_1x_2 - (W+T)(x_1+x_2) - (1+\tau)(x_1+x_2)\,dT/d\tau}{(1+\tau)(x_1x_2 + (x_1+x_2))}.$$

Moreover,

$$\begin{aligned}\text{DWL} \;=\; & \frac{\alpha W}{g} + \frac{\alpha}{g}(1-\tau)^{-1}\varepsilon Y - \left(\frac{\alpha}{W+T}\right)\frac{dT}{d\tau} \\ & -\beta\frac{T}{g} + \beta\frac{(1-\tau)^{-1}\varepsilon Y}{g} - \beta\frac{1+\tau}{g}\frac{dT}{d\tau} \\ & -\gamma\frac{T}{g} + \gamma\frac{(1-\tau)^{-1}\varepsilon Y}{g} - \gamma\frac{1+\tau}{g}\frac{dT}{d\tau}. \end{aligned} \tag{4.20}$$

Suppose that all parameters and variables that enter the DWL equation are observed, but utility is incorrectly assumed to be given by the sum of $\alpha\log(W+T)$ and $\beta\log(\ell)$. In this case, the last line of the DWL is not taken into account, and DWL is therefore overestimated if and only if

$$-\gamma\frac{T}{g} + \gamma\frac{(1-\tau)^{-1}\varepsilon Y}{g} - \gamma\frac{1+\tau}{g}\frac{dT}{d\tau} < 0.$$

Substituting in for ε and then for $dW/d\tau$ implies that the previous is true if and only if

$$\begin{aligned}
& -(W+T)-(1+\tau)\frac{dW}{d\tau}-(1+\tau)\frac{dT}{d\tau}<0 \\
\iff & -(W+T)+\frac{Wx_1x_2+(W+T)(x_1+x_2)+(1+\tau)(x_1+x_2)\,dT/d\tau}{x_1x_2+x_1+x_2}-(1+\tau)\frac{dT}{d\tau}<0 \\
\iff & -(W+T)\,x_1x_2-(W+T)(x_1+x_2)+Wx_1x_2+(W+T)(x_1+x_2) \\
& +(1+\tau)(x_1+x_2)\frac{dT}{d\tau}-(x_1x_2+x_1+x_2)(1+\tau)\frac{dT}{d\tau}<0,
\end{aligned}$$

which holds if and only if

$$-Tx_1x_2-x_1x_2\,(1+\tau)\frac{dT}{d\tau}<0.$$

Again, this is true for $T, dT/d\tau > 0$, in which cases DWL is indeed overestimated when assuming away the existence of the amenity.

Over/under estimation in linear versus convex cost function scenarios. From above, it is straightforward that if the cost function is convex but the amenity is assumed away, then the DWL is overestimated by,

$$\frac{T(1+q)\,x_1x_2+(1+\tau)(1+q)\,x_1x_2dT/d\tau}{(1+q)\,x_1x_2+(1+q)\,x_2+x_1}$$

while in the linear case it is overestimated by

$$\frac{Tx_1x_2+x_1x_2\,(1+\tau)\,dT/d\tau}{x_1x_2+x_1+x_2}.$$

Assuming identical parameter values and $dT/d\tau$, it follows that overestimation under the convex case is higher if and only if

$$\frac{T(1+q)x_1x_2+(1+\tau)(1+q)x_1x_2dT/d\tau}{(1+q)x_1x_2+(1+q)x_2+x_1}>\frac{Tx_1x_2+x_1x_2(1+\tau)dT/d\tau}{x_1x_2+x_1+x_2}$$
$$\iff T(1+q)(x_1x_2)^2+T(1+q)x_1^2x_2+T(1+q)x_1x_2^2$$
$$+\left((1+\tau)(1+q)(x_1x_2)^2+(1+\tau)(1+q)x_1^2x_2+(1+\tau)(1+q)x_1x_2^2\right)dT/d\tau$$
$$>T(1+q)(x_1x_2)^2+T(1+q)x_1x_2^2+Tx_1^2x_2$$
$$+\left((1+\tau)(1+q)(x_1x_2)^2+(1+\tau)(1+q)x_1x_2^2+(1+\tau)x_1^2x_2\right)dT/d\tau,$$

which therefore is true if and only if

$$Tqx_1^2x_2+(1+\tau)qx_1^2x_2\frac{dT}{d\tau}>0;$$

this inequality always holds for $T, dT/d\tau>0$. Thus, DWL under the convex cost function case is smaller than under the linear case (note that in reality, optimal endogenous variables, and therefore, ε differ in the linear and convex cases).

This raises an interesting, related question. Even if the amenity is accounted for, how is DWL estimation affected by incorrectly assuming a linear rather than convex cost, or vice versa? Suppose that all parameters and variables that enter the DWL equation are observed, but the cost of the amenity is incorrectly assumed to be linear. Then, using equations (4.15) and (4.20) DWL is underestimated if and only if

$$-\gamma\frac{T}{g}+\gamma\frac{(1-\tau)^{-1}\varepsilon Y}{g}-\gamma\frac{1+\tau}{g}\frac{dT}{d\tau}+\gamma\frac{T}{(1+q)g}-\gamma\frac{(1-\tau)^{-1}\varepsilon Y}{(1+q)g}+\gamma\frac{1+\tau}{(1+q)g}\frac{dT}{d\tau}<0$$
$$\iff -T-(1-\tau)^{-1}\varepsilon Y-(1+\tau)\frac{dT}{d\tau}<0,$$

which always holds. Of course, it follows that if all parameters and variables that enter the DWL equation are observed, but the cost of the amenity is incorrectly assumed to be convex, then DWL is overestimated.

4.7.2 Linear Utility

Convex amenity cost function: $\phi = \left(\frac{1}{A}\right)\exp\left(q\right)$. Assume ℓ is inelastic at some level $\bar{\ell}$,

$$u = (W + T) + \bar{\ell} + q,$$

and utility maximization occurs subject to $\pi\left(1 - \bar{\ell}\right) = W\left(1 + \tau\right) + \phi$. Of course,

$$W = \frac{\pi\left(1 - \bar{\ell}\right) - \phi}{1 + \tau} \implies W d\tau + \left(1 + \tau\right) dW = -\phi' dq \implies \frac{dW}{d\tau} = \frac{-W}{1 + \tau} - \frac{\phi}{1 + \tau}\frac{dq}{d\tau},$$

which follows by use of $\phi' = \phi = \left(\frac{1}{A}\right)\exp(q)$. Substituting the wage equation into the utility function, the first-order condition (FOC) with respect to q implies that

$$\phi = 1 + \tau \implies \phi' dq = d\tau \implies \frac{dq}{d\tau} = \frac{1}{\phi}.$$

Substituting this into the wage equation total differential implies that

$$\frac{dW}{d\tau} = \frac{-\left(1 + W\right)}{1 + \tau},$$

which follows by use of $\phi' = \phi$. As before,

$$\varepsilon = \frac{d\log\left(\left(1 + \tau\right)W\right)}{d\log\left(1 - \tau\right)} = -\frac{dW}{d\tau}\left(\frac{1 - \tau}{W}\right) - \left(\frac{1 - \tau}{1 + \tau}\right).$$

Substituting in for $dW/d\tau$ implies that

$$\varepsilon = \frac{1 - \tau}{W\left(1 + \tau\right)} > 0.$$

Now, totally differentiate the utility function; this yields

$$\begin{aligned}\text{DWL} &= -\frac{du}{d\tau} = -\frac{dW}{d\tau} - \frac{dT}{d\tau} - \frac{dq}{d\tau} \\ &= \frac{W}{1+\tau} + \frac{(1-\tau)^{-1}\varepsilon Y}{1+\tau} - \frac{dT}{d\tau} - \frac{1}{\phi},\end{aligned}$$

where the second line follows, as before, by rearrangement of the definition of ε, and by use of the total differential of the first-order condition for q. If utility is assumed to be $u = W + T$ by assuming away the amenity, then the associated DWL is

$$\frac{W}{1+\tau} + \frac{(1-\tau)^{-1}\varepsilon Y}{1+\tau} - \frac{dT}{d\tau},$$

which, conditional on observing all variables and parameters correctly, overestimated the true DWL if and only if $1/\phi > 0$, which always holds.

Linear amenity cost function: $\phi = q$. Assume ℓ is inelastic at some level $\bar{\ell}$,

$$u = (W + T) + \bar{\ell} + q,$$

and utility maximization occurs subject to $\pi\left(1 - \bar{\ell}\right) = W(1+\tau) + q$. In (q, W) space the slopes of the indifference curves and budget line, are, respectively

$$dq/dW = -1 \text{ and } dq = -(1+\tau).$$

Hence, an interior solution can only occur if $\tau = 0$. For any $\tau > 0$, $W = 0$ optimally and the worker's entire compensation occurs through amenity provision.

Of course, given the budget constraint

$$q = \pi\left(1 - \bar{\ell}\right) - W(1+\tau).$$

Suppose $\tau = 0$ and increases to $\tau > 0$. Then, in (q, W) space the budget line becomes steeper, and utility is maximized at $W = 0$ and $q = \pi\left(1 - \bar{\ell}\right)$. Because this point was on the original budget line, then after implementation of the tax, holding T constant utility does not change, hence, DWL=0, otherwise DWL= $dT/d\tau$. Clearly, if $\tau > 0$ and increases to $\tau' > \tau$, utility is maximized both before and after the change at $W = 0$ and $q = \pi\left(1 - \bar{\ell}\right)$, so it is again the case that holding T constant DWL= $-dT/d\tau$. Note that in fact, under either case, as long as T is a fraction of taxable income, then in the former case $dT/d\tau < 0$ and in the later case $dT/d\tau = 0$. In the former case, this occurs because after the increase in taxes, taxable income drops to zero. In the later case, taxable income is zero both before and after the change in taxes. It immediately follows that in the later case the implied value of ε is infinity (decreasing taxes makes taxable income go from zero to a positive amount), whereas in the later it is zero (decreasing taxes results in no change in taxable income, which is zero both before and after the change in taxes).

If the amenity is assumed away, then u is incorrectly assumed to be $W + T$, and use of the budget constraint yields

$$W = \frac{\pi\left(1 - \bar{\ell}\right)}{1 + \tau},$$

implying that $dW/d\tau = -W/\left(1 + \tau\right)$. In this case

$$\text{DWL} = -dW/d\tau = \frac{W}{1 + \tau} + \frac{\left(1 - \tau\right)^{-1}\varepsilon Y}{1 + \tau}$$

by definition of ε, meaning that in reality $\varepsilon = 0$ and DWL= $W/\left(1 + \tau\right)$. Note that if presence of the amenity is incorrectly assumed away, as long as relevant variables and parameters are correctly observed, including the value of ε, whether this be infinity or zero (see discussion above), then DWL is overestimated; if $\varepsilon = 0$, then the "incorrect" DWL equation form above implies DWL= $W/\left(1 + \tau\right)$. If $\varepsilon = \infty$, then the "incorrect" DWL equation form above implies DWL= ∞.

Table 4.1: Summary statistics

	Mean[1]
Age	29.5 years
Sex	49.7% female
Race	7.6% Hispanic, 13.1% black, 79.3% non-black,non-Hispanic
Education	12.6 years
Tenure	30.2 months
Wage	$15.8/hour

1. Data are from the NLSY79.

Table 4.2: Simulated and empirical moments

Moments	**Simulated**	**Target**
mean tenure	32.8	30.5
variance of tenure	1414	1419
skewness of tenure	2.32	2.29
b as a fraction of flow output	0.370	0.500
unemployment rate	0.040	0.065
job-to-job probability	0.013	0.069
unemployment-to-employment probability	0.230	0.193
wage tenure correlation	0.164	0.225
wage coefficient of variation	0.480	0.489

Table 4.3: Effects of wage taxation with endogenous amenities

Wage tax	Social welfare	Taxable income	Social welfare elasticity[1]	Taxable income elasticity
0 percent	100.0	100.0		
10 percent	99.6	90.0	0.04	1.00
20 percent	98.9	83.6	0.06	0.63
30 percent	97.9	79.3	0.08	0.39
40 percent	96.2	73.2	0.11	0.52
50 percent	92.6	60.7	0.21	1.02
60 percent	87.3	42.6	0.27	1.59
70 percent	78.3	15.5	0.37	3.51
80 percent	63.3	-	-	-
90 percent	31.2	-	-	-

1. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.
2. Elasticities are with respect to the net-of-tax rate.
3. Amenities are endogenously supplied.

Table 4.4: Effects of wage taxation with exogenous amenities

Wage tax	Social welfare	Taxable income	Social welfare elasticity[1]	Taxable income elasticity
0 percent	100.0	100.0		
10 percent	99.6	92.2	0.04	0.77
20 percent	98.9	89.1	0.05	0.29
30 percent	98.1	86.7	0.06	0.20
40 percent	97.0	83.4	0.08	0.26
50 percent	94.5	75.4	0.14	0.55
60 percent	91.9	67.8	0.12	0.47
70 percent	87.8	56.5	0.16	0.63
80 percent	82.2	41.9	0.16	0.74
90 percent	75.9	26.1	0.12	0.68

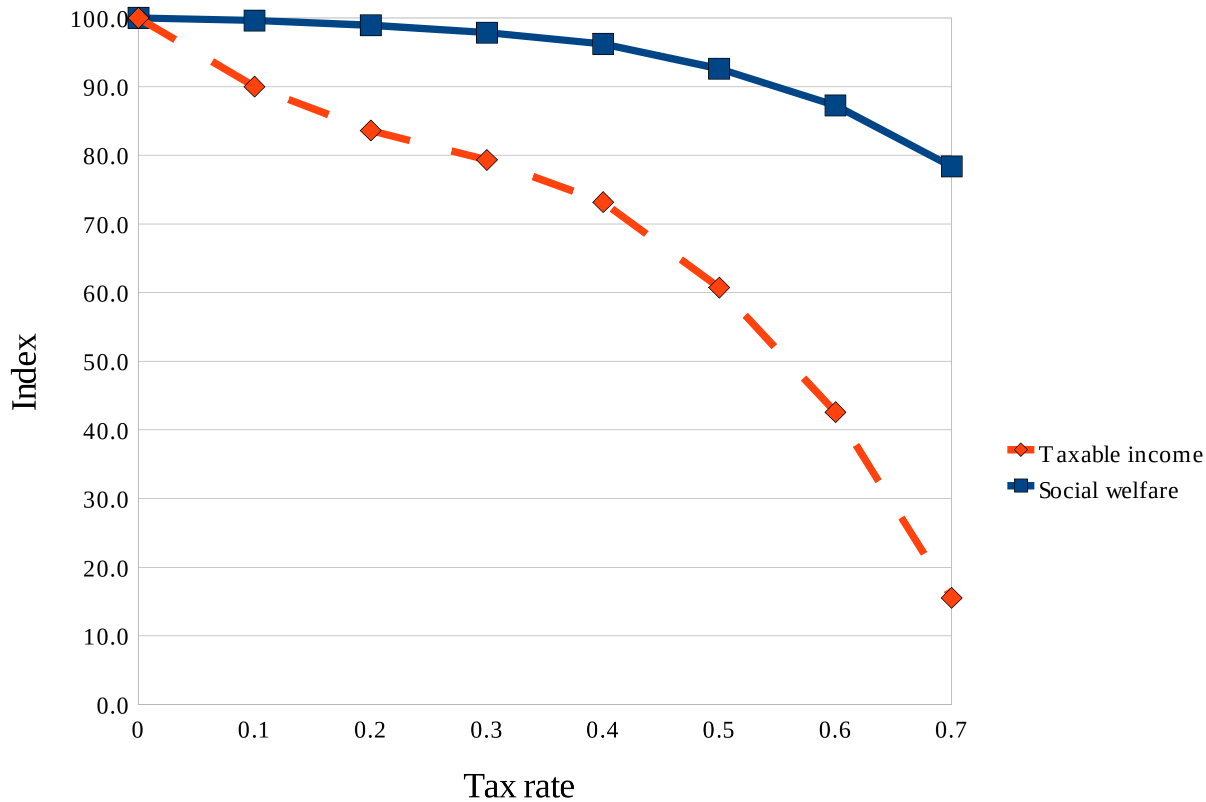
1. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.
2. Elasticities are with respect to the net-of-tax rate.

Table 4.5: Dynamic effect of wage taxation

Wage tax	Social welfare	Taxable income
Pre-increase baseline	100.0	100.0
1 month	98.1	98.9
6 months	98.1	98.6
12 months	98.2	98.1
18 months	98.2	97.7
24 months	98.4	97.4
30 months	98.2	97.1
36 months	98.1	96.8
42 months	98.4	96.6
48 months	98.5	96.1

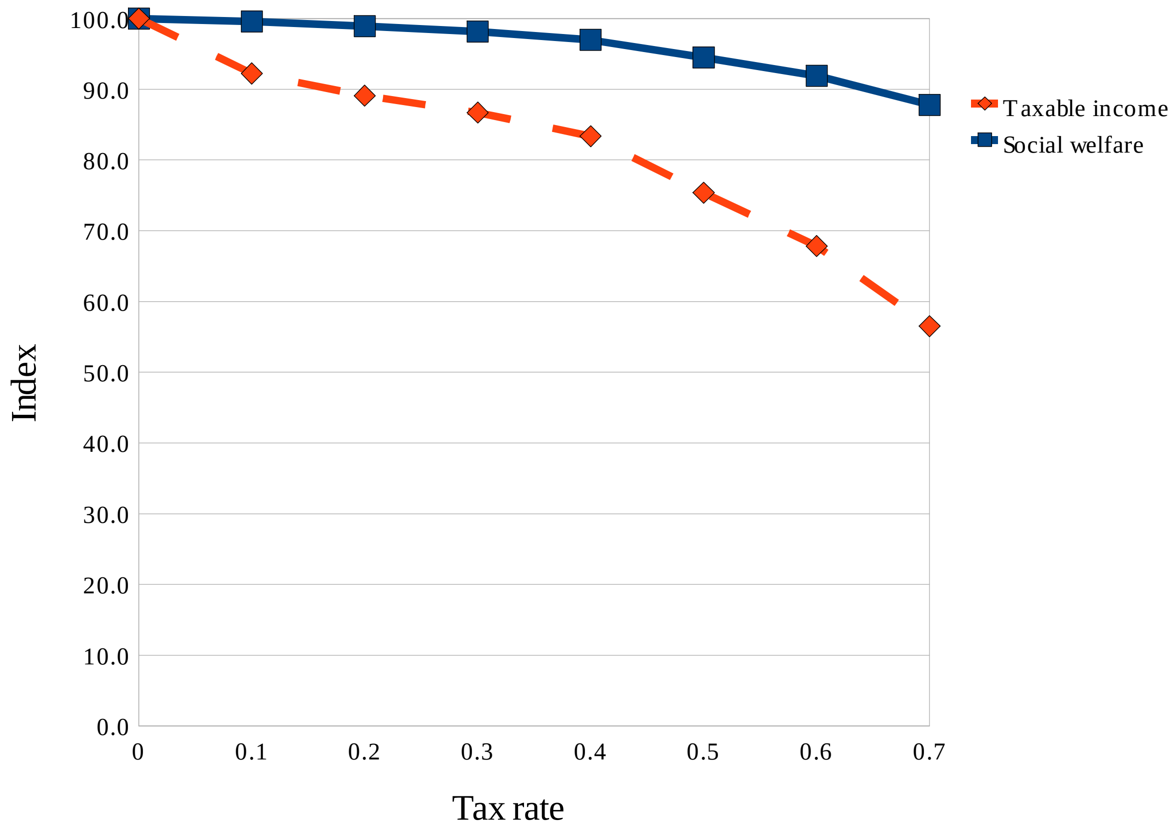
1. Social welfare and taxable income are indexed to 100 at the original 24.8 percent tax rate.
2. The wage tax rate unexpectedly rises 10 percentage points in all periods after the initial.
3. Amenities are endogenously supplied.

Figure 4.2: Effects of wage taxation with endogenous amenities



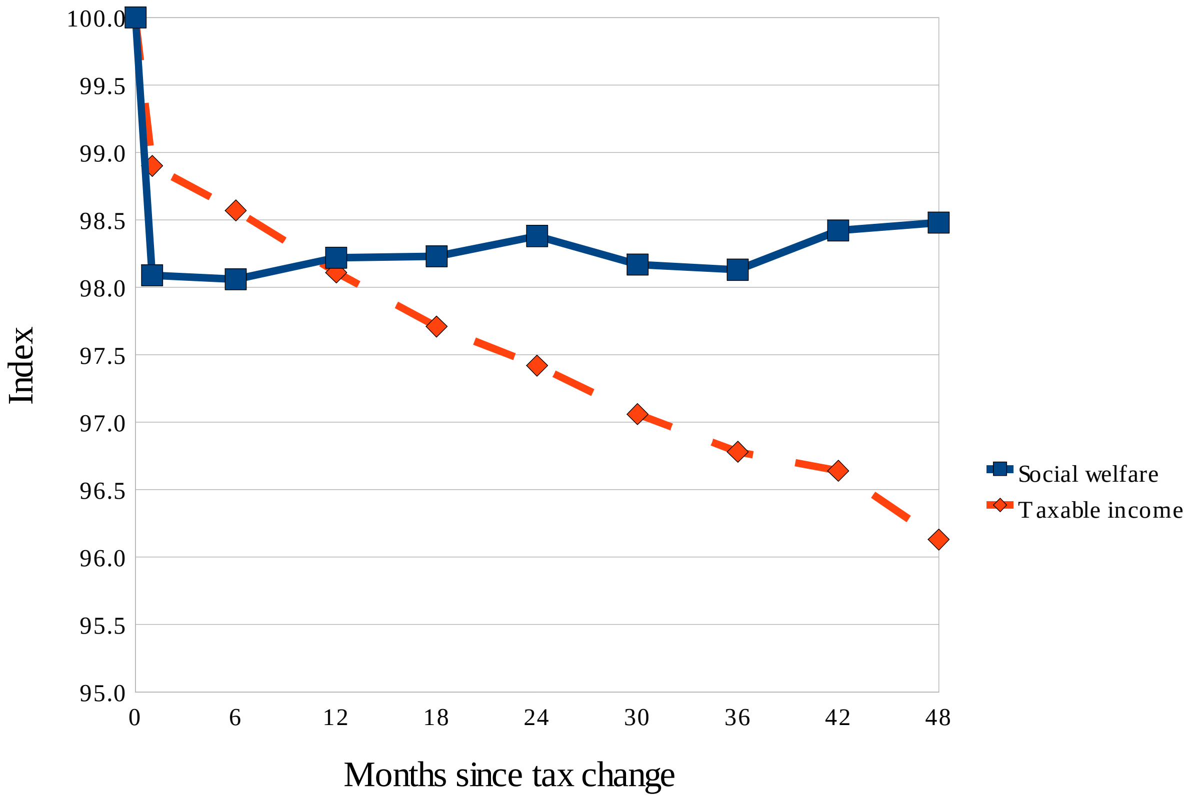
1. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.
2. Amenities are endogenously supplied.

Figure 4.3: Effects of wage taxation with exogenous amenities



1. Social welfare and taxable income are indexed to 100 at a 0 percent tax rate.

Figure 4.4: Dynamic effect of wage taxation



1. Social welfare and taxable income are indexed to 100 at the original 24.8 percent tax rate.
2. The wage tax rate unexpectedly rises 10 percentage points in all periods after the initial.
3. Amenities are endogenously supplied.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abowd, J., F. Kramarz, S. Perez-Duarte, and I. Schmutte (2009): "A formal test of assortative matching in the labor market," *NBER Working Paper*, 15546.

Acemoglu, D. and R. Shimer (1999): "Efficient unemployment insurance," *Journal of Political Economy*, 107, no. 5, 893–928.

Ackerberg, D. (2001): "A new use of importance sampling to reduce computational burden in simulation estimation," *NBER Technical Working paper*, 273.

Andrews, F. and S. Withey (1976): *Social indicators of well-being: Americans' perceptions of life quality*, New York: Plenum.

Barlevy, G. (2002): "The sullying effect of recessions," *The Review of Economic Studies. Vol. No.1,*, 69, no. 1, 65–96.

——— (2008): "Identification of search models using record statistics," *Review of Economic Studies*, 75, 2964.

Becker, D. (2009): "Non-wage job characteristics and the case of the missing margin," *Working paper*.

Belzil, C. (2001): "Unemployment insurance and subsequent job duration: job matching versus unobserved heterogeneity," *The Journal of Applied Econometrics*, 16, no. 5, 619–36.

Benjamin, D., O. Heffetz, M. Kimball, and A. Rees-Jones (2010): "Do people seek to maximize happiness? Evidence from new surveys," *NBER Working Paper*, 16489.

Blackwell, D. (1965): "Discounted dynamic programming," *The Annals of Mathematical Statistics*, 36, 226–235.

Bottan, N. and R. Truglia (2011): "Deconstructing the hedonic treadmill: is happiness autoregressive?" *Journal of Socio-Economics*, 40, 22436.

Bowlus, A. (1995): "Matching workers and jobs: cyclical fluctuations in match quality," *Journal of Labor Economics*, 13, no. 2, 335–50.

Bradburn, N. (1969): *The structure of psychological well-being*, Chicago: Aldine.

Burdett, K. and D. Mortensen (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 39, no. 2, 257–73.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage bargaining with on-the-job search: theory and evidence," *Econometrica*, 74, 323–64.

Campbell, A., P. Converse, and W. Rodgers (1976): *The quality of American life*, New York: Russell Sage.

Card, D., F. Devicienti, and A. Maida (2010): "Rent-sharing, holdup, and wage: evidence from matched panel data," *NBER Working Paper*, 16192.

Centeno, M. and A. Novo (2006): "The impact of unemployment insurance generosity on match quality distribution," *Economics Letters*, 93, 235–41.

——— (2009): "Reemployment wages and UI liquidity effect: a regression discontinuity approach," *Portugese Economic Journal*, 8, 45–52.

Chetty, R. (2008): "Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance," *NBER*, 13844.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from Danish tax records," *NBER Working Paper*, No. 15617.

Clark, A., E. Diener, Y. Georgellis, and R. Lucas (2008a): "Lags and leads in life satisfaction: a test of the baseline hypothesis," *The Economic Journal*, 118, F222–F243.

Clark, A., P. Frijters, and M. Shields (2008b): "Relative income, happiness, and utility: an explanation for the Easterlin aradox and other puzzles," *Journal of Economic Literature*, 46, 95–144.

Deaton, A., J. Fortson, and R. Tortora (2008): "Life (evaluation), HIV/AIDS, and death in Africa," *Working paper.*

Di Tella, R., R. MacCulloch, and A. Oswald (2001): "Preferences over inflation and unemployment: evidence from surveys of happiness," *The American Economic Review*, 91, 335–41.

Diener, E., R. E. Lucas, and C. Napa (2006): "Beyond the hedonic treadmill: revising the adaptation theory of well-being," *American Psychologist*, May-June, 305–14.

Dolan, P. and D. Kahneman (2008): "Interpretations of utility and their implications for the valuation of health," *The Economic Journal*, 118, 215–34.

Feldstein, M. (1999): "Tax avoidance and the deadweight loss of the income tax," *The Review of Economics and Statistics*, 81, 674–80.

Finkelstein, A., E. Luttmer, and M. Notowidigdo (2008): "What good is wealth without health? The effect of health on the marginal utility of consumption," *NBER Working Paper*, 14089.

Frederick, S. and G. Loewenstein (1999): *Well-being: the foundations of hedonic psychology*, Russell Sage, chap. Hedonic adaptation.

Friedman, M. and S. Kuznets (1954): *Income from Independent Professional Practice*, NBER.

Fujita, S. and M. Nakajima (2009): "Worker flows and job flows: a quantitative investigation," *Federal Reserve Bank of Philadelphia Research Department Working Papers*, 09-33.

Gourieroux, C. and A. Monfort (1996): *Simulation-based econometric methods*, Oxford University Press.

Gruber, J. and S. Mullainathan (2005): "Do cigarette taxes make smokers happier," *Advances in Economic Analysis and Policy*, 5.

Hagedorn, M. and I. Manovskii (2010): "Search frictions and wage dispersion," *Working paper.*

——— (forthcoming): "Job selection and wages over the business cycle," *The American Economic Review.*

Hall, R. (2005): "Employment efficiency and sticky wages: evidence from flows in the labor market," *NBER Working Paper*, 11183.

Headey, B., R. Muffels, and G. Wagner (2010): "Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness," *Proceedings of the National Academy of Sciences, Early Edition.*

Jovanovic, B. (1979): "Job matching and the theory of turnover," *The Journal of Political Economy*, 87, No. 5, 972–90.

Kahn, L. (2008): "Job durations, match quality and the business cycle: what we can learn from firm fixed effects," *Working paper.*

Kimball, M. and R. Willis (2010): "Utility and Happiness," *University of Michigan manuscript.*

Lucas, R., A. Clark, Y. Georgellis, and E. Diener (2004): "Unemployment alters the set point for life satisfaction," *Psychological Science*, 15, 8–13.

McCall, B. (1990): "Occupational matching: a test of sorts," *Journal of Political Economy*, 98, 45–69.

Menzio, G. and S. Shi (2008): "On-the-job search and business cycles," *Working paper.*

Mortensen, D. (2003): *Wage dispersion: why are similar people paid differently*, MIT Press.

Mortensen, D. and C. Pissarides (1994): "Job creation and job destruction in the theory of unemployment," *The Review of Economic Studies*, 61, no. 3, 397–415.

Nagypal, E. (2007): "Learning by doing vs. learning about match quality: can we tell them apart?" *The Review of Economic Studies*, 74, 537–66.

——— (2008): "Worker reallocation over the business cycle: the importance of job-to-job transitions," *Working paper.*

Nunn, R. (2012): "Match Quality with Unpriced Amenities," *Working paper.*

Oswald, A. and N. Powdthavee (2008): "Does happiness adapt? A longitudinal study of disability with implications for economists and judges," *Journal of Economics*, 92, 1061–77.

Ours, J. V. and M. Vodopivec (2008): "Does reducing unemployment insurance generosity reduce job match quality?" *Journal of Public Economics*, 92, 684–95.

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica Vol.70 No.6, 2295-350*, 70, 2295–350.

Powdthavee, N. (2009): "What happens to people before and after disability? Focusing effects, lead effects, and adaptation in different areas of life," *Social Science & Medicine*, 69, 1834–44.

Riis, J., G. Loewenstein, J. Baron, C. Jepson, A. Fagerlin, and P. Ubel (2005): "Ignorance of hedonic adaptation to hemodialysis: a study using ecological momentary assessment," *Journal of Experimental Psychology*, 134, 3–9.

Rosen, S. (1974): "Hedonic prices and implicit markets: product differentiation in pure competition," *Journal of Political Economy*, 82, 34–55.

——— (1986): *The theory of equalizing differences*, Elsevier.

Saez, E., J. Slemrod, and S. Giertz (2009): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *NBER Working Paper*, 15012.

Sargent, T. (1987): *Dynamic macroeconomic theory*, Harvard University Press.

Schwarz, N. (1987): *Mood as Information: On the Impact of Moods on Evaluation of Ones Life*, Heidelberg: Springer-Verlag.

Schwarz, N. and F. Strack (1999): *Well-being: The foundations of hedonic psychology*, New York: Russell-Sage, chap. Reports of subjective well-being: Judgmental processes and their methodological implications, 61–84.

Shimer, R. (2003): "Dynamics in a model of on-the-job search," *Working paper.*

——— (2005): "The cyclical behavior of equilibrium unemployment and vacancies," *The American Economic Review*, 95, 25–49.

——— (2006): "On-the-job search and strategic bargaining," *European Economic Review*, 50, 811–30.

Stevenson, B. and J. Wolfers (2009): "The Paradox of Declining Female Happiness," *American Economic Journal: Economic Policy*.

Sullivan, P. and T. To (2011): "Search and non-wage job characteristics," *Working paper*.

Tauchen, G. (1986): "Finite state markov chain approximations to univariate and vector autoregressions," *Economic Letters*, 20, 177–81.

Wolpin, K. (1987): "Estimating a structural search model: the transition from school to work," *Econometrica*, 55, 801–17.

Woodcock, S. (2007): "Match effects," *SFU Discussion Paper*, dp07-13.

——— (2008): "Wage differentials in the presence of unobserved worker, firm, and match heterogeneity," *Labour Economics*, 15, 772–94.