Essays on Urban and Labor Economics

by

Aurel Hizmo

Department of Economics
Duke University

Date: ____________________

Approved:

____
Patrick Bayer, Supervisor

____
Peter Arcidiacono

____
Andrew Patton

____
James Roberts

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2011

Abstract

Essays on Urban and Labor Economics

by

Aurel Hizmo

Department of Economics
Duke University

Date: ________________________

Approved:

____
Patrick Bayer, Supervisor

____
Peter Arcidiacono

____
Andrew Patton

____
James Roberts

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2011

Copyright © 2011 by Aurel Hizmo
All rights reserved except the rights granted by the
Creative Commons Attribution-Noncommercial Licence

Abstract

In the first chapter of this dissertation I develop a flexible and estimable equilibrium model that jointly considers location decisions of heterogeneous agents across space, and their optimal portfolio decisions. Merging continuous-time asset pricing with urban economics models, I find a unique sorting equilibrium and derive equilibrium house and asset prices in closed-form. Risk premia for homes depend on both aggregate and local idiosyncratic risks, and equilibrium returns for stocks depend on their correlation with city-specific income and house price risk. In equilibrium, very risk-averse households do not locate in risky cities although they may have a high productivity match with those cities. I estimate a version of this model using house price and wage data at the metropolitan area level and provide estimates for risk premia for different cities. The estimated risk premia imply that homes are on average about \$20000 cheaper than they would be if owners were risk-neutral. This estimate is over \$100000 for volatile coastal cities. I simulate the model to study the effects of financial innovation on equilibrium outcomes. For reasonable parameters, creating assets that correlate with city-specific risks increase house prices by about 20% and productivity by about 10%. The average willingness to pay for completing markets per homeowner is between \$10000 and \$20000. Productivity is increased due to a unique channel: lowering the amount of non-insurable risk decreases the households' incentive to sort on these risks, which leads to a more efficient allocation of human capital in the economy.

The second chapter of this dissertation studies ability signaling in a model of employer learning and statistical discrimination. In traditional signaling models, education provides a way for individuals to sort themselves by ability. Employers in turn use education to statistically discriminate, paying wages that reflect the average productivity of workers with the same given level of education. In this chapter, we provide evidence that graduating from college plays a much more direct role in revealing ability to the labor market. Using the NLSY79, our results suggest that ability is observed nearly perfectly for college graduates. In contrast, returns to AFQT for high school graduates are initially very close to zero and rise steeply with experience. As a result, from very beginning of the career, college graduates are paid in accordance with their own ability, while the wages of high school graduates are initially unrelated to their own ability. This view of ability revelation in the labor market has considerable power in explaining racial differences in wages, education, and the returns to ability. In particular, we find a 6-10 percent wage penalty for blacks (conditional on ability) in the high school market but a small positive black wage premium in the college labor market. These results are consistent with the notion that employers use race to statistically discriminate in the high school market but have no need to do so in the college market.

To Khana

Contents

List of Tables

List of Figures

Acknowledgements

I am extremely grateful to my advisor, Patrick Bayer, for providing me with guidance, inspiration, and love for economics. I also thank the committee members Peter Arcidiacono, Andrew Patton and James Robers for their encouragement and support. I also thank Robert McMillan, Tim Bollerslev, Vish Viswanathan, Chris Timmins, and the participants at Duke Seminars and the ERID Housing Dynamics Conference for their helpful comments.

The second chapter of my dissertation is joint work with Patrick Bayer and Peter Arcidiacono.

I am also grateful to my parents, Teuta and Maksim Hizmo, for their infectious blind faith that I would someday succeed, and my partner Regina for providing me with perspective on life.

1

Risk in Housing Markets: An Equilibrium Approach

1.1 Introduction

Throughout their lifetimes, households face many long-term economic risks, most of which may be difficult to insure against.[1] Social safety nets and insurance mechanisms are particularly limited for long-term income and house price risk. The inability to efficiently spread risk across many households has direct adverse welfare effects since individual households may bear most of the burden of negative income and housing wealth shocks. Also, because of risk-aversion, many talented individuals may forgo brilliant careers and investment opportunities that are deemed too risky to undertake. Effective risk sharing can therefore limit downside risk and also enable individuals to make better choices, which leads to higher productivity and welfare in the economy. Such considerations have motivated Shiller (1993, 2003) and others to call for the creation of financial instruments to enable widespread sharing of risks

[1] Many studies test and strongly reject the hypothesis that people share risks effectively. Some examples include Zeldes (1989), Cochrane (1991), Hayashi, Altonji and Kotlikoff (1996), Athanasoulis and van Wincoop (2000).

that are not directly traded in equity markets.

Risks management considerations are especially important when a young household chooses a combined labor and housing market. Since most homeowners live near their workplace, they are exposed to a considerable amount of location-specific income and house price risk. From a portfolio management point of view, households are very much concerned about location-specific risk since a disproportionate share of their wealth is invested in one particular house, and they cannot reoptimize very frequently due to large reallocation costs.[2] What exacerbates this problem even more is that much of the local income and house price risk may not be easily diversified away by individual homeowners.[3] The riskiness of a particular location can therefore simultaneously affect house prices and households' location and portfolio choice decisions. In turn, financial innovation that allows households to better control their exposure to these risks may have a major impact on labor market choice, home prices and welfare.

The goal of this paper is to show both theoretically and empirically how exposure to uninsurable city-specific risks affects house prices, as well as household location and portfolio decisions. Understanding the nature and magnitude of these effects is crucial in evaluating the benefits of creating new financial instruments that facilitate risk sharing, such as those proposed by Shiller (1993, 2003) . To that end, I first develop an equilibrium theory that shows in a transparent way how location-specific

[2] Tracy, Schneider, and Chan (1999) report housing wealth comprises about two-thirds of the typical households portfolio.

[3] The correlation between the stock market and average labor income is usually estimated to be close to zero. See for example Cocco, Gomes, and Maenhout (2005). Campbell and Viceira (2002) do find a positive correlation of income with lagged stock returns that is as high as .5 for college graduates. The correlation between housing returns and the aggregate stock market or REITs is also found to be very low (see Flavin and Yamashita (2002), and Hinkelmann and Swindler (2006)). Hinkelmann and Swindler (2006) also find very little correlation between prices of futures contracts (not including the new housing futures) and house price returns, stressing that the creation of housing futures should be very beneficial for hedging purposes. Constructing local stock price indexes for largest employers from 16 cities in California, Hizmo (2010) reports a correlation of .36 between the local house price returns and the local stock price index.

risk is capitalized into house prices and how it affects productivity and welfare in the economy. The model is estimated using wage and price data for individual metropolitan areas in the US, providing empirical support for the idea that risk is priced in housing markets. I then simulate the model to study the benefits from creating financial assets that correlate with city-specific house prices and income.

The backbone of the paper consists of a novel micro-founded dynamic equilibrium model that merges standard methodologies used in urban economics, which study sorting and spatial properties of the problem, with models from continuous-time finance that study financial assets. To my knowledge, this is the first flexibly estimable model that simultaneously considers risk-aversion, multiple sources of uncertainty, rich agent heterogeneity, sorting, portfolio choice and asset prices in one unified framework. The main features in the model that drive most of the interesting results are that markets are incomplete, in that there is no perfect hedging of risks, houses are indivisible, and agents are heterogeneous in terms of their productivity and risk preferences.

The general setting of the model is very similar to that in Ortalo-Magne and Prat (2010) (OMP hereafter).[4] Each instant, a generation of households are born and live for T periods. Households initially choose a single labor and housing market from a system of cities. Once households choose a city to live in, they are assumed to live there until the end of their life. Unlike the main specifications of OMP, I

[4] OMP is the first paper to theoretically show the links between location and portfolio choice in an equilibrium model. Building on their elegant work, I extend the theory in number of important dimensions. Since the main focus in OMP is not empirical, income in different cities follows a random walk (in discrete time) with no drift and so do stock prices. I consider city-specific income processes that are richer and match empirical patterns of wage processes. Here, the growth in wages is modeled as a city-specific mean reverting process. Also, because of the continuous time setting, this model can handle stock prices that follow geometric Brownian motions with drift, which is in line with the modern continuous-time finance literature. In addition, this model also includes amenities, heterogeneity in preferences for local amenities, and heterogeneity in risk-aversion. In terms of the equilibrium, OMP conjecture a price process and show the existence of an equilibrium, while saying relatively little about about uniqueness. This paper proves that a linear stationary equilibrium is unique in its class, and does not rely on the existence of "hyper-marginal" households.

model houses as indivisible assets. Each household must buy one house in order to live in a city, which gives them access to local wages and amenities. Households also have access to a risk-free asset and the stock market, and can adjust their portfolios continuously. At the end of their lives, households sell their homes and their financial assets, and consume their terminal wealth. Because they cannot hold shares of homes in different cities, households have to resort to using stocks to insure against their income and house price risks. Because stocks may not be correlated with every dimension of risk that households face, households may face city-specific uninsurable risks.

The solution to the household problem involves two steps. Conditional on location choice, households decide how to invest their wealth in the financial market given their local income and house price process. Taking into account the utility derived from these optimal portfolio decisions, households choose the city that provides them with the highest expected utility. The portfolio choice problem with exogenous income that households face in this model is studied extensively in the literature. Generally the solutions are either found numerically and/or under the assumption that markets are complete.[5] Using transformations of the Hamilton-Jacobi-Bellman equation similar to Henderson (2005), I derive a closed-form solution of the portfolio choice with housing, exogenous income and incomplete markets. The explicit solution is crucial here since it is used to solve for the equilibrium allocation of individuals across space.

Given optimal portfolio decisions, I then solve for the spatial allocation of households. The equilibrium concept used here follows in the tradition of classic urban

[5] Examples that use numerical methods to solve the portfolio choice problem with housing and exogenous income include Cocco (2004), Yao and Zhang (2005), and Van Hemert (2009). Because these models focus on numerical solutions to single agent problems, they allow for general income processes and utility function, as well as own/rent decisions. Kraft and Munk (2010) solve the optimal portfolio choice in closed-form but they assume that all individuals are renters and markets are complete.

models of Rosen (1979) and Roback (1982) where spatial equilibrium is found by equalizing utility differences across cities. The more modern version of these models is the static horizontal sorting model of Bayer, McMillan, and Rueben (2005) or Bayer and Timmins (2005). Under a static setting, these models prove existence and uniqueness of a spatial equilibrium under very general assumptions about household heterogeneity. I make direct use of these horizontal sorting results in my model.[6]

Merging closed-form results from the optimal portfolio choice problem with a horizontal sorting model of individuals across space, I prove the existence and uniqueness of a linear stationary equilibrium. One key theoretical result is that home prices are derived to be a closed-form function of the underlying productivity of the economic base of a city minus a city-specific premium, which is a function of agent heterogeneity, sorting and risks in the particular city. Home prices can also be interpreted as the expected discounted sum of future dividends of the marginal household in that city plus a risk premium. In equilibrium, risk premia of homes takes a linear factor structure where the tradable part of risk is priced at the market price of risk. The non-insurable part of the local risk is also priced in equilibrium but its price is the risk-aversion parameter of the marginal person who lives there. In general, cities exposed to higher amounts of non-insurable risks have lower home prices in order to compensate residents for the extra risk they are taking.

Financial asset portfolio decisions are also found to be a generalized version of classic results in portfolio choice in finance. Stock holdings consist of a classic myopic term as found in Merton (1969), and a hedging demand term that depends on the correlation of the particular stock with income in the city a household lives in. The equilibrium rate of return for a particular stock also depends on its correlation with

[6] In terms of dynamic settings, the closest models are Glaeser and Gyourko (2010) and Van Nieuwerburgh and Weill (2010). These models are designed to study cross-sectional and time-series properties of house prices and wages. This paper differs from these models as it includes risk-aversion, portfolio choice and asset markets in the typical location choice equilibrium model.

income in all of the cities and the distribution of the risk-aversion parameter over the population.[7] Stocks that are negatively correlated with income and house prices are in large demand for hedging purposes. In order for the stock market to clear the returns for these stocks are lower in equilibrium.

An interesting insight to emerge is that households sort not only on income and amenities, but also on the uninsurable local risk. Households that are more risk-averse are more likely to locate in cities with less uninsurable risk. Sorting on risk leads in turn to missallocation of human capital. Risk-averse households do not move to the labor market where they are most productive because they may find that city too risky. Instead, some other less risk-averse and less productive household will move to the risky city in question. The creation of financial instruments that can be used as a hedge against noninsurable local risk can therefore reduce the incentives to sort on risk and increase productivity and welfare in the economy. In the limiting case where all risk is tradable, households do not sort on risk and they locate in the labor market where their productivity is highest.

I estimate a version of the model using wage and house price data for 216 metropolitan areas in the US under the assumption that agents are homogeneous.[8] Instead of looking for all the underlying factors that drive income and house prices in the US I focus on three factors that capture a large share of the common variation across metropolitan areas. These factors are constructed in the spirit of the three

[7] Other studies have also found that the presence of nontradable income and housing wealth in the household's portfolio can affect asset prices. Examples that study the impact of housing decisions and prices on financial asset prices include OMP, Piazzesi, Schneider, and Tuzel (2007), Lustig and van Nieuwerburgh (2005), and Yogo (2006). The interaction of labor income risk and asset prices is studied in Constantinides, Donaldson, and Mehra (2002), Santos and Veronesi (2006), and Storesletten, Telmer, and Yaron (2007).

[8] The model could be fully estimated using individual migrations decisions data available from the Decennial Census or the PSID. Using individual migration decisions allows for direct estimation of heterogeneity in preferences and in productivity. In a companion paper, I estimate the full structural model using a two-stage procedure closely related to the popular random coefficient logit model of Berry, Levinsohn and Pakes (1995).

Fama-French factors. The first factor (HMKT) is the average price growth across metropolitan areas. The second (SMBH) is price growth in high-price metropolitan areas minus growth in low-price ones. The third (HMLH) is growth in metropolitan areas with high price-to-wage ratios minus growth in ones with low price-to-wage ratios. The national HMKT factor turns out to have a correlation of .6 with aggregate REIT returns, while the other two factors are uncorrelated with REITs or the three Fama-French factors. This suggests that the national factor may be spanned by traded assets while the other two most likely are not. I estimate the model under different assumptions about the factors' tradability and I find that areas with higher nontradable variance demand a house price risk premium, which is reflected in lower house prices. The estimated risk premia imply that on average homes are about $20000 cheaper than they would be if owners were risk-neutral, although these estimates vary from $135000 for San Francisco, CA to -$20000 for Denver, CO. To my knowledge no other estimates of housing risk premia exist in the previous literature.

Using the estimates for 216 US metropolitan areas, I then simulate the model to study the effect of financial innovation on house prices, household sorting across space and on overall productivity and welfare in the economy. I start from the baseline where only the national HMKT factor is spanned by traded stocks. I then consider the cases where new financial instruments are created that correlate with the two other factors proposed above, and the case when the financial instrument allows for perfect insurance. The creation of tradable financial instruments that correlate with housing and income risk improves households' ability to hedge risk and consequently lowers housing risk premia. Due to heterogeneity in risk preferences, this leads to a different sorting of households across space. In the new sorting equilibrium, human capital is allocated more efficiently, leading to higher overall productivity welfare.

For a reasonable range of risk-aversion parameters, I find that completing markets can increase home prices by about 20 percent, increase productivity in the economy by 10 percent, and significantly improve welfare.[9] The average willingness to pay for access to financial instruments that correlate with all of the sources of risk in the economy is between \$10000 and \$20000 per homeowner.[10] The willingness to pay is, however, much higher for households that are very risk-averse, or households who live in locations with high noninsurable volatility. Overall, these findings do depend on the distribution of the risk-aversion parameter over the population. When heterogeneity in risk-aversion is large, productivity increases, but prices and welfare respond less to completing markets because of sorting effects. When households are homogeneous, all of the benefits of completing markets are directly reflected into higher house prices, and there is no gains in productivity.

The rest of the paper is organized as follows: Section 2 develops a joint equilibrium theory of location and portfolio choice. In Section 3, I describe the estimation procedure and fit the model to annual wage and house price data for 216 metropolitan areas in the US for the last 25 years. Section 4, uses the model to conduct a series of general equilibrium counter-factual simulations designed to study the impact of creating new financial instruments that allow individuals to hedge housing risk; and Section 5 concludes.

1.2 The Model

The classical Capital Asset Pricing Model (CAPM) states that only variation related to the market factor should be priced in equilibrium. In contradiction with CAPM,

[9] Under the popular "standard incomplete markets" framework in macroeconomics, Heathcote, Storesletten and Violante (2008) also find that insuring wage risk has large welfare and productivity implications. The gain in productivity is through a different channel in their model. Under market incompleteness the less productive workers work too much while high productivity agents work too little. Similar results are also found in Pijoan-Mas (2006). My paper proposes a different channel through which market completeness affects productivity. Completing markets increases productivity because workers are matched with jobs more efficiently because of sorting effects.

[10] To my knowledge there exist no previous estimates in the literature for the willingness to pay for access to these new financial instruments.

Hizmo (2010) and Case et al. (2010) find that factors not related to market returns and idiosyncratic risks are priced in housing market returns. The likely reason why CAPM fails to describe housing returns fully is that housing markets are characterized by many frictions that limit arbitrage. Because houses are large and indivisible, most households only own a home in one city, which exposes them to a large amount of local risk. This, combined with households' inability to perfectly hedge income and house price risks, gives rise to a unique problem that typical asset pricing models are not well-suited to tackle.

This paper develops a micro-founded dynamic equilibrium model that accounts for all of the frictions mentioned above and fits the deviations from the CAPM. This flexible model simultaneously considers risk-aversion, multiple sources of uncertainty, rich household heterogeneity, sorting, portfolio choice and asset prices in one unified framework. In equilibrium, risk premia of homes turns out to take a linear factor structure where the non-tradable part of the local risk is priced. Interestingly, equilibrium financial asset returns are also affected by the frictions in housing markets. Because of the explicit nature of the solution, the main equations of the model can be directly estimated using house price and income data. The model is also well suited for counter-factual simulations since the derived equilibrium is unique for any given set of parameters. In the rest of this section, I describe the setting of the model, prove existence and uniqueness of the equilibrium, and discuss the theoretical results.

1.2.1 *The Setting*

Cities and population

There is a measure one of households born at time t and they live until $t + \overline{T}$.[11] I consider an overlapping generations model where a new cohort of people is born in every instant. Households must buy only one house to live either in the countryside or in any city. They choose a location after they are born and live there until the end of their lives, at which point they sell their house to the new incoming generation. Such transactions occur continuously since at every instant in time a generation is born and a generation dies.

Suppose there are L cities denoted $l = 1, ..., L$ with countryside $l = 0$. Each city l has n^l houses available for people to move into at any given period of time and cities are fixed in size. The total supply of a city is $\int_0^{\overline{T}} n^l ds = \overline{T} n^l$. Assume that one household must occupy exactly one house. Not all of the households from the same cohort can locate in cities since housing is scarce there. This means that a share of the population must locate in the countryside. The reason why the countryside is important here is that it serves as the outside choice that can be used the set the level of utility and prices in the spatial equilibrium. This will become obvious when we discuss uniqueness of the equilibrium.

Financial asset market

Suppose that the whole economy is driven by m independent Brownian motions $\mathbb{B}_t = (B_t^1 ... B_t^m)'$. Also suppose that people can invest in n risky financial assets and

[11] Allowing for random life spans would complicate the model since households would want to hedge their mortality risk on top of their income and house price risk. We abstract from these complications in the current setting since our main goal is to understand the effects of wage and house price risk.

one risk-less asset.[12] The risky assets do not pay any dividends.[13] Households can hold any amount of these assets and there are no transaction costs or other frictions. The evolution of the risky asset prices is given by:

$$dP_t^i/P_t^i \quad = \quad \mu_t^i dt + \sigma_{i1} dB_t^i + \sigma_{i2} dB_t^i + ... + \sigma_{im} dB_t^i \qquad i = 1, 2, ...n$$

For compactness, we write this in matrix form as:

$$D_{P_t}^{-1} dP_t = \mu_t dt + \Sigma_P d\mathbb{B}_t$$

where $P_t = [P_t^1, ..., P_t^m]$, $D_{P_t}$is a $n \times n$ matrix with diagonal equal to P_t, $\mu_t = [\mu_t^1, ..., \mu_t^m]$, Σ_P is the $n \times m$ matrix whose rows are the volatilities of P_t^i.

Although we could potentially deal with a more general case, for simplicity we assume that the coefficients $\mu_t^i = \mu^i$ and $\sigma_{ijt} = \sigma_{ij}$ are constant over time so we have:

$$D_{P_t}^{-1} dP_t = \mu dt + \Sigma_P d\mathbb{B}_t$$

The matrix Σ_P, which captures the exposure of stocks to the underlying risk factors $\mathbb{B}_t$, is given exogenously. In contrast, the expected return vector μ is an equilibrium object. Given the volatility matrix Σ_P, I solve for the equilibrium rate of return vector μ that clears the financial asset market.

For now, we do not consider other kinds of assets or derivatives in this economy. The inclusion and pricing of other sorts of assets or derivatives could also be easily handed in this framework because of the standard normality assumptions that we are making as to how shocks evolve in the economy. Therefore, the set of stock used here can be thought of as the basis that spans all the other financial assets in the economy that we do not consider.

[12] Notice that depending on m and n the stock market could be complete or incomplete. We will have $m > n$ in our model which will imply that markets are not complete.

[13] The assumption that stocks do not pay any dividends is made for analytical simplicity and is standard in the finance literature. For the purpose of our model it doesn't really matter if returns from the financial assets are coming from dividends or from price appreciation. All we are interested in is what returns are and how they are correlated with other assets in the economy.

The housing dividend

Households must buy only one house in order to live either in the countryside or in any city. As mentioned, they decide where to live when they are born and live in the same place until the end of their lives. The countryside does not pay any wages and does not offer any amenities.[14] When a household moves to a city, it gets utility from the local amenities as well as wages from local firms. The per period housing dividend a household living in city l receives in terms of dollars is:

$$D_{it}^{l} = w_{it}^{l} + \beta_i M^{l}$$

where w_{it} is the wage household i receives at time t and $\beta_i M^l$ is the value it receives from amenities in city l.[15] The taste parameter β_i capture heterogeneity in the population of preferences for city-specific amenities, such as school quality, crime or weather. The wages household i receives are given by:

$$w_{it}^{l} = y_t^{l} + \xi_i + \varepsilon_i^{l}$$

where y_t^l is time-varying city productivity, ξ_i is a worker-specific fixed effect that does not vary across cities and ε_i^l is a city-worker match fixed effect. We can think of y_t as the part of the city-specific productivity that does not depend on who works there. The fixed effect ξ_i can be thought of as the effect of an individual's general education or expertize on wages, which does not depend on where the household locates. The last term ε_i^l can be interpreted as the effect of industry or firm-specific human capital on wages. Workers specialized in the auto industry will have a higher ε_i^l in Detroit,

[14] This assumption is not as stringent as it may seem. House prices and utility in equilibrium will be relative to the outside choice i.e. the countryside. The utility and prices of the outside choice are set to zero for simplicity. In principle, the outside utility could be set to any level and this does not affect the equilibrium allocation of households or asset prices. Changing the level of the outside choice utility only increases the price level in every city by a constant. This is a standard property of standard discrete choice spatial models.

[15] Note that this should be interpreted as the net dividend from living in one house, which is the part of wages and amenities left over after households pay taxes and other costs to live in the city.

MI, and workers who specialize in the high tech industry will have a higher ε_i^l in San Jose, CA.

The only part of wages that is stochastic from the household's point of view is y_t. Suppose that the productivity of city l at time t denoted by y_t^l evolves according to:[16]

$$
\begin{aligned}
dy_t^l &= s_t^l dt \\
ds_t^l &= \phi^l \left(m^l - s_t^l\right) dt + \Sigma_s^l d\mathbb{B}_t
\end{aligned}
$$

This process means that the city-specific productivity y_t grows with a stochastic drift s_t. This drift follows a mean-reverting stochastic process also known as the Ornstein-Uhlenbeck process, which is the continuous-time analog of a discrete time AR(1) process. The stochastic shocks that govern this process are the underlying sources of risk in the economy $\mathbb{B}_t$. The vector Σ_s^l, which is $m \times m$, governs the exposure of city l to these underlying sources of risk. The city-specific parameter m^l is the long-run average of the drift s_t, while the parameter ϕ^l governs the speed at which the process reverts to the mean m^l. High values of ϕ^l imply that the process reverts to the long run mean quickly, and low values imply that the process is more likely to wander off far from the mean for extended periods of time. This process implies that income will on average increase by the average amount m^l, but there will be "business cycles" where income increases faster or slower than the long run average m^l.

The utility specification

A household is born at some time t, buys a house in a city, starts working and receives wages, and continuously invests his wealth in financial assets. At the end of

[16] This exact form of the evolution of the city-specific productivity is not essential for the model. The choice of this particular stochastic process is motivated by empirical patterns of the metropolitan area-specific average income series, which are used to estimate of the model in the next section.

his lifetime, the household sells his home and all of his assets and consumes all of his accumulated wealth. It is assumed that there is no intermediate consumption; all of the wealth is consumed at the end of his life. Allowing for intermediate consumption complicates the solution of our problem to the point where analytical solutions are not possible. In fact, there are no known analytical solutions in the literature to the portfolio choice problem with exogenous income and incomplete markets, unless the income process is extremely simplistic.[17]

At birth time t_0, households maximize their lifetime utility given by:

$$V\left(X_{t_0}, w_{t_0}, s_{t_0}, l\right) = \sup_{\theta_t, l} - E_{t_0} e^{-\gamma_i \left[X_{t_0+\overline{T}} + M^l \beta_i \frac{1}{r}\left(e^{r\overline{T}}-1\right)\right]} \tag{1.1}$$

where $X_{t_0+\overline{T}}$ is the financial wealth accumulated through the lifetime and the second term is the utility accumulated from access to the local amenities. Each household chooses the optimal location l in which to live and each period chooses the optimal θ, which is the dollar amount invested in stocks. The above maximization problem is constrained by the wealth evolution equation, which for $t \in (t_0, t_0 + \overline{T})$ is:

$$dX_t = \theta_{it} \frac{dP_t}{P_t} + r\left(X_{it} - \theta_{it}\right) dt + w_{it} dt$$

where X_t is total wealth and θ_t is amount of money invested in stocks. Suppose an individual will die at time T, which means he is born at $t_0 = T - \overline{T}$. Using the constraint, we can solve for the terminal wealth and write the optimization problem for any $t \in (t_0, T)$ as:

$$V_t^l = \sup_{\theta_t, l} - E_t e^{\bar{U}} \tag{1.2}$$

[17] Merton (1971) solves this problem by assuming deterministic income. Svensson and Werner (1993) provide solutions in the case of infinitely lived agents with intermediate consumption in the case where income is drawn from an iid normal distribution, meaning that all of the shocks to income are temporary. See Henderson (2005) for a more detailed discussion.

where:

$$\bar{U} = e^{r(T-t)}X_t + \int_t^T e^{r(T-u)} \left(\theta'_u D_{P_u}^{-1} dP_u - 1'\theta_u r du + w_u du\right) + \frac{1}{r}\left(e^{r(T-t)} - 1\right)\beta_i M^l + p_T^l$$

Households therefore choose a city l to work and live in, and a series of stock holdings θ_t that maximize their lifetime utility, which is defined over their terminal wealth. This wealth is composed of their captial gains on the home they bought, the wealth accumulated from wages they received, the wealth accumulated from investing in the stock market, and the dollar value of the utility accumulated from having access to local amenities.

1.2.2 Equilibrium

An equilibrium in this economy is a set of home prices, asset prices and portfolio and location decisions such that households maximize utility and the asset markets and the housing markets clear. I consider a stationary equilibrium where the mass of houses available to a new cohort n^l is constant over time.[18] This means that there are n^l households of every possible age in one city.

Because of the complexity of the problem considered, I focus only on equilibrium prices that are linear in the states:

$$p_t^l = A^l y_t^l + B^l s_t^l + C^l \tag{1.3}$$

where A^l, B^l, and C^l are city-specific constants to be determined in equilibrium. Starting with prices of this form, I show that there exists a unique equilibrium and that the values for these three vectors $\mathbf{A} = [A^1, A^2, ..., A^L]$, $\mathbf{B} = [B^1, B^2, ..., B^L]$

[18] The assumption of fixed supply is not empirically realistic for many metropolitan areas, although it may fit the most volatile and supply constrained coastal metropolitan areas. Allowing for elastic supply is first on the list of many extentions to the current model.

and $\mathbf{C} = [C^1, C^2, ..., C^L]$ are determined uniquely in equilibrium. The price for the countryside is normalized to zero.[19]

Several steps are needed to prove that there exists a unique equilibrium . First, conditional on a household having located in a specific city, I solve for optimal value function and for optimal portfolio choice. Next, I show that given vectors $\mathbf{A}$ and $\mathbf{B}$, the equilibrium stock returns are uniquely determined by asset market clearing. Then, I turn to the location decision which gives unique values for $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ and a unique sorting equilibrium. These vectors $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ ensure equilibrium both in the housing and the asset market.

Portfolio choice and stock market equilibrium

Conditional on living in some location l, each household continuously chooses where to allocate his money. The problem given in equation (1.2), conditional on a choice of l, is the optimal portfolio choice problem for an investor facing imperfectly hedgeable stochastic income. The returns on income and stocks are imperfectly correlated, so the market is incomplete. This is a complicated class of problems to solve and very few instances in the finance literature have been successful in finding a closed-form solution.[20]

[19] The assumption on the linearity of equilibrium prices is not as limiting as it seems. The equilibrium prices found from this assumption turn out to be the expected discounted sum of future dividends of the marginal person that lives in the city plus a city-specific risk premium as shown later in a proposition.

[20] Henderson (2005) is the closest example to this paper in providing closed-form results. Under the special case of iid and normally distributed income, Svensson and Werner (1993) obtain explicit results in an infinite horizon problem. Similar results are also found by Duffie and Jackson (1990) and Tepla (2000) in a finite horizon problem. Other cases that study the portfolio choice problem with exogenous income either solve the model numerically or make the assumption that markets are complete. Examples that use numerical methods include Cocco (2004), Yao and Zhang (2005), and Van Hemert (2009). Kraft and Munk (2010) solve the optimal portfolio choice in closed-form but they assume that markets are complete.

To find the value function, I follow the logic laid out by Henderson (2005).[21] First, I take the first order conditions of the Hamilton-Jacobi-Bellman (HJB hereafter) and substitute out θ_t and X_t. This gives rise to a nonlinear partial differential equation we need to solve. The nonlinear HJB equation is then transformed to a linear partial differential equation. At this point, we do a change of measure and use the Feynman-Kac theorem to get the value function with the portfolio choice substituted out. Taking the expectation, we achieve the result in the following proposition, which is proved in the appendix.

Proposition 1. (Value Function) *The value function of an household at some time $t \in (t_0, T)$ conditional on being in city l is:*

$$V(t, y, s, l) \quad = \quad -e^{U_t^l} \tag{1.4}$$

with:

$$\begin{aligned} U_t^l \quad = \quad & -\gamma_i \left[e^{r(T-t)} X_t + \frac{1}{r} \left(e^{r(T-t)} - 1 \right) \left(\beta_i M^l + \xi_i + \varepsilon_i^l \right) \right] \\ & -\gamma_i \left[\hat{k}_{t1}^l y_t^l + \hat{k}_{t2}^l s_t^l + \hat{k}_{t3}^l + \int_t^T \hat{k}_4 \hat{k}_{u2}^l \sqrt{\Sigma_{ss}^l} du \right] \\ & + \frac{1}{2} \gamma_i^2 \left(\Sigma_{ss}^l - \Sigma_{sP}^l \Sigma_{PP}^{-1} \Sigma_{Ps}^l \right) \int_t^T \left(\hat{k}_{u2}^l \right)^2 du - \frac{1}{2} \left(\mu_t' - 1'r \right) \Sigma_{PP}^{-1} \left(\mu_t - 1r \right) (T-t) \end{aligned}$$

where $\hat{k}_{t1}^l ... \hat{k}_4$ are defined as

$$\hat{k}_{t1}^l = A^l - \frac{1 - e^{r(T-t)}}{r}$$

[21] The problem is different from that in Henderson (2005) since here risks and stocks are multidimensional, the income process is governed by a state variable that does not depend on the level of income, and agents hold a risky house in their portfolio.

$$\hat{k}^l_{t2} = \left\{ A^l \frac{1-e^{-\phi^l(T-t)}}{\phi^l} + B^l e^{-\phi^l(T-t)} - \frac{1-e^{r(T-t)}}{\phi^l r} + \frac{e^{-\phi^l(T-t)} - e^{r(T-t)}}{\phi^l\left(r+\phi^l\right)} \right\}$$

$$\hat{k}^l_{t3} = m^l \left\{ -\hat{k}^l_{t2} + A^l\left(T-t\right) + B^l + \frac{t\left(1-e^{r(T-t)}\right)}{r} - \frac{1}{r^2}\left(rT + 1 - e^{r(T-t)}\left(rt+1\right)\right) \right\} + C^l$$

$$\hat{k}^l_4 = -\frac{\Sigma^l_{sP}\Sigma^{-1}_{PP}\left(\mu - 1r\right)}{\sqrt{\Sigma^l_{ss}}}$$

and A^l, B^l, C^l, m^l,ϕ^l,ε^l_i and Σ^l_s are city-specific constants and $\Sigma_{KL} = \Sigma_K \Sigma'_L$.

Notice that the value function above is the exponential of a linear function in states. All of the functions $\hat{k}^l_{1t}...\hat{k}^l_4$ do not depend on the states or individual heterogeneity parameters. This simple structure is very helpful in solving the portfolio and location decisions given in the next proposition and proven in the appendix.

Proposition 2. (Portfolio Choice) *A household of type i living in city l at time $t \in (t_0, T)$ holds stocks in the dollar amount given by the vector:*

$$\theta_{ilt} = \left(\Sigma_{PP}\right)^{-1}\left[\frac{\left(\mu - \mathbf{1}r\right)}{\gamma_i e^{r(T-t)}} + \frac{\hat{k}^l_{t2}}{e^{r(T-t)}}\Sigma^l_{Ps}\right] \tag{1.5}$$

where $\hat{k}^l_{t2}$ is a city-specific nonrandom function of time that depends on vectors **A** *and* **B**.

The optimal portfolio in the stock, θ_{ilt} given in (1.5) is comprised of two components. The first is the same term that Merton (1969) finds as an optimal strategy in the absence of stochastic income. This strategy is myopic as it is the portfolio choice for an investor who ignores income risk and only looks one period ahead. Since income is risky and is correlated with stock returns, the optimal portfolio also includes a hedging component. This second term hedges the change of stochastic income and can be interpreted as the inter-temporal hedging demand as in Merton (1971).[22] Notice that the second term does not depend on the risk-aversion parameter. This is because income is normally distributed and it is instantaneously riskless, which means that investors do not hedge income but the change in income s_t.

Now we turn to equilibrium in the stock market conditional on values of **A** and **B**. There is a fixed supply of measure one of stocks and in equilibrium this should be equal to the aggregate demand for stocks. The aggregate demand for stocks of every household type i at age t in city l should equal the total stock supply:

$$\sum_{l=1}^{L}\left\{\int_{\gamma_i:V^l\geqslant V^k,\forall k}\int_0^T \Sigma_{PP}^{-1}\left[\frac{(\mu-\mathbf{1}r)}{\gamma_i e^{r(T-t)}}+\frac{\hat{k}_{t2}^l}{e^{r(T-t)}}\Sigma_{Ps}^l\right]dtd\Gamma(\gamma_i)\right\}=1$$

where $\Gamma(\gamma_i)$ is the cumulative density function for γ_i and n^l is the number of homes available in city l. Simplifying and solving for μ, we get the equilibrium market returns.

Proposition 3. (Stock Market Equilibrium) *Given a set vectors* **A** *and* **B** *the stock returns are uniquely determined in equilibrium*:

$$\mu=\frac{r\Sigma_{PP}+r\sum_{l=1}^{L}\left(n^l\Sigma_{Ps}^l\int_0^T\frac{\hat{k}_{2t}^l}{e^{r(T-t)}}dt\right)}{\left(1-e^{-r(T-t)}\right)\int_{\gamma_i}\frac{1}{\gamma_i}d\Gamma(\gamma_i)}+1r \tag{1.6}$$

where $\hat{k}_{2t}^l$ *is a city-specific constant that depends on vectors* **A** *and* **B** *and the age* t *of the individual.*

The equilibrium rates of returns for stocks in this case do depend on the distribution of the risk-aversion parameter over the population but do not depend on the

[22] A similar hedging demand is also found in Henderson (2005), Duffie and Jackson (1990) and Tepla (2000).

location decisions of individuals.[23] Equilibrium returns do not only depend on the variance-covariance matrix of the stock but also on the covariance of that particular stock with the local shocks in each city and the size of those cities. This finding is a deviation from standard equilibrium models where local nondiversifiable risk is ignored. If stocks are positively correlated with local shocks, they will have a higher rate of return than otherwise in order for the market to clear. Further interpretation of this result is postponed until after we solve for the equilibrium value of $\hat{k}^l_{2t}$, which makes this equation much more transparent.

Housing market equilibrium

We turn now to equilibrium in housing markets. At birth time t_0, households maximize the value function in equation (1.4) with initial wealth $X^l_{t_0} = X_{0i} - p^l_{t_0} = X_{0i} - A^l y^l_{t_0} - B^l s^l_{t_0} - C^l$. Given their optimal portfolio decisions given by Proposition 2, at time t_0 they choose the optimal location l that maximizes their lifetime utility in equation (1.4). Any monotonic transformations of the value function do not change the maximization problem a household faces. After taking the log of the value function and dropping any terms that are the same across choices, the maximization problem for a household born at time t_0 the log value function takes a very simple form.

Proposition 4. (Optimal Location) *A household i born at time t_0 chooses his/her optimal location by solving for the maximum log value function:*

$$max_l U^l_{t_0 i} = \beta_i M^l + k^l_1 y^l_{t_0} + k^l_2 s^l_{t_0} + k^l_3 - rC^l - \gamma_i k^l_4 + \varepsilon^l_i \tag{1.7}$$

[23] The reason that equilibrium rates of return do not depend on the location decisions of individuals is that, once the move to a city, all individuals face the same exposure to the city's risk factors given by Σ^l_s. If I were to allow the variance of wages to be individual specific then the location decisions of individuals would affect asset prices directly. I have solved the model in that case but do not present it here since allowing for individual heterogeneity in exposure to factors makes the solution much clumsier and harder to interpret.

where $k_1^l ... k_4^l$ are location-specific constants that do not depend on states, individual heterogeneity, or the vector $\mathbf{C}$*:*

$$k_1^l = 1 - rA^l$$

$$\begin{aligned} k_2^l &= \frac{r}{\left(e^{r\overline{T}} - 1\right)} \left\{ A^l \frac{1 - e^{-\phi^l \overline{T}}}{\phi^l} + B^l \left(e^{-\phi^l \overline{T}} - e^{r\overline{T}} \right) \right. \\ & \left. - \frac{1 - e^{r\overline{T}}}{\phi^l r} + \frac{e^{-\phi^l \overline{T}} - e^{r\overline{T}}}{\phi^l \left(r + \phi^l\right)} \right\} \end{aligned}$$

$$\begin{aligned} k_3^l &= \frac{r}{\left(e^{r\overline{T}} - 1\right)} \left[A^l m^l \overline{T} - m^l \frac{1}{r^2} \left(r\overline{T} + 1 - e^{r\overline{T}} \right) \right] \\ & -m^l k_2^l - \Sigma_{sP} \Sigma_{PP}^{-1} \left(\mu - 1r\right) \int_0^{\overline{T}} k_5^l du \end{aligned}$$

$$k_4 = \frac{1}{2} \left(\Sigma_{ss}^l - \Sigma_{sP}^l \Sigma_{PP}^{-1} \Sigma_{Ps}^l \right) \int_0^{\overline{T}} \left(k_{5u}^l + Be^{ru} \right)^2 du$$

$$\begin{aligned} k_{5u}^l &= \frac{r}{\left(e^{ru} - 1\right)} \left\{ A^l \frac{1 - e^{-\phi^l u}}{\phi^l} + B^l \left(e^{-\phi^l u} - e^{ru} \right) \right. \\ & \left. - \frac{1 - e^{ru}}{\phi^l r} + \frac{e^{-\phi^l u} - e^{ru}}{\phi^l \left(r + \phi^l\right)} \right\} \end{aligned}$$

The constant $k_1^l ... k_4^l$ are found by setting $T - t_0 = \overline{T}$ in the constants$\hat{k}_{1t}^l ... \hat{k}_4^l$ in Proposition 1. The $\hat{k}_{it}$ in Proposition 1 depend on time t because they are functions of the number of years left until death $T - t$. Since at birth there are $\overline{T}$ years left until death, the constants $k_1^l ... k_4^l$ do not depend on time but only on total lifespan $\overline{T}$.

Given the maximization problem (1.7), household i chooses location l if the utility it receives from this choice exceeds the utility it receives from all other choices, or when:

$$U_i^l \geqslant U_i^k \quad \Rightarrow \quad \bar{U}_i^l + \varepsilon_i^l \geqslant \bar{U}_i^k + \varepsilon_i^k \quad \Rightarrow \quad \varepsilon_i^l - \varepsilon_i^k \geqslant \bar{U}_i^k - \bar{U}_i^l \qquad \forall l, k$$

where $\bar{U}_i^l = U_i^l - \varepsilon_i^l$. The location-worker specific fixed effect ε_i^l is drawn from some probability distribution for each household. Because we have a continuum of households, every possible value of ε_i^l is drawn in each generation. Therefore we can interpret the mass of individuals with ε_i^l in some given range as the probability of drawing ε_i^l from that range. If we interpret ε_i^l as a random variable, the probability that a household chooses location l can be written as a function of everything that enters in U_i^l:[24]

$$Pr_i^l = f_l\left(\beta_i, \gamma_i, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{K}\right) \tag{1.8}$$

Aggregating these probabilities in equation (1.8) for all households gives the demand for each city:

$$D_t^l = \int_{\beta_i, \gamma_i} f_l\left(\beta_i, \gamma_i, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{K}\right) dF(\beta_i, \gamma_i)$$

where $F(\beta_i, \gamma_i)$ is the joint distribution of the β_i and γ_i parameters over the population. In equilibrium, demand for each city has to equal supply in each city:

$$D_t^l = n^l \quad \forall l \quad \Rightarrow \int_{\beta_i, \gamma_i} f_l\left(\beta_i, \gamma_i, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{K}\right) dF(\beta_i, \gamma_i) = n^l, \quad \forall l \tag{1.9}$$

The sorting problem described here is very similar to horizontal sorting models in urban economics. The closest models in structure are those developed by Bayer, McMillan and Rueben (2005) (BMR) and Bayer and Timmins (2005). At birth, which is when households choose where to locate, $k_1^l ... k_4^l$ are city-specific characteristics. This means that the current framework maps directly to the BMR sorting

[24] Again the probability that a household of type i chooses location l is the same as the mass of households of type i that choose location l. I interpret this as a probability only to relate these results to the urban horizontal sorting models as shown below.

model with the vector **C** serving as a price that clears the market. If demand exceeds supply, C^l is bid up until demand equals supply, and vice versa. Just as in the BMR model, given fixed city characteristics, there is a unique "price" vector **C** that leads to a unique sorting equilibrium.

Proposition 5. (Sorting) *Given a set of vectors* **A** *and* **B** *and a set of city characteristics* $k_1^l ... k_4^l$, *if* ε_i^l *is drawn from a continuous distribution, a unique vector* **C** *solves the system of equations given in (1.9). Moreover, the equilibrium spatial allocation of households is unique (within the class of linear stationary equilibria).*

The first part of the Proposition 5 is proved in Proposition 1 of BMR and the second part in Proposition 2 of Bayer and Timmins (2005). Their proofs will not be reproduced here.

Now we turn to determine the unique values of the vectors **A** and **B**. Note that the distribution $F(\beta_i, \gamma_i, \varepsilon_i)$ is not time-dependent. The quantity of available houses in each neighbourhood n^l is also constant across time by assumption. That means that the same set of heterogeneous individuals will be drawn each period and they will sort across the same set of homes. The sorting problem will be identical in each period except that the states y_t and s_t may change. Because **A** and **B** are constant across time, in order for the constant vector **C** to clear the markets each period, we need the utility from living in a city not to depend on states y_t and s_t. If the utility function given in (1.7) depended on the states, the vectors **A**, **B** and **C** could not be constant. Setting $k_1^l = k_2^l = 0, \forall l$ gives the unique values for **A** and **B**. The result is given in the next proposition which is proved in the appendix.

Proposition 6. (Equilibrium Prices) *Given a time invariant distribution of household characteristics* $F(\beta_i, \gamma_i, \varepsilon_i)$, *the equilibrium house prices are given by:*

$$p_t^l = \frac{1}{r} y_t^l + \frac{1}{r\left(r + \phi^l\right)} s_t^l + \frac{\phi^l m^l}{r^2\left(r + \phi^l\right)} + \pi^l \tag{1.10}$$

where $\pi^l = C^l - \frac{\phi^l m^l}{r^2(r+\phi^l)}$. *This price function gives a unique sorting equilibrium in housing markets as shown in Proposition 5 and gives unique equilibrium rates of return for stocks as shown in Proposition 3.*

Proposition 6 shows that the price function given in (1.10) ensures market clearing in both the stock and the housing market. Not only is this equilibrium price unique among the 'linear-in-the-states' class of functions, but it also results in a unique equilibrium in the housing and asset markets.

1.2.3 Analysis of the Equilibrium

Equilibrium house prices

The linear structure of equilibrium house prices is not as limiting as it may first seem. The price function given in (1.10) has also the appealing feature that it can be expressed as the expected discounted sum of productivity in one city plus a city-specific constant, π^l, which captures risk premia and heterogeneity in preferences in each city.

Proposition 7. *Under the assumed processes for the evolution of* y_t *and* s_t*, the equilibrium price function (1.10) is equivalent to the price function:*

$$p_t = E_t\left(\int_t^{\infty} e^{-r(u-t)} y_t^l dt\right) + \pi^l$$

where π^l *is a city-specific constant determined in equilibrium that captures housing risk premia and the effect of individual heterogeneity in prices.*

In order to prove the proposition above, I plug in the assumed process for y_t^l and solve the double integral that discounts the cash flows to the present and takes the expectation. Given the properties of the y_t^l process, we can switch the order of integration and solve for the double infinite integral to get equation (1.10). The details of the proof can be found in the appendix.

Sorting and the factor structure for housing risk premia

The analysis so far has been conducted under a general form for Σ_P, which is the co-movement of stock returns in relation to the underlying sources of uncertainty. Without loss of generality, consider the case where stock i's returns are driven by only one underlying Brownian motion. Suppose there are n such stocks and that stock i is only driven by Brownian motion B_t^i. This means that all the off-diagonal elements of Σ_P are zero. There are a total m Brownian motions that drive the economy so if $m > n$, perfect hedging cannot be achieved. The reason for this transformation of the model is that it greatly simplifies the interpretation of the results, and it delivers equations that I can easily estimate using house price and wage data.

Proposition 8. *If the n stocks are driven by the first n sources of risk $B^1 \ldots B^n$ and all the off-diagonal elements of Σ_P are zeros, then the location-specific value function is:*

$$max_l U_i^l = \beta_i M^l - r\pi^l - \sum_{j=1}^{n} \lambda_j \sigma_s^{lj} \frac{1}{r\left(r+\phi^l\right)} - \gamma_i \sum_{j=n+1}^{m} \left(\sigma_s^{lj}\right)^2 \frac{\left(e^{r\overline{T}}-1\right)}{4r^2\left(r+\phi^l\right)^2} + \varepsilon_i^l \quad (1.11)$$

where λ^j is the Sharpe ratio of stock j given in equilibrium by:

$$\lambda^j = \frac{\mu^j - r}{\sigma_P^j} == \frac{r\sigma_P^j - \overline{T}\sum_{l=1}^{L}\left(\sigma_s^{lj}\frac{n^l}{(r+\phi^l)}\right)}{\left(1-e^{-r\overline{T}}\right)\int_{\gamma_i}\frac{1}{\gamma_i}d\Gamma(\gamma_i)} \quad (1.12)$$

and σ_P^j and σ_s^{lj} are the (j,j)th and jth element of Σ_p and Σ_s^l respectively.

Proposition 8 can be derived by plugging the equilibrium house price function in equations (1.6) and (1.7). First note how exposure to a source of risk affects utility and in turn prices. A city can be exposed to an underlying source of risk that is spanned by traded stocks or to one that is not hedgeable by traded assets. If the underlying source of uncertainty j is spanned, the standard deviation σ_s^{lj} of the city's exposure to it enters utility in a linear form and is multiplied by the price of risk (Sharpe's ratio) for that source of uncertainty. Otherwise, the unspanned risk factors enter jointly in the term that is multiplied by the risk-aversion parameter γ_i. Notice that if all of the sources of risk were spanned, the latter term would drop out of the utility function and not affect prices in equilibrium. If this is not the case, however, idiosyncratic risk specific to a city that is not spanned by the traded assets will affect utility and it will be priced in equilibrium.

The equilibrium market prices of risk λ^j for stock j in this case depend not only on the interest rates and the volatility σ_P^j, but also on its correlation with the income and house price processes in all of the cities captured by σ_s^{lj}, which is weighted by the size of the city n^l. If stock j is positively correlated with local shocks for a large share of the homes in the economy, they will have a higher rate of return and a higher market price of risk than otherwise. This is because most people will want to short the stock in question for hedging purposes. For every person who shortsells the stock, there must be someone who holds it. Therefore, in order for it to be worth it for people to hold this stock, the rate of return and the market price of risk need to be higher. The market price of risk does not depend on the location decisions of individuals since σ_s^{lj} is not individual specific. If it was, the allocation of individuals across space would directly affect equilibrium asset returns. The distribution of the risk-aversion heterogeneity also affects the market prices of risk as can be seen from the integral in the denominator.

Talent allocation

In this equilibrium, individuals do not always choose the location where they are most productive or where their ε_i^l is the highest. As a result, productivity in the economy is not maximized and there are two reasons for this. Firstly, workers do not go where they are most productive because of amenities that different places possess.

A worker who is very productive in low amenity city A but not so productive in a high amenity city B may decide to locate in B if he places a lot of weight on amenities.

The same kind of productivity inefficiencies also arise when cities have different levels of nondiversifiable risk. An individual who is very risk-averse may not decide to locate in a city with a high level of nondiversifiable risk. If all of the sources of uncertainty were hedgeable, however, this kind of sorting inefficiency would not be present. In our counterfactual simulations in the section 4, we look at how much more productive the US would be as a result of creating assets that span presently hedgeable sources of risk.

1.3 Estimation

The equilibrium model developed in the previous section can be fully estimated by combining individual migration data from the decennial Census, housing price data from FHFA, wage data available from the Bureau of Economic Analysis, and amenity data such as crime, weather and quality of schooling from other sources. Instead of using migration data, I take a simple approach in this paper and estimate the model assuming homogeneity in preferences. I later use the estimated parameters of the model in a series of counterfactual simulations in which I allow for various degrees of preference heterogeneity.

I estimate the model using house price and wage data at the metropolitan area level for the US. Home prices are constructed by using the annual FHFA home price indices at the metropolitan area level, which are scaled up by using median home price levels from the 2000 Census. The wage data used consist of the personal income per capita measure made available by the Bureau of Economic Analysis. I also construct several proxies for amenities, which are: crime levels from the FBI's publication "Crime in the United States", population, population density, and house density from the decennial Census, and a series of weather variables from the Area Resource

File maintained by the Quality Resource Systems. The final sample analyzed here consists of annual data for the period from 1985 to 2008 for 216 metropolitan areas (MSAs hereafter).

Under the assumption of homogeneity in preferences for amenities and for risk, the price equation becomes linear in risk factors. The main estimating equation therefore is:

$$p_t^l = \frac{1}{r} y_t^l + \frac{1}{r\left(r + \phi^l\right)} s_t^l + \frac{\phi^l m^l}{r\left(r + \phi^l\right)} + \pi^l \tag{1.13}$$

where

$$\pi^l = \frac{1}{r} M - \sum_{i=1}^{n} \lambda_i \frac{\sigma_s^{il}}{r^2\left(r + \phi^l\right)} - \gamma \sum_{i=1}^{n} \frac{\left(\sigma_s^{il}\right)^2 \left(e^{rT} - 1\right)}{4r^3\left(r + \phi^l\right)} + \varepsilon \tag{1.14}$$

Empirically, p^l is the price level in any particular metropolitan area l, y_t^l is the net income received, s_t is the instantaneous growth in income, and m^l is the expected growth in income. The price premium π^l is composed of the effect of amenities M, the effect of the exposure to traded sources of risk, and the exposure to local noninsurable sources of risk. Our goal is to estimate all of the parameters of the price equation in order to get estimates for risk premia. Estimating the price equation is achieved in several steps. First, I use time-series regressions to estimate ϕ, as well as to identify common risk factors that drive the economy. Using these factors, I can estimate the factor loadings σ_s^{il} for each metropolitan area. In the second step, I use cross-sectional regressions to estimate equation (1.14) and get estimates of the market prices of risk λ_i and the risk-aversion parameter γ. Using these parameters, I then calculate the implied risk premia and amenities for each of the MSAs.

1.3.1 Time-Series Factor Decomposition

In order to estimate the model, we would first need estimates of the underlying factors that drive the economy in each metropolitan area, which are given by the

Brownian motions $\mathbb{B}_t$ in the model. Using price and wage data, we can only identify the total effect that these factors have on the prices and wages of any particular area. Breaking this total effect up into the actual factors would require a deeper analysis into the forces that drive the local industry of each metropolitan area, which is beyond the scope of this paper. Instead, here I estimate the total effects of these underlying factors for each metropolitan area and I decompose them into common empirical factors that are constructed in the spirit of the Fama-French factors for the stock market.

Total Effect of Factors

First, I estimate the total effect of the underlying risk factors on wages and prices. As in the model I assume that the growth in average income in each metropolitan area evolves according to:

$$\begin{aligned} dy_t^l &= s_t^l dt \\ ds_t^l &= \phi\left(m^l - s_t^l\right)dt + \Sigma_s^l d\mathbb{B}_t \end{aligned}$$

which means that the change in income dy_t^l has a mean reverting drift s_t^l. The instantaneous income drift s_t^l is driven by the exposure of the metropolitan area to the underlying factors $\mathbb{B}_t$. What we are after here is estimating the total effect of these factors on income as given by the term $\Sigma_s^l d\mathbb{B}_t$. This effect cannot be identified by using only income data unless income is observed continuously over time. Using equation (1.13), however, we can estimate the total effect by exploiting both wage and house price data even though they are observed over discrete time intervals.

The procedure to estimate the magnitudes of interest is straight-forward and follows a series of linear regressions. Note that rearranging the equation (1.13), we can define x_t as the observable part of our model:

$$p_t^l - \frac{1}{r}y_t^l = x_t^l = \frac{1}{r\left(r+\phi^l\right)}s_t^l + \frac{\phi^l m^l}{r^2\left(r+\phi^l\right)} + \pi^l$$

The variables on the left hand side are the observable annual prices and wages in city l. All of the variables in the right hand side are unobservable and are to be estimated. Since s_t^l is assumed to follow an Ornstein-Uhlenbeck process, we know what process the observable x_t^l should follow. The evolution of x_t^l can be described by the SDE:

$$dx_t = \phi(\bar{x} - x_t)\,dt + \frac{1}{r(r+\phi)}\Sigma_s^l d\mathbb{B}_t \tag{1.15}$$

$$= \phi(\bar{x} - x_t)\,dt + \frac{\kappa^l}{r(r+\phi)}dZ_t^l \tag{1.16}$$

where $\bar{x} = \frac{1}{r(r+\phi)}m - \frac{\phi m^l}{r^2(r+\phi)} - \pi^l$, $\kappa^l = \Sigma_{ss}$ and $dZ_t^l = \Sigma_s^l/\Sigma_{ss}^l dB_t$. Notice that in the second equality we collapsed the separate effects of every risk factor $\mathbb{B}_t$ on one single metropolitan specific factor Z_t^l, which is a one-dimensional Weiner process.

The problem now is to estimate the parameters and the residual of an Ornstein-Uhlenbeck process x_t, which is observed at discrete time periods. To do this, first note that the solution to the stochastic differential equation (1.15) is:

$$x_{t+1} = x_t e^{-\phi} + \bar{x}\left(1 - e^{-\phi}\right) + \frac{\kappa^l}{r(r+\phi)}\int_t^{t+1} e^{-\phi(t+1-u)}dZ_t$$

Since Z_t is a Weiner process, the integral above follows a normal distribution. We can estimate the above equation by running the following regressing for each MSA:

$$x_{t+1}^l = a^l + b^l x_t^l + u_t^l \tag{1.17}$$

We can estimate the ϕ and κ^l by setting the parameters:

$$\hat{\phi} = -ln(\hat{b})$$

$$\kappa^l = r(r+\phi)\sqrt{\frac{-2ln(\hat{b}^l)}{1-\left(\hat{b}^l\right)^2}var(\hat{u}^l)}$$

Notice also that we can get an estimate of $\int_t^{t+1} \kappa^l dZ_t^l$, which is the total sum of factors that drive the economy from t to $t+1$, from the residual of the regression above. The estimate is:

$$F_t^l = \left(\widehat{\int_t^{t+1} \Sigma_s^l d\mathbb{B}_t}\right) = \left(\widehat{\int_t^{t+1} \kappa^l dZ_t^l}\right) = \frac{r\,(r+\phi)\,\kappa e^{\phi}}{std(\hat{u_t}^l)}\hat{u_t}^l$$

where $std(\hat{u_t}^l)$ is the standard deviation of $\hat{u_t}^l$. We interpret F_t^l as the MSA-specific underlying driver of prices and wages.

Using annual house price and income data for 216 metropolitan areas in the US for the period of 1985-2008, I estimate the parameters ϕ, κ^l and a vector F_t^l for each metropolitan area. I deflate prices and wages by the CPI with base year 2008 in order to eliminate any effects of inflation. Although inflation risk can significantly affect lifetime wealth, I abstract from it in the model and estimate the equations of interest using real variables. Also, since the model abstracts from time-varying interest rates, I set interest rate to $r = 0.04$.

The first regression that I estimate separately for each metropolitan area is:

$$Hprice_t^l - \alpha Income_t^l = a^l + b^l \left(Hprice_{t-1} - \alpha Income_{t-1}^l\right) + u_t^l \tag{1.18}$$

where $Hprice^l$ is the median house price level in MSA l and $Income_t$ is the per capita income. In the model α would be equal $1/r$ if all of the income that an individual receives can be saved or consumed. In other words the model states that the coefficient on *net* income should be 1/r. In reality only a portion of the labor income can be considered as net benefit since individuals pay taxes and also incur utility losses from sacrificing leisure in order to receive that income. Estimating the parameter α directly can be difficult since wages are correlated with unobserved amenities in each MSA leading to serious endogeneity problems. Because nicer places generally offer higher wages the estimated parameter will be biased upwards. For my sample, the estimated parameter is 3.7, which is very high, implying very negative amenities for many metropolitan areas that have relatively low price to wage ratios. In this analysis I set α to 2.5, which implies reasonable levels and ranking of amenities across metropolitan areas as it will be shown later. The qualitative nature of the results does not depend on the exact value of this parameter. Setting α between 2 and 3 yields very similar results.

In principle, ϕ could be estimated separately for each MSA. Here, I fix ϕ to be the same across metropolitan areas in order to keep the interpretation of the other parameters of the model simple. The estimate parameter here is $\hat{\phi} = .0315$ and is significantly different from zero at the 99% level. The interpretation of this parameter is that x_t, which is the difference between prices and wages, does not revert very quickly to the mean. If the difference between prices and wages is too high, we would expect it to remain high for quite some time before it eventually reverts back to the normal level. The parameter b from the discrete time AR(1) regression (1.18) is .9689 (.0041) meaning that x_t is close to being non-stationary. The total variance parameter κ varies widely across metropolitan areas. The most volatile MSAs are San Francisco-San Mateo-Redwood City, CA (105.35), Santa Ana-Anaheim-Irvine, CA (97.88) and Honolulu, HI (97.21). The least volatile MSAs are South Bend-Mishawaka, IN-MI (3.98), Huntington-Ashland, WV-KY-OH (4.47) and McAllen-Edinburg-Mission, TX (5.26). I also estimate F_t^l for each metropolitan area from the residual of the regression (1.18). While there is no particular interest in the values of F_t^l, decomposing it into different factors is crucial for the estimation of the model.

Three Common Factors for Housing Markets

In order to estimate equation (1.14), we need estimates for σ_i^l which are the MSA's exposure to the underlying factors that drive the economy. This means that we need to decompose F_t^l in several factors and estimate the factor loadings σ_i^l for each metropolitan area. As it is very difficult to uncover all of the factors that drive the local economy for each city, I decompose F_t^l into a few factors that capture much of the variation in the data.[25] These factors are constructed very similarly to the popular factors that Fama and French (1993) use to explain common variation in the financial asset markets.[26]

The first factor, denoted housing market (HMKT hereafter), is the annual house price returns for the whole US housing market. The second factor, denoted housing small-minus-big (SMBH), is defined as the average returns in MSAs in the bottom half of the price level distribution in a given year minus and the average return in MSAs in the top half. This factor replicates a self-financing diversified portfolio that holds houses in low priced metropolitan areas and shorts houses in high price ones. The third factor, denoted housing high-minus-low (HMLH), is defined as the average returns in MSAs in the bottom 30 percent of the price level distribution in a given year minus the average return in MSAs in the top 30 percent. This factor replicates a self-financing diversified portfolio that holds houses in metropolitan areas with high price-to-wage ratios and shorts houses in areas with low price to wage ratio. Similar to the small-minus-big Fama-French factors, SMBH intends to capture size effects in MSA price returns. On the other hand, HMLH is intended to capture growth effects since a high price to wage ratio can be an indicator of high expected growth.

[25] Hizmo (2010, 2) shows that the same three factors considered here explain about 90 percent of the time-series and cross-sectional variation in the house price returns for twenty five diversified housing portfolios constructed by sorting metropolitan areas on price level and price over wage ratios. For individual metropolitan areas these three factors explain about 50 percent of the variation.

[26] The Fama/French factors are constructed using 6 portfolios formed on size (market capitalization) and book to market value ratios. The first factor is the excess return on the stock market. The second factor, small minus big, is the average return on the three small portfolios minus the average return on the three big portfolios. The third factor, high minus low, is the average return on the two value portfolios minus the average return on the two growth portfolios. These three factors combined can explain over 90% of the time-series and cross-sectional variation of any well diversified stock portfolio.

For ease of interpretation I orthogonalize the three factors through a series of regressions. First, I regress HMKT on HMLH and SMBH and redefine HMKT as the regression residual. Then I regress HMLH on SMBH and redefine HMLH as the regression residual. This procedure gives three factors that are orthogonal to each other. The qualitative nature of the results is not affected by this orthogonalization.

I decompose the local growth term F_t^l down into the effect of three factors and a residual term ε. The estimated regression for each metro area is:

$$F_t^l = \alpha^l + \sigma_{HMKT}^l HMKT_t + \sigma_{SMBH}^l \cdot SMBH_t + \sigma_{HMLH}^l \cdot HMLH_t + \varepsilon_t^i \quad (1.19)$$

The results are summarized in Table 1.1. The coefficients in this table are presented at the mean across metropolitan areas. In parentheses I show the number of times a coefficient is found to be significantly different from zero at the 90% level. The distribution of the estimated R-squared is summarized by its mean, minimum and maximum value. While the magnitude of these coefficients is hard to interpret, their sign is straightforward. On average, when the housing market factor HMKT increases, so do wages and house prices. The higher the increase in this factor the more prices will deviate from wages. A similar result holds for the HMLH factor. If the growth in high price to wage metropolitan areas is higher than in low price to wage ones, then on average prices and wages will increase, and the price-wage gap will increase. The opposite is found when low priced MSA appreciate faster than high priced MSAs as it can be seen by the average coefficient on the SMBH factor. The HMKT factor alone on average explains about 13% of the time-series variation, the SMB factor about 22%, and the HMLH factor about 16%. The three factors combined explain about 50% of the time-series variation on average, although the R-squared for particular metropolitan areas ranges from .02 to .90.

Table A.1 in the Appendix displays all of the factor loadings for the three factors and the estimated R-squared coefficients for all of the 216 metropolitan areas. The

Table 1.1: Time-Series Regressions of the Local Economic Base on Three Risk Factors

	(1)	(2)	(3)	(4)
HMKT	7.8257			7.8363
	(102)			(147)
SMBH		-11.7965		-11.7404
		(107)		(121)
HMLH			7.7780	7.7919
			(98)	(115)
R^2 distribution				
Mean	.1290	.2172	.1583	.5049
Min	.0001	.0001	.0001	.0264
Max	.4810	.7466	.7392	.9036
Groups	216	216	216	216

Note - The dependent variable in all of the regressions is the estimated process F that drives wages and prices in each metro area . Specifications (1)-(4) show the mean coefficients from time-series regressions that are run separately for each MSA. In parentheses is shown the number of times a coefficient is found to be significantly different from zero at the 90% level. The sample consists of annual data from 1985 to 2008.

MSAs with high R-squared coefficients are generally metropolitan areas with high volatility, growth, price levels, population and density. The opposite is true for MSAs with low R-squared coefficients. The top MSA's in terms of the R^2 are Washington-Arlington-Alexandria, DC-VA-MD-WV (.90), Bakersfield, CA (.87) Riverside-San Bernardino-Ontario, CA (.86). The bottom MSA's are South Bend-Mishawaka, IN-MI (.026), Charlotte-Gastonia-Concord, NC-SC (.027) and Elkhart-Goshen, IN (.047). In terms of factor loadings, cities like New York, Boston, Washington DC, San Francisco have large positive loadings on the HMKT factor, and large negative loadings on SMBH. MSA's near California or Florida generally have high loadings on the HMLH factor, while other large MSAs in Massasuchets, Connecticut or New York have negative loadings. Overall there is a large amount heterogeneity in the exposure of particular metropolitan areas to the three factors, which fits the general observation that local housing markets are very different from each-other, and do not usually follow the national market.

Are the Factors Spanned by Traded Assets?

While the three proposed factors do explain a large share of the variance in the underlying local growth term F_t^l, it may not be possible for homeowners to use them to hedge house price and income risk. First of all, it is not feasible for any household to hold multiple homes in several metropolitan areas to replicate any of the three factors. Perhaps, the only feasible way to hedge against these sources of risk would be to invest in financial assets that correlate with the three factors. In order to get an idea about how much of the variance of these three factors can be spanned by using stock returns I regress each of the three factors on the three original Fama-French factors and on REIT returns.

The REIT returns used here come from the aggregate NAREIT index for REITs that invest in mortgage issued on real estate and construction. The results are almost identical if we use Equity or Hybrid REIT indexes. The Fama-French factors are three portfolios constructed using 6 portfolios formed on market capitalization (size) and book-to-market value ratios. The first factor, Mkt-Rf is the excess return on the stock market. The second factor, small minus big (SMB), is the average return on the three small portfolios minus the average return on the three big portfolios. The third factor, high minus low (HML), is the average return on the two high book-to-market ratio portfolios minus the average return on the two low book-to-market ratio portfolios. These three factors combined can explain over 90% of the time-series and cross-sectional variation of any well diversified stock portfolio.

The results are presented in Table (1.2). The housing market factor HMKT is the only one that seems to be very correlated with traded assets. In the first specification, HMKT is regressed on REIT returns. The coefficient on REIT is positive and statistically significant with the magnitude and standard error of .0396 (.0116). The correlation coefficient between the HMKT and the REIT returns is about 0.6. As it can be seen in the second specification, adding the Fama-French factors does not increase the R-squared and all of the coefficients on these factors are not statistically significant. In both specifications, the coefficient on REIT is positive and statistically different from zero. When the same regressions are estimated for the SMBH and HMLH factor, all of the coefficients are found to be small in magnitude and not statistically significant from zero. The R-squared coefficients from these regressions are also very low. I interpret these results as evidence that homeowners are able to hedge most the HMKT factor risk by using tradable assets. On the other hand, tradable assets do not seem to be helpful at all at hedging the risk that is due to the SMBH and the HMLH factors.

1.3.2 Market Prices of Risk and Risk Premia

After estimating the factor loadings in the previous section, we have all of the ingredients to estimate the cross-sectional equation (1.14). For this we need an estimate of π^l, which can be given by:

$$
\begin{aligned}
\pi &= E\left(p_t^l - \alpha w_t^l - \frac{1}{r\left(r+\phi^l\right)} s_t^l + \frac{\phi^l m^l}{r^2\left(r+\phi^l\right)}\right) \\
&= E\left(p_t^l - \alpha w_t^l\right) - \frac{m^l}{r^2} \\
&= E\left(p_t^l - \alpha w_t^l - \frac{p_t^l - p_{t-1}^l}{r}\right)
\end{aligned}
$$

Table 1.2: Predicting Housing Factors with Stock Returns

	HMKT		SMBH		HMLH	
REIT	.0396**	.0488**	.0033	.0097	.0028	.0059
	(.0116)	(.0130)	(.0191)	(.0206)	(.0140)	(.0155)
Mkt-Rf		-.0111		-.0174		.0039
		(.0213)		(.0338)		(.0253)
SMB		.0028		.0150		-.0091
		(.0332)		(.0526)		(.0394)
HML		-.0081		-.0565		-.0242
		(.0276)		(.0438)		(.0328)
R^2	0.345	0.355	0.001	0.084	0.002	0.042
Years	24	24	24	24	24	24

Note - The dependent variables are the three housing factors. The independent variables are aggregate mortgage REIT returns from the FTSE NAREIT series, and the three Fama-French factors downloaded from Kenneth French's website. The sample consists of annual data from 1985 to 2008. The standard errors are given in parentheses.

** statistical significance at the 95% level

where the second equality uses the fact that $E(s_t^l) = m^l$ and the third equality uses the fact that $E\left(p_t^l - p_{t-1}^l\right) = m^l/r$. Instead of first estimating π and then estimating equation (1.14), I do it all in one step by using a between-effects panel data estimator that only uses cross-sectional variation in the data. The estimating equation is:

$$\pi_t^l = \frac{1}{r}\beta M^l - \frac{\lambda_1 \hat{\sigma}_{HMKT}^l + \lambda_2 \hat{\sigma}_{SMBH}^l + \lambda_3 \hat{\sigma}_{HMLH}^l}{r^2\left(r + \hat{\phi}\right)} - \gamma \frac{\left(\hat{\sigma}_{err}^l\right)^2}{4r^3\left(r + \hat{\phi}\right)} + e_t^l \tag{1.20}$$

where

$$\pi_t^l = p_t^l - \alpha w_t^l - \frac{p_t^l - p_{t-1}^l}{r}$$

The parameters $\hat{\sigma}$ are the factor loadings from the previous section and σ_{err} is the standard deviation of the residual from regression (1.19). The variable M captures amenities, which in this case are average temperature in January and July, average hours of sun in January, average humidity in July, a crime index that weights violent crimes ten times more than non-violent ones, population size, population density, housing density and percent are of water of MSA. The parameters to be estimated are β, λ and the risk-aversion parameter γ.

I estimate equation (1.20) to get estimates for the market prices of risk for the three proposed factors and for the parameter of risk aversion. The results are presented in Table 1.3. Specification (1) assumes that all the three factors are spanned and the residual variance $\left(\hat{\sigma}^l_{err}\right)^2$ is not. The estimated risk prices and the risk-aversion parameter are all statistically significant at the 95% percent level. As it can be seen from the estimate of the risk-aversion parameter, higher idiosyncratic variance leads to cheaper house prices. Households need to be compensated by one dollar decrease in current house prices for every increase of $300000 in the variance to their lifetime wealth. The R-squared coefficient is fairly high at .6859 meaning that the three factors explain a large share of the cross-sectional variation in prices and wages. Even if we didn't control for amenities, the R-squared would be rather high at .5165. In the next three columns, I repeat the same procedure under different assumptions about which factors are spanned and which aren't. Specification (2) assumes that both HMKT and SMB are spanned by traded assets while HMLH isn't. In specification (3) only the HMKT factor is spanned, and in specification (4) all of the volatility in each metropolitan areas is not spanned by any traded asset. The results are very similar under these different assumptions both in terms of magnitudes and in terms of their significance. The estimated absolute risk-aversion parameter is around $3 \cdot 10^{-6}$, which implies a relative risk-aversion parameter of 1 on average if we were to assume that a household lives in the median city in terms of volatility and earns \$20000 in net wages.

Using the estimates in Table 1.3 together with the factors loadings for each metropolitan area, we can estimate the MSA specific risk premia and the implied amenities from equation (1.20). Using risk prices from specification (1) of Table1.3, the appendix Table A.2 shows the estimated risk premia and the implied amenity value for all the metropolitan areas in alphabetical order. A sample of the fifty most populated MSAs sorted by risk premia is displayed in Table 1.4. Negative estimates for risk premia are interpreted as the dollar amount a homeowner needs to be compensated for living in a risky MSA. Santa Ana-Anaheim-Irvine, CA is the MSA with the highest risk premium of -\$140048. This means that prices there are \$140048 cheaper than they would be if all homeowners were risk-neutral. Since homeowners are risk-averse, they are compensated by cheaper home prices for taking the risk of owning. On the bottom of the table we can see that Denver-Aurora-Broomeld, CO experiences home prices than are about \$20000 higher than they would be if homeowners were risk-neutral. This is because Denver has almost opposite factor loadings to Santa Anna or San Francisco and lower overall volatility. Overall, coastal cities have large risk premia while the Midwest and the southern metropolitan areas have the lowest. The part of home prices that is not due to wages or risk premia is used as an estimate for amenities in that MSA. For example, after taking out the effects of wages, expected growth and risk premia from home prices in Santa Ana, the implied value of amenities is about \$300000. Out of the fifty largest cities in Table 1.4, the worst MSA in terms of amenities is Detroit-Livonia-Dearborn, MI with a value of about -\$32000 and the best metropolitan area is San Francisco-San Mateo-Redwood City, CA with a value of amenities of \$316297.

Table 1.3: Cross-Sectional Estimates of Risk Prices for Traded Factors

	(1)	(2)	(3)	(4)
$\hat{\lambda}_{HMKT}$	-.1166**	-.2093**	-.0473	
	(.0591)	(.0629)	(.0648)	
$\hat{\lambda}_{SMBH}$	-.1285**	-.1223**		
	(.0241)	(.0263)		
$\hat{\lambda}_{HHML}$	.1649**			
	(.0302)			
$\hat{\gamma}/10^5$	-.3462*	-.5932**	-.3163**	-.2618**
	.1944	(.1973)	(.0735)	(.0575)
Amenities	Yes	Yes	Yes	Yes
R^2	0.6859	0.6471	0.5596	0.5559
Groups	177	177	177	177
N. Obs.	4230	4230	4230	4230

Note - The estimated coefficients represent the market price of risk for each of the factors if they are traded. The dependent variable is the part of house prices that is not due to wages and expected growth. The independent variables are the factor loadings estimated from time-series regressions multiplied by $1/\left(r^2\left(r+\lambda\right)\right)$. The absolute risk-aversion parameter $\hat{\gamma}$ is the coefficient on the variable $\hat{\sigma}_{ERR}$, which is the standard deviation of the residual from each time-series regression multiplied by $e^{30r}/\left(4*r^3\left(r+\lambda\right)^2\right)$ as the model predicts. The coefficients in this table are estimated by a panel data between estimator that only uses cross-sectional variation. The amenities included are weather, crime, population, density, and percent water area in MSA. The standard errors are shown in parentheses.

* statistical significance at the 90% level

** statistical significance at the 95% level

Table 1.4: The Fifty Most Populated MSAs Sorted on Risk Premia

	Price	Wages	Amenities	Risk Premia
SantaAna-Anaheim-Irvine,CA	482828	51894	316297	-140048
SanFrancisco-SanMateo-RedwoodCity,CA	714716	76042	382048	-135615
SanJose-Sunnyvale-SantaClara,CA	520378	58531	304677	-130572
LosAngeles-LongBeach-Glendale,CA	340842	42265	200361	-97654
SanDiego-Carlsbad-SanMarcos,CA	361444	46649	208206	-90688
Oakland-Fremont-Hayward,CA	351486	53093	177091	-88480
Sacramento-Arden-Arcade-Roseville,CA	235755	41119	137359	-77007
Riverside-SanBernardino-Ontario,CA	191675	30634	115268	-62994
WestPalmBeach-BocaRaton-BoyntonBeach,FL	243844	58358	49911	-59038
FortLauderdale-PompanoBeach-DeerfieldBeach,FL(MSAD)	239411	41974	94247	-57586
Miami-MiamiBeach-Kendall,FL	251186	35887	89587	-49034
LasVegas-Paradise,NV	178390	39920	60215	-44996
Baltimore-Towson,MD	267441	47881	78048	-43632
Orlando-Kissimmee,FL	192993	35717	70500	-42265
Phoenix-Mesa-Scottsdale,AZ	156698	36156	37786	-34795
Edison-NewBrunswick,NJ	330532	51865	112694	-34683
Wilmington,DE-MD-NJ	244404	43643	86808	-32946
Newark-Union,NJ-PA	390079	56655	140479	-31607
NewYork-WhitePlains-Wayne,NY-NJ	402722	54540	140388	-31317
Nassau-Suffolk,NY	411170	57617	132911	-30188
Portland-Vancouver-Beaverton,OR-WA	276563	39942	85134	-29251
Camden,NJ	195479	42626	44754	-27447
Providence-NewBedford-FallRiver,RI-MA	227287	40887	69347	-22110
Philadelphia,PA	190433	47361	17739	-21036
Chicago-Naperville-Joliet,IL	248207	45510	70747	-15429
Milwaukee-Waukesha-WestAllis,WI	230988	42824	67918	-11209
LakeCounty-KenoshaCounty,IL-WI	218836	51782	33225	-10426
Boston-Quincy,MA	336747	55220	92699	-7879
RockinghamCounty-StraffordCounty,NH	239595	45231	57195	-6151
Minneapolis-St.Paul-Bloomington,MN-WI	186015	47653	7047	-5723
Cambridge-Newton-Framingham,MA	387527	60093	127056	-4981
Peabody,MA	318301	50895	107415	-4670
Gary,IN	117579	35922	330	-3063
St.Louis,MO-IL	137199	41823	-2587	-2983
Nashville-Davidson-Murfreesboro-Franklin,TN	167671	39768	16522	-1207
Warren-Troy-FarmingtonHills,MI	123421	44488	-13735	314
Detroit-Livonia-Dearborn,MI	58617	32094	-32071	533
SanAntonio,TX	151913	34937	33157	1326
Columbus,OH	165190	38741	37593	2023
Charlotte-Gastonia-Concord,NC-SC	197281	39621	41230	2221
Indianapolis-Carmel,IN	122092	39297	-1721	2270
Cincinnati-Middletown,OH-KY-IN	166943	39066	37096	2389
KansasCity,MO-KS	144052	40396	8017	2919
Pittsburgh,PA	126616	42104	-15350	3052
Cleveland-Elyria-Mentor,OH	144077	40118	21777	5183
Atlanta-SandySprings-Marietta,GA	191194	38336	47676	6041
Houston-SugarLand-Baytown,TX	159217	45835	1872	6079
Dallas-Plano-Irving,TX	149710	43458	19554	7398
NewOrleans-Metairie-Kenner,LA	163257	41740	-5321	13291
Denver-Aurora-Broomfield,CO	258493	48010	51834	20281

Note - All the variables are for year 2008 and given in 2008 dollars. The risk premia is estimated by the exposure of a metropolitan area to the three factors and to the idiosyncratic variance. A negative risk premium should be interpreted as a cheaper house price due to the exposure of the MSA to risk. The amenities are calculated as the amount left over from prices after removing the effect of wages, expected growth, and risk premia.

1.4 Simulations

Using estimates for 216 US metropolitan areas I simulate the model to study the effect of financial innovation on house prices, household sorting across space and on overall productivity in the economy. In order to simulate the model I need estimates of the individual-MSA productivity match ε^l. I take estimates of this distribution from the previous literature.

1.4.1 The wage equation parameters

Using very detailed confidential migration data on a large set of individuals, Bishop (2008) estimates a wage regression:

$$w_{it}^l = f(age_i) + \omega_i \gamma + \mu_t^l + \theta_i^l + \eta_i + e_t$$

where wages of individual i who works in city l at time t are regressed on age dummies and individual specific characteristics ω_i and city-by-year fixed effects μ_t^l. The parameters of interest that are used here are the standard deviations for the individual-city match θ_i^l and the individual fixed effect η_i. In the model presented here these variables correspond to ε_i^t and ξ_i respectively. The estimated standard deviations that are used in my model are $\hat{\sigma}_\varepsilon = \$15499.21$ and $\hat{\sigma}_\xi = \$7073.40$ in year 2000 dollars.

1.4.2 Risk-Aversion Parameters

While there is some consensus in the literature about the magnitude of the relative risk-aversion (RRA) parameter at least for Constant Relative Risk Aversion utility functions, there is no agreement on the absolute risk-aversion (ARA) coefficient.[27] For these simulations I experiment with a range of absolute risk-aversion parameters

[27] See for example Vigna (2009)

anchored to be not to far from the estimates of Table 1.3. I simulate the model for ARA coefficients from $3 \cdot 10^{-6}$ to $3 \cdot 10^{-5}$, which in the simulations imply average RRA coefficients from .41 to 4.15. These values for the RRA are in line with estimates of the previous literature that find that the RRA coefficient is between 1 and 5. In the case where I allow for heterogeneity in risk-aversion I draw these parameters from a uniform distribution. The average ARA and RRA parameters are the same as above except for that I allow for bounds around the values given.

1.4.3 Results

I first simulate the model under the assumption that there is no heterogeneity in risk-aversion. I start from the baseline where there is no correlation between stocks and and factors that drive the local economy. I then consider the cases when new financial instruments are created that correlate with three factors proposed above and the case when the financial instrument allows for perfect insurance. The results from a series of simulations are presented in Table 1.5. In the first panel I study the effects of market completeness on house prices. In the first row the model is simulated for ARA of $.3 \cdot 10^{-5}$, which on average implies a RRA of .41 for the simulated wealth levels. Starting from the case where all variance is noninsurable, if we create an asset that spans the HMKT factor, prices will increase by 2.4%.[28] If in addition we create another asset that also spans the HMLH factor, prices will increase by a total of 3.14% in relation to what they where when no factors were spanned. If we span all the three factors, prices will go up by 4.24%. If we create assets that span not only the three factors but all of the remaining variation in prices and wages, prices would go up by about 6%. In the next few rows I simulate the model again with higher

[28] We have reason to believe that households can hedge against the HMKT risk factor by using existing financial assets as shown from the high R-squared estimates when we regress HMKT on traded financial assets in the previous section. If we agree that the HMKT factor is already spanned, we can alternatively interpret the negative of the magnitudes from these simulations as the effect of prohibiting the households from using financial assets to hedge the HMKT risk.

ARA and RRA parameters. Price changes are very sensitive to the risk-aversion parameter. In the extreme case of a RRA of 4.15 prices increase by 40% when we span all sources of variance.

The second panel in Table 1.5 shows welfare effects of completing the market in terms of the compensating variation. The compensating variation is the amount of dollars a household should be compensated after a policy change in order to reach the initial utility level. Here I interpret the compensating variation as the maximum payment a homeowner is willing to make in order to have access to a financial asset that spans a particular factor. In the first row the model is simulated for ARA of $.3 \cdot 10^{-5}$, which on average implies a RRA of .41 for the simulated wealth levels. Starting form the case where all variance is noninsurable, on average homeowners would be willing to pay \$858 dollars in order to gain access to a financial asset that spans the HMKT factors, \$2680 for a financial asset that spans all three factors, and \$3600 for one that spans all of the volatility in the market. The willingness to pay is again closely tied to the risk-aversion parameter. Homeowners are willing to pay \$7760, \$15097 and \$22125 if their RRA parameters were 1.38, 2.77 and 4.15 respectively. The willingness to pay is as high as \$31342 for an asset that spans all sources of volatility for the case of high RRA of 4.15.

Next I turn to simulating the model by allowing heterogeneity in risk-aversion over the population. The results from a series of simulations are displayed in Table 1.6. In each row the mean of the risk-aversion parameters is set to be the same as in Table 1.5. The first panel shows effect of completing markets on prices. In the first row for example, I draw the absolute risk-aversion parameter from a continuous uniform distribution $U(1, 59) \cdot 10^{-7}$, which for the average wealth implies RRA coefficients as if they were drawn from the uniform distribution $U(.01, .8)$. The effects on prices are similar to those previous table. The range of price increases goes from 2.67% when homeowners are given access to assets that span the HMKT factor to 7.12% when all

Table 1.5: The Effects of Completing Markets on Prices and Welfare when Households have Identical Risk Preferences

		HMKT	HMKT+ HMLH	HMKT+ SMBH	HMKT+ HMLH+SMB	All Variance
Price %						
ARA	RRA					
$.3 \cdot 10^{-5}$	0.41	2.64	3.14	3.74	4.24	5.89
$1 \cdot 10^{-5}$	1.38	4.08	5.75	7.77	9.43	14.92
$2 \cdot 10^{-5}$	2.77	6.15	9.48	13.52	16.85	27.83
$3 \cdot 10^{-5}$	4.15	8.21	13.21	19.27	24.26	40.73
Willingness to Pay $						
ARA	RRA					
$.33 \cdot 10^{-5}$	0.41	856	1243	2292	2680	3600
$1 \cdot 10^{-5}$	1.38	1622	2915	6467	7760	10895
$2 \cdot 10^{-5}$	2.77	2775	5411	12461	15097	21339
$3 \cdot 10^{-5}$	4.15	3901	7771	18255	22125	31342

Note - The model is simulated for 216 MSAs using parameters from estimated house price and wage processes. Each column displays the effect of creating financial instruments that correlate with the given housing factors. The last column displays the effects of creating financial instruments that span all of the variance in wages and prices. The first panel shows percent changes in average house prices, and the second panel shows the average compensating variation in dollars. The willingness to pay is the average amount a household would be willing to pay for the creation of the financial instrument. ARA stands for the coefficient of absolute risk-aversion. RRA is the coefficient of relative risk-aversion implied by the ARA for the average household.

possible sources of risk are spanned. Prices can jump up by 43% when all variance is spanned if the RRA is drawn from $U(.02, 12)$. For a more reasonable range of RRA parameters drawn from $U(.02, 7)$, prices increase from 5.5% when only the HMKT is spanned to 30% when all the sources of volatility are spanned.

The second panel in Table 1.6 simulates the effects of market completeness on productivity increases. In general the creation of tradable financial instruments that correlate with housing and income risk improves the households' ability to hedge risk and consequently lowers housing risk premia. In addition when there is heterogeneity in risk preferences, lowering the amount of noninsurable volatility leads to a different sorting of households across space. In the new sorting equilibrium, human capital is allocated more efficiently leading to higher overall productivity or wage level in the economy. The simulated effects on productivity are sensitive to the assumption on the distribution of the risk-aversion parameters. For a reasonable range, when the RRA parameters are distributed according to the uniform distribution $U(.02, 7)$, wages increase by 8.39% when all of the three factors are spanned, and they increase by 10.86% when all of the sources of volatility are spanned.

The last panel of Table 1.6 shows welfare effects of completing the market in terms of willingness to pay in dollars. The average willingness to pay for access to a tradable asset varies widely with the risk-aversion parameters as well. For the range of RRA parameters that are drawn from the uniform distribution $U(.02, 7)$ on average homeowners would be willing to pay $2009 dollars in order to gain access to a financial asset that spans the HMKT factors, $9431 for a financial asset that spans all three factors, and $13758 for one that spans all of the volatility in the market.

Taken together, results from these simulations can be taken as support to the idea that creating financial instruments that improve the households ability to manage risk is very beneficial in many levels. Particularly, under different assumptions about parameters values, simulations show that on average creating assets correlate with all of the sources of volatility significantly increase productivity, wages and welfare.

1.5 Conclusion

In order to understand the links between underlying risk factors, house prices and household location decisions I develop a micro-founded equilibrium model. The approach used is unique in that it merges standard methodologies used in urban economics, which study sorting and spatial properties of the problem, with models from continuous-time finance that study financial assets. This flexible and estimable model simultaneously considers risk-aversion, multiple sources of uncertainty, rich agent heterogeneity, sorting, portfolio choice and asset prices in one unified model. One key theoretical result is that home prices are derived to be a closed-form function of the underlying productivity of the economic base of a city minus a city-specific risk premium, which is a function of agent heterogeneity, sorting and risks in the economy. The problem of household sorting across space turns out to be very similar to that studied by Bayer et. al. (2005). Asset portfolio decisions are also found to be a generalized version of classic results in portfolio choice in finance.

The model is then estimated using US price and wage data and is simulated to study the effect of completing markets on house prices, household sorting across space and on overall productivity in the economy. The estimated risk premia imply that on average homes are about \$20000 cheaper than they would be if owners were risk-neutral, although there is large heterogeneity across metropolitan areas. Creating financial instruments that can be used for hedging purposes, lowers housing risk premia, increases welfare and increases productivity in the economy through sorting effects. For a reasonable range of risk-aversion parameters, I find that completing markets can increase home prices by about 20 percent, increase productivity in the economy by 10 percent and significantly improve welfare. The average willingness to pay for access to financial instruments that correlate with all of the sources of risk in the economy is between \$10000 to \$20000 per homeowner, depending on the assumed risk-aversion. This willingness to pay also varies widely across individuals and the sources of risk they are exposed to. Taken together, these findings draw attention to the potential benefits for creating financial instruments that correlate with house prices and income in every metropolitan area.

Table 1.6: Then Effects of Completing Markets on Prices, Productivity and Welfare with Heterogeneous Risk Preferences

		HMKT	HMKT+ HMLH	HMKT+ SMBH	HMKT+ HMLH+SMB	All Variance
Price %						
ARA	RRA					
$U(1,59)\cdot 10^{-7}$	$U(.01,.8)$	2.67	3.40	4.60	5.36	7.12
$U(1,199)\cdot 10^{-7}$	$U(.01,3)$	3.90	6.15	9.14	11.54	16.87
$U(1,399)\cdot 10^{-7}$	$U(.02,7)$	5.50	9.74	14.95	19.38	29.15
$U(1,599)\cdot 10^{-7}$	$U(.02,12)$	7.27	13.65	21.51	28.15	42.84
Productivity %						
ARA	RRA					
$U(1,59)\cdot 10^{-7}$	$U(.01,.8)$	0.29	0.74	2.20	2.43	2.62
$U(1,199)\cdot 10^{-7}$	$U(.01,3)$	0.78	1.84	6.69	8.04	9.78
$U(1,399)\cdot 10^{-7}$	$U(.02,7)$	0.91	2.29	8.39	10.86	14.59
$U(1,599)\cdot 10^{-7}$	$U(.02,12)$	1.01	2.67	9.16	12.47	18.58
Willingness to Pay $						
ARA	RRA					
$U(1,59)\cdot 10^{-7}$	$U(.01,.8)$	894	1280	2260	2665	3554
$U(1,199)\cdot 10^{-7}$	$U(.01,3)$	1344	2388	4561	5665	8119
$U(1,399)\cdot 10^{-7}$	$U(.02,7)$	2009	3910	7469	9431	13758
$U(1,599)\cdot 10^{-7}$	$U(.02,12)$	2782	5617	10953	13869	20257

Note - The model is simulated for 216 MSAs using parameters from estimated house price and wage processes. Each column displays the effect of creating financial instruments that correlates with the given housing factors. The last column displays the effects of creating financial instruments that span all of the variance in wages and prices. The first panel shows percent changes in average house prices, the second shows percent changes in average wages, and the third panel shows the average compensating variation in dollars. The willingness to pay is the average amount a household would be willing to pay for the creation of the financial instrument. ARA stands for the coefficient of absolute risk-aversion. RRA is the coefficient of relative risk-aversion implied by the ARA for the average household.

2

Beyond Signaling and Human Capital: Education and the Revelation of Ability

In traditional models of ability signaling (Spence (1973), Weiss (1995)), education provides a way for individuals to sort into groups (education levels) that are correlated with ability. Employers in turn use education to statistically discriminate, paying wages that depend in part on the average ability of the individuals with the same level of education. Building on these models, Farber and Gibbons (1996) and Altonji and Pierret (2001) develop a framework in which employers do not initially observe the ability of a worker but learn about it over time. As employers gather more information about the ability of a worker, they rely less on education and more on the new information in determining the wages. In these dynamic learning models, education serves as a tool for workers to signal their unobserved ability, although its role in determining wages decreases with experience.

In this paper, we argue that education (specifically, attending college) plays a much more direct role in revealing ability to the labor market. Rather than simply sorting individuals into broad ability groups, our results suggest that college allows individuals to directly reveal key aspects of their own ability to the labor market. Following in the tradition of the employer learning literature, the evidence that we provide is based on an examination of the returns to ability over the first 12 years of an individual's career.[1] Specifically, using data from the National Longitudinal Survey of Youth (NLSY), we show that the returns to the Armed Forces Qualification Test (AFQT), our measure of ability, are large for college graduates immediately upon entering the labor market and do not significantly change with labor market experience. In contrast, returns to AFQT for high school graduates are initially very close to zero and rise steeply with experience. These results suggest that key aspects of ability are observed nearly perfectly for college graduates but are revealed to the labor market more gradually for high school graduates.

There are a number of potential factors that likely contribute to ability revelation in the college labor market. Resumes of recent college graduates typically include information on grades, majors, standardized test scores and, perhaps even more importantly, the college attended.[2] In this way, our analysis leaves open the possibility that sorting of individuals across colleges may play a significant role in the revelation of ability in the college market. It does, however, imply a more limited role of educational attainment per se in signaling (as opposed to revealing) ability in the college market.[3]

[1] The main analysis presented in the paper limits the sample to males. Conducting a similar analysis for females is slightly more complicated due to greater concerns about selection into the labor market. Preliminary results for females that use the procedure outlined in Neal (2004) to deal with selection reveal similar patterns to those for males.

[2] In the analysis presented below, we show that this type of information explains a large portion of the variation in AFQT scores.

[3] This has important consequences for the large empirical literature that examines the extent to which the college wage premium is due to productivity enhancement versus ability sorting. See, for example, Fang (2006), Altonji and Pierret (1998), Lange (2007), Ashenfelter and Krueger (1994), Weiss (1995), Lang (1994), Stiglitz (1975), Mincer (1974) and Becker (1964). Our analysis also naturally suggests a reinterpretation of the findings of the employer learning literature following Altonji and Pierret (2001).

The insight that ability is revealed in the college market but not in the high school market has a great deal of power in explaining racial wage differences. In the college market, consistent with the notion that ability is almost perfectly revealed, we find that, if anything, blacks earn *more* than whites in the college market.[4] The lack of evidence for statistical discrimination in the college market is especially noteworthy given the large differences in the AFQT distributions for college-educated blacks and whites.[5] In contrast, we estimate that blacks initially earn 6-10 percent less than whites with the same AFQT scores in the high school labor market. Such a wage difference would arise naturally if employers use race to statistically discriminate when setting wages in the high school market. These results then provide an alternative explanation for the finding of Bjerk (2007) that blacks earn significantly less in blue collar occupations than whites but there is no racial wage gap in white collar occupations.

These results on the evolution of racial wage differences also provide a compelling explanation for the fact that, conditional on ability, blacks obtain more education than whites (Neal and Johnson (1996), Lang and Manove (2006)). Facing a wage penalty in the high school labor market (possibly due to statistical discrimination) but not in the college labor market, blacks clearly have stronger incentives to obtain a college degree than whites with comparable AFQT scores.[6]

[4] That blacks earn a premium in the college market is being driven by particularly high AFQT blacks. This may be due to affirmative action in the labor market operating most heavily where blacks are the scarcest.

[5] The mean AFQT for blacks is approximately one standard deviation lower than that of whites in both the high school and college samples.

[6] Another possibility is that the AFQT is racially-biased. With a racially-biased test, blacks would be more likely to attend college conditional on AFQT, all else equal. While it is very difficult to identify whether a test is racially-biased, Neal and Johnson (1996) note that the AFQT has been subject to rigorous examination to ensure that it is a racially-fair test. Further, if ability was higher for blacks once we netted out AFQT, we would expect blacks to perform better in college than whites. After controlling for SAT (which has received much more attention for racial bias than the AFQT), blacks still have much lower grade point averages than whites (see Betts and Morell (1999)).

The rest of the paper is organized as follows. Section 2 gives a general overview of the data we use for our empirical analysis. Section 3 presents our main empirical findings, which consist of a series of wage regressions. To fully interpret these findings, Section 4 uses the resulting coefficients to estimate a simple model of employer learning and statistical discrimination. Section 5 presents some additional specifications of our main estimating equations and Section 6 concludes.

2.1 Data

The data used in this study are drawn from the 1979-2004 waves taken from NLSY79. In selecting the sample, we follow the criteria used in Altonji and Pierret (2001) and Lange (2007) as closely as possible. Our main analysis is restricted to white or black men who have completed 12 years or 16 years of education, i.e. who have exactly a high school or a college degree. We consider a respondent to have entered the labor market the moment that he reports to have left school for the first time. Actual experience is the weeks worked divided by 50 and potential experience is defined as years since the respondent first left school.[7] If the respondent leaves the labor market and goes back to school, we subtract the added years of schooling from the experience measures. Military jobs, jobs at home or jobs without pay are excluded from the construction of experience and from the analysis.

The wage variable is the hourly rate of pay at the most recent job from the CPS[8] section of the NSLY.[9] In order to make our measure of ability, the AFQT, comparable across individuals, we standardize the AFQT score to have a mean zero

[7] Lange (2007) argues that this way of constructing potential experience captures time spent in the labor market better than age minus education minus seven.

[8] The CPS is a section of the NLSY79 that includes variables that establish activity during the survey week, job characteristics, global job satisfaction, hourly pay and hours worked per week for current/most recent job and job search behavior.

[9] The real wage is created using deflators from the 2006 Economic Report of the President. We limit real wages to more than one dollar and less than one hundred dollars per hour.

Table 2.1: Summary Statistics for College and High School Graduates by Race

	Blacks			Whites		
	Total	HS Grad	Col Grad	Total	HS Grad	Col Grad
Observations†	7,177	6,070	1,047	16,449	11,976	4,473
AFQT						
Mean	-.663	-.839	.358	.484	.253	1.102
Std. Dev.	.877	.769	.762	.805	.784	.460
Urban Residence (%)	83.40	81.83	92.52	72.59	68.67	83.08
Region (%)						
Northeast	14.79	15.23	12.24	21.38	20.22	24.49
North Central	15.64	14.81	20.46	35.60	36.71	32.63
South	62.54	63.94	54.40	28.31	28.09	28.92
West	7.03	6.02	12.91	14.70	14.98	13.96
Log of Real Wage						
Ages <25	6.47	6.45	6.84	6.61	6.58	6.83
Ages 25-30	6.65	6.58	7.02	6.88	6.80	7.07
Ages 30-35	6.71	6.61	7.13	7.02	6.91	7.26
Ages >35	6.80	6.71	7.23	7.13	6.98	7.45
Actual Experience						
Cum. Weeks Worked/52						
Ages <25	2.42	2.46	1.71	2.75	2.87	1.82
Ages 25-30	5.51	5.73	4.48	5.97	6.59	4.72
Ages 30-35	8.73	8.67	8.88	9.56	9.71	9.27
Ages >35	12.20	12.03	13.03	13.54	13.21	14.20
Potential Experience						
Years Since Left School						
Ages <25	3.37	3.46	1.76	3.30	3.52	1.54
Ages 25-30	7.67	8.28	4.70	7.16	8.40	4.56
Ages 30-35	12.36	12.95	9.69	11.89	13.20	9.30
Ages >35	17.41	18.00	14.46	17.02	18.21	14.57

† Individual by year observations coming from a panel from 1979-2004. In terms of individuals we have 1,917 whites and 798 blacks.

and standard deviation one for each age at which the test was taken.[10] We use data from the main and the supplementary sample of the NLSY79, which oversamples blacks and disadvantaged whites.[11]

We restrict the sample to observations where potential experience was less than thirteen years. The reason for this, as explained in the online appendix that replicates AP, is that there exists a nonlinear relationship between log wages, AFQT and potential experience. In order to keep the analysis simple, we focus on the approximately linear region of this relation. This region seems to correspond to experience levels less than thirteen. Another reason for this sample selection is attrition in the NLSY79, which implies that the number of observations falls noticeably with experience. A more detailed explanation of the sample construction is given in the online data appendix.

Table 2.1 summarizes the main variables in our sample. Notable from Table 2.1 are the differences in AFQT scores for blacks and whites of the same education level. For both college and high school graduates, this gap extends to about one standard deviation of the AFQT population distribution. It is also clear from the table that conditional on age, blacks generally earn lower wages and accumulate less labor market experience than whites.

An important exception to the general pattern of racial differences in wages in Table 2.1 is the fact that blacks and whites earn almost identical wages at the time of initial entry into the college labor market. At first glance, this unconditional statistic may seem surprising given that the average AFQT scores of college-educated whites

[10] Here we use the original definition of AFQT. We also estimated analogous specifications to those reported in the paper using AFQT89, which weights the underlying ASVAB sections that make up the AFQT differently; this had no effect on our results.

[11] All of the statistics in this study are unweighted. As shown in the online appendix, using the sampling weights has no effect on the qualitative resuls.

are about a standard deviation higher than those of their black counterparts.[12] As shown in Arcidiacono et al. (2008), this pattern is driven by the fact that college-educated blacks in the top decile of the AFQT distribution earn a substantial initial wage premium that declines to zero over the first ten years of labor market experience. We return to a more detailed discussion of racial differences in wages later in the paper.

2.2 Baseline Results

Given limited information, employers have incentives to rely on easily observed characteristics such as education and race to assess the productivity of a potential worker. In pure signaling models of education (Spence (1973), Weiss (1995)) education serves as a (costly) mechanism for workers to sort on ability. Employers then use the average group ability of the education level the worker belongs to determine wages. In many cases race can also be a predictor of ability, so employers use race when determining wages.

The employer learning literature argues that if AFQT is not directly observable by firms, it will have a limited relationship to initial wages. As workers spend more time in the labor market, employers become better informed about their ability, leading to an increased correlation between wages and AFQT with experience. As employers learn directly about ability, they need to rely less on correlates of ability and, therefore, the returns to education decline over time. These predictions have been shown to hold in Altonji and Pierret (2001) (AP thereafter) and Farber and Gibbons (1996). We replicate the main results of AP using our sample and present their results in the online appendix.

In this paper we argue that education is more than a tool for workers to signal

[12] Using Census data, Neal (2006) also documents that college-educated blacks and whites initially have similar wages upon entering the labor market.

their ability. Our hypothesis is two-fold: (i) that employers learn slowly about the ability of high school graduates and (ii) that ability is directly revealed for college graduates. If our hypothesis is true, pooling all education levels in wage regressions can lead to biases and the misinterpretation of the results. Examples of papers that pool all the education levels and analyze employer learning and statistical discrimination include AP, Bauer and Haisken-DeNew (2001), Farber and Gibbons (1996), Galindo-Rueda (2003) and Lange (2007). We test our hypothesis and analyze racial differences in wages and returns to ability by splitting the sample into college and high school graduates. We formulate a simple econometric model similar to that of AP, and estimate it separately for each of the two education levels. For each group, the log wage equation is:

$$\begin{aligned} w_i = \beta_0 + \beta_2 r_i + \beta_{AFQT} AFQT_i + \beta_{r,x}(r_i \times x_i) + \beta_{AFQT,x}(AFQT_i \times x_i) \\ + \beta_{r,AFQT}(r_i \times AFQT_i) + \beta_{r,AFQT,x}(r_i \times AFQT_i \times x_i) + f(x_i) + \beta'_{\Phi}\Phi_i + \varepsilon_i \quad (2.1) \end{aligned}$$

Log wages of individual i, w_i, are given as a linear interacted function of race r_i, AFQT, experience x_i, and other controls Φ_i. In all of our specifications we control for urban residence and for year fixed effects. We also report White-Huber standard errors that take into account correlation at the individual level over time.

2.2.1 Education and learning

Following the interpretation of AP, if employers do not initially observe ability but learn about it over time, the weight placed on AFQT should be small initially and increase with experience. This means that β_{AFQT} should be close to zero, and $\beta_{AFQT,x}$ should be positive and sizable. On the other hand, if employers directly observe AFQT, the returns to AFQT should be high initially and should not change much over time. This case translates to a large β_{AFQT} and a relatively small $\beta_{AFQT,x}$. We estimate equation (2.1) separately for high school and for college graduates and present the results in Table 2.2. Because we are working with log wages, β_{AFQT} is the percent change in real wages as a response to an increase of AFQT by one standard deviation. We divide the interaction of any variable with experience by ten so the coefficient $\beta_{AFQT,x}$ is the change in the wage slope between the periods when $x = 0$ and $x = 10$.

Specification (1) in Table 2.2 estimates equation (2.1) for our high school sample by setting $\beta_{r,AFQT}$ and $\beta_{r,AFQT,x}$ to zero. This is the equivalent specification of AP for our high school sample. The coefficient on AFQT is very small and statistically insignificant, suggesting that there are no returns to AFQT at the time of initial entry into the labor market. This is consistent with the view that AFQT is not readily observable to employers when they set wages.

In contrast, the coefficient on AFQT interacted with experience is positive and significant. The coefficient estimate implies that an individual with 10 years experience would see an increase in wages of almost 13 percent from a one standard deviation increase in AFQT. The results do not change under specification (2), which includes additional controls for region of residence and part time jobs. These results for the high school sample are consistent with standard hypothesis put forth in the employer learning literature: employers initially observe ability imperfectly but learn about it over time.[13]

Specifications (3) and (4) of Table 2.2 repeat the same empirical exercise for the college market, revealing a very different experience profile for the returns to AFQT. In specification (3), the coefficient on AFQT is large and statistically significant while the coefficient on AFQT x exper/10 is small and not statistically significant. In con-

[13] There is another potential explanation for the AFQT-experience profile revealed in specifications (1) and (2) of Table 2.2. In particular, the observed profile may simply reflect the actual impact of AFQT on the productivity of high school graduates as they gain experience in the labor market. Perhaps AFQT does not matter for the entry-level jobs performed by high school graduates but matters more as workers gain experience. We take up this issue formally in Section 4 below, where we develop a model of employer learning and statistical discrimination.

Table 2.2: The Effects of AFQT on Log Wages for High School and College Graduates

	High School		College		Test: College=HS P-values	
Model	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	.0060	.0078	.1485**	.1420**	0.000	0.000
	(.0130)	(.0129)	(.0350)	(.0354)		
AFQT x exper/10	.1261**	.1183**	.0122	.0198	0.026	0.050
	(.0176)	(.0173)	(.0480)	(.0472)		
Black	-.0628**	-.0483*	.1098*	.1125**	0.006	0.007
	(.0267)	(.0259)	(.0563)	(.0543)		
Black x exper/10	-.0358	-.0340	-.1304*	-.1264*	0.223	0.255
	(.0350)	(.0345)	(.0694)	(.0677)		
R^2	0.1631	0.1874	0.1678	0.1821		
No. Observations	11795	11772	4112	4112		
Add. controls	No	Yes	No	Yes	No	Yes
Experience measure: Years since left school for the first time <13						

Note - All specifications control for urban residence, a cubic in experience and year effects. Specifications (2) and (4) also control for region of residence and for part time vs full time jobs. In specification (5) we report the P-values for the difference in the coefficients of specifications (1) and (3). Similarly specification (6) compares (2) and (4). The White/Huber standard errors in parenthesis control for correlation at the individual level.
* statistical significance at the 90% level
** statistical significance at the 95% level

trast to the high school sample, there are substantial returns to AFQT immediately upon entry into the labor market: a one standard deviation increase in AFQT is associated with an almost 15 percent increase in wages. Moreover, the returns to AFQT are only slightly affected by experience, rising only an additional percentage point after ten years. Interpreted through the lens of the employer learning literature, this AFQT-experience profile suggests that employers observe AFQT nearly perfectly at the time of initial entry into the college labor market and learn very little additional information with experience.[14]

In specifications (5) and (6) of Table 2.2 we test if the coefficients presented are

[14] While AFQT represents only a single dimension of ability, remarkably similar patterns emerge for an alternative correlate of ability - father's education. Controlling for father's education show the same qualitative patterns: it does not correlate with high school wages initially but becomes correlated over time while in college the correlation starts strong and does not change over time. Although all the patterns are the same as with AFQT, the estimates were generally noisy and only sometimes significant.

significantly different in the college versus the high school market. Specification (5) presents the P-values of the difference between specification (1) and (3). We find significant differences between the college and high school coefficients for AFQT and AFQT x exper/10. Similar results can be seen in specification (6) where we include additional controls. Overall, it is clear from all specifications that there are significant differences between the college and high school samples in both the initial returns to AFQT and its experience profile.

2.2.2 How College Reveals AFQT

There are a number of potential factors that likely contribute to ability revelation in the college labor market. Resumes of recent college graduates, for example, typically include information on grades, majors, standardized test scores and, perhaps even more importantly, the college from which the individual graduated. Some of this information likely to be found on resumes is available to us so we attempt to understand how it reveals underlying ability. We do not have information on grades or the quality of the college attended, but standardized test scores such as SAT, PSAT and ACT that we observe should serve as a proxy for college quality.

Table 2.3 displays the coefficient estimates for regressions of AFQT on college major, years in college and standardized test scores interacted with race. We are interested in the general fit of the regressions rather than in the individual coefficients. The R-squared in specifications (1) through (3) ranges from 0.6225 when we control for PSAT scores, to 0.7024 when we control for SAT scores. The number of the observations in these regressions, however, is low since these test scores are available for very few college graduates in the NLSY79. In specification (4) we include only years in college and major fixed effects interacted with race. The R-squared here is lower than when we include test scores, but it is still at a sizable magnitude of .3940. Overall, these regressions, which include variables describing only some of

Table 2.3: Predicting the AFQT for College Attendees

Dep. Variable: AFQT	(1)	(2)	(3)	(4)
SAT Math Sect./10	.0258**			
	(0031)			
SAT Verbal Sect/10	.0093**			
	(.0035)			
PSAT Math Sect./10		.2403**		
		(.0257)		
PSAT Verbal Sect./10		.0686**		
		(.0251)		
ACT Math Sect./10			.2862**	
			(.0449)	
ACT Verbal Sect./10			.3434**	
			(.0676)	
Years in college	.0121	.0145	-.0013	.1560**
	(.0268)	(.0217)	(.0216)	(.0159)
R^2	0.7024	0.6225	0.6494	0.3940
No. Observations	224	311	276	1173

Note. - All the specifications above are interacted with race and twenty four college major dummies.
* significance level at the 90% level
** significance level at the 95% level

the information contained on a typical resume of individuals who attended college, indicate that this dimension of ability may be essentially revealed at the time of initial entry into the labor market.

2.2.3 Racial Differences

Racial differences in wage profiles

There are significant differences in the average AFQT of whites and blacks in both the high school and college samples. As shown in figure 2.1, the mean and the median of the black distribution lie about one standard deviation below the white distribution for both high school and college graduates.[15] Indeed, the median AFQT

[15] Similar findings about achievement tests gaps are documented earlier in the literature; see Neal (2006) for a detailed discussion.

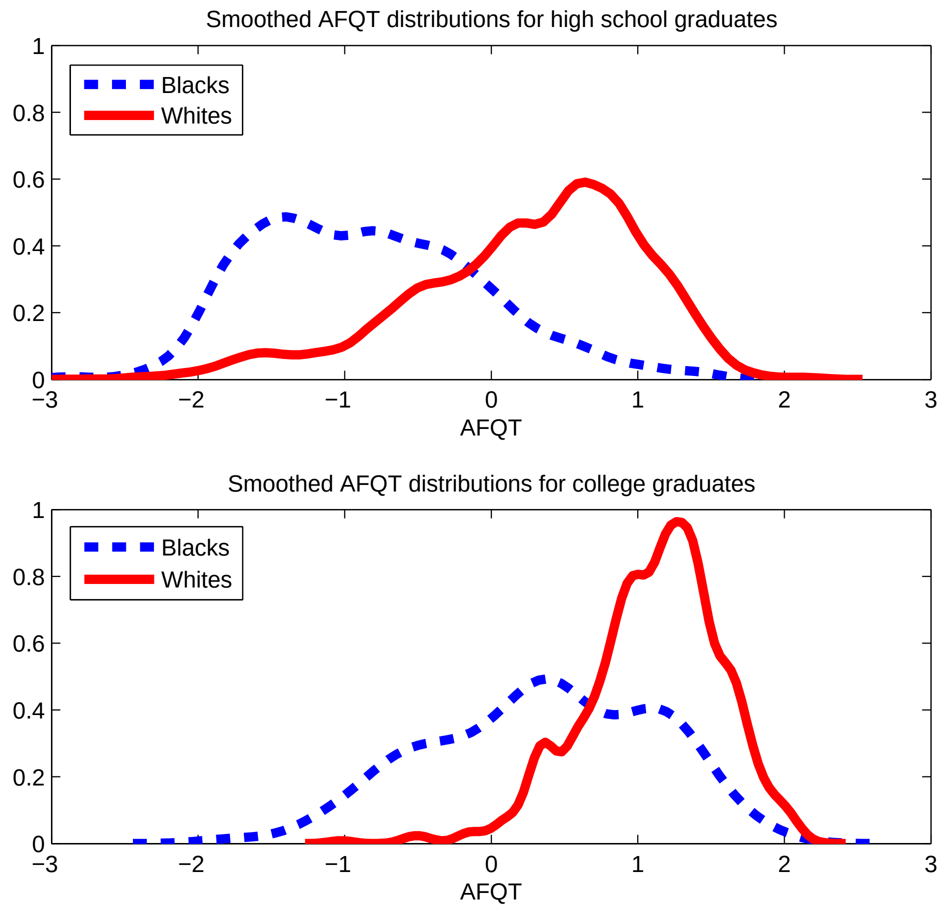

FIGURE 2.1: AFQT Distributions by Race and Education

score for blacks who graduate from college equals the median AFQT score for whites who only attend high school. As a result, if employers do not directly observe ability, there are strong economic incentives to statistically discriminate on the basis of race.

Given the results for employer learning discussed in the previous subsection, we would expect the incentives for statistical discrimination to be strong in the high school market, where ability is initially unobserved. This is reflected in the results presented in Table 2.2, which imply that blacks earn wages that are about 6 percent lower than those received by whites with the same AFQT score at the time of initial entry into the labor market. This gap increases (insignificantly) with labor market experience so that the estimated racial wage gap at ten years of experience, conditional on AFQT, is 10 percent.[16]

Conversely, given the results for employer learning, we would expect incentives for statistical discrimination to be weak in the college market, as ability appears to be nearly perfectly observed at the time of entry into the college market. Specification (3) in Table 2.2 shows that, conditional on AFQT, college-educated blacks earn eleven percent *higher* wages than their white counterparts upon initial entry into the labor market.[17] This premium declines to zero after about 9 years of labor market experience.[18] It is important to note here that the black college premium does not become negative after 10 years but is a result of imposing linearity on the experience interaction.

The results reported in Table 2.2 strongly suggest that statistical discrimination is not present in the college market. This fits well with the notion that college reveals ability. The lack of statistical discrimination in the college market is especially remarkable given the sizable differences in the distributions of AFQT between whites and blacks.

[16] That the racial wage gap in the high school market increases with experience is inconsistent with standard models of employer learning and statistical discrimination (see, for example, Altonji and Pierret (2001)). These models would predict that employers should weight race less as they learn more directly about worker productivity. In the next section we formulate and estimate a model of employer learning and statistical discrimination that can accommodate an increasing racial wage gap. Our model differs from the existing models in that it allows for the true productivity of AFQT to change with experience. We decompose the coefficients on *Black* and AFQT into the part that comes from employer learning, and the part that comes from the true productivity of AFQT increasing over time. We find that AFQT becomes more important for productivity over time, and this generates increasingly stronger incentives for employers to statistically discriminate. If employers decide to statistically discriminate, the racial wage gap may indeed widen with experience.

[17] In Arcidiacono et al. (2008) we show that much of the large initial premium and its sharp decline is mainly driven by especially high wages of a small number of blacks in the top decile AFQT distribution.

[18] The asymmetry of racial differences in the college and high school markets documented here is similar to results reported in Arcidiacono (2005). Using the NLS72 sample and analogous controls, Arcidiacono finds that blacks that attend at least some college earn more than their white counterparts, while blacks with a high school degree earn less, all else equal.

It is important to emphasize that the existence of an initial wage premium for college-educated blacks, conditional on ability, is a robust feature of the US labor market. Using much larger samples drawn from US Census data and the CPS, Neal (2006) shows that college-educated blacks and whites have similar wages at the time of initial entry into the labor market. Given racial differences in average AFQT scores, this pattern implies the existence of a substantial black wage premium.

Explaining racial differences in education attainment

Lang and Manove (2006) (LM, hereafter) report that, conditional on AFQT, blacks obtain more education than whites.[19] Having documented this key empirical fact, LM attempt to explain it.[20] They develop a model in which employers generally observe a noisier signal for blacks than whites but this racial difference in precision declines with education. This mechanism certainly provides an increased incentive for blacks to earn more education but also implies that, conditional only on AFQT, blacks should earn more than whites, an implication generally not supported by the data.[21] Statistical discrimination arises through a different channel in our model. We do not need the crucial assumption made in LM that employers observe a noisier signal for blacks. In our view, statistical discrimination arises because employers have some idea of the racial differences in the AFQT distributions in Figure 1. In Arcidiacono et al. (2008) we test if there are any racial differences in the initial signal observed and in employer learning by adding interactions of race with AFQT and

[19] A similar result can be seen for our sample in Figure 2.1, which reports the AFQT distributions for blacks and whites in both high school and college. This fact has important implications for how one thinks about racial wage differences, implying, for example, that estimating the black-white wage gap properly requires one to control for both AFQT and education.

[20] LM first rule out differences in school quality as a potential explanation. They reason that because blacks generally attend lower quality schools they may require more education in order to reach a given level of cognitive ability. They conclude, however, that school quality differences while present cannot possibly explain the observed racial differences in educational attainment.

[21] Neal and Johnson (1996) shows that, conditional on only AFQT, the racial wage gap is smaller but blacks continue to earn less than otherwise identical whites.

AFQT × experience. We did not find any significant differences in the returns to AFQT between whites and blacks in either the high school or college samples though the results are noisy.

The view of the labor market suggested by our main findings provides a related and more direct explanation for why blacks obtain more education than whites with the same AFQT score. Facing statistical discrimination in the high school labor market (where ability is initially unobserved) blacks have a greater incentive to enter the college labor market and thereby revealing their AFQT. Thus, education symmetrically improves the precision of the signals that employers get for blacks and whites but, because the value of that increased precision is greater for blacks, blacks obtain more education.[22]

2.2.4 Structural change at college graduation

In this section we investigate whether there is a discrete jump in our coefficients of interest at college graduation. To do this we estimate a similar model to AP where we pool all education levels 8-20 in the same regression and include linear interactions of education with all the variables. In order to test for a structural shift at college graduation, we interact our variables of interest with an indicator variable that takes the value of one if the individual has sixteen or more years of education. To make comparisons between these results and the results in Table 2.2, we subtract twelve off of years of education so that when education is zero we get back results for the high school market.

The baseline coefficients as well as their interactions with education and education

[22] Our explanation is also consistent with the fact that, conditional on only AFQT, blacks earn lower wages than whites on average. In the college labor market, our results suggest that blacks earn more than whites with identical AFQT scores, while in the high school labor market blacks at least initially earn 6 percent less than identical whites regardless of their AFQT score. On average whites can earn more than blacks with the same AFQT score, provided that the college attendance of blacks is not enough to offset the wage penalty that blacks face in the high school market.

Table 2.4: Testing for Structural Break at College Graduation

Interacted with:	None	Education	$I(Educ \geqslant 16)$
Standard. AFQT	.0148	-.0109	.1874**
	(.0109)	(.0066)	(.0437)
AFQT x exper/10	.1061**	.0084	-.0939*
	(.0508)	(.0089)	(.0508)
Black	-.0504**	.0102	.1658**
	(.0211)	(.0145)	(.0771)
Black x exper/10	-.0587**	-.0018	-.0369
	(.0265)	(.0179)	(.0857)
R^2	0.3080		
No. Observations	25692		
Experience Measure: Years since left school for the first time<13			

Note - In order to make the baseline coefficients in the first column comparable to Table 2 we subtract 12 from our education measure. We pool all education levels -4 to 8 and estimate a pooled version of specification (1) in Table 2 by adding linear interactions of education with everything (Education column). We also interact our four variables of interest with a dummy that equals one if education is bigger or equal to sixteen (I(Educ=16) column). The White/Huber standard errors in parenthesis control for correlation at the individual level.

* statistical significance at the 90% level

** statistical significance at the 95% level

greater than or equal to sixteen are presented in Table 2.4. The baseline effects are similar to those for the high school market found in Table 2.2. Namely, AFQT is uncorrelated with wages initially but becomes correlated with experience and blacks face a significant wage penalty. None of the education interactions shown in column 2 are statistically significant. In contrast, column three shows evidence of a statistically significant shift in the AFQT, AFQT $\times$ exper/10 and Black coefficients, but not in the coefficient on Black $\times$ exper/10. The same patterns emerge as in Table 2 with strong initial returns to AFQT for college graduates that change little as individuals acquire experience. Further, black college graduates see a wage premium the first few years in the labor market.

Figure 2.2 presents graphical evidence of the shift in parameters following college graduation. Here we split the sample in 4 education levels: high school drop-outs (nine to eleven years of education), high school graduates (twelve years), college drop outs (thirteen to fifteen years), and college graduates (16 years). We estimate the parameters separately for each group and for the categories that include more than one level of education we also include controls for years of education.[23] Again the same patterns emerge for both the relationship between race and wages and the relationship between AFQT and wages, with college being a clear turning point in the relationships.

2.3 Statistical Discrimination

We argue that our results for the high school sample can be reconciled with statistical discrimination on the basis of race.[24] One scenario that rationalizes an increasing racial wage gap under the existence of statistical discrimination is the case when the true returns to AFQT increase with experience.[25] This is motivated by the intuition that AFQT should be more important for jobs further down the career path rather than for jobs taken upon initially entering the labor market. Under this scenario, blacks would be paid less initially since employers do not observe ability and therefore put weight on average group productivity. But because the true productivity of AFQT is increasing with time, employers have even stronger incentives to statistically discriminate over time. Thus, even though employers might learn about the productivity of their workers to some degree, they might increase the

[23] The results are not sensitive to whether we control for years of education.

[24] There could be other reasons why the wage gap between blacks and whites increases with experience such as increasing taste based discrimination or racial differences in on the job training. Without ruling out these explanations, we focus on whether the observed patterns can be explained by statistical discrimination alone.

[25] One may argue that if the true returns to AFQT are increasing in the high school market, why would they not also increase in the college market? As we will show, the true returns to AFQT level off in the high school market after 10 years of experience. Hence, obtaining 10 years of experience after high school puts those in the high school market in the same types of jobs as those who receive a college degree.

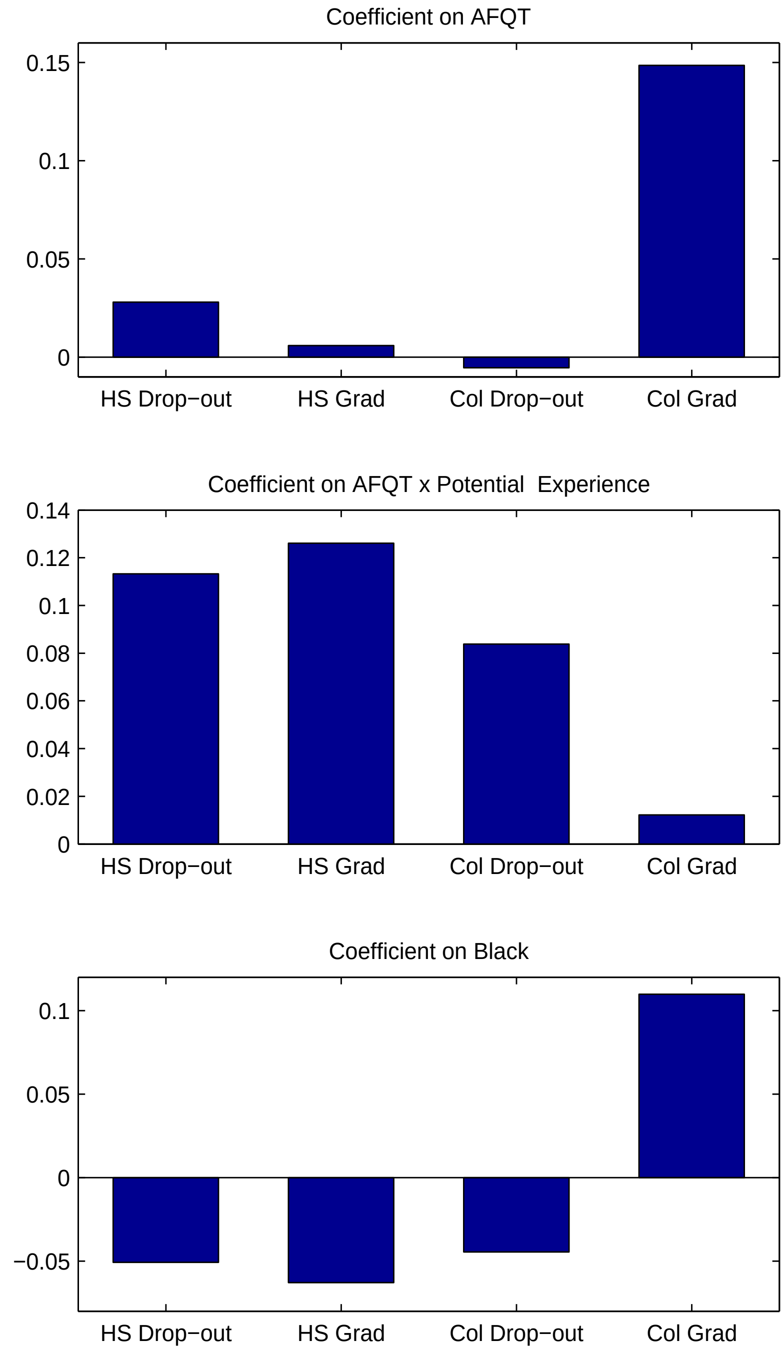

FIGURE 2.2: Plots of coefficients on AFQT, AFQT x potential experience and Black.

Note: There are 2471, 11795, 5090, and 4112 observations in each of the education categories starting from high school drop-outs to college graduates.

weight they put on race with time as a result of the increased incentive to statistically discriminate.

We estimate a model of statistical discrimination that incorporates the insights developed in this scenario. We estimate the model for the high school sample where employer learning appears to be relevant. The starting point for our model is the standard employer learning model formalized by Farber and Gibbons (1996). A model of learning closely related to ours is developed in Lange (2007). Lange estimates the speed of employer learning assuming symmetrical learning and a competitive labor market. We maintain these crucial assumptions in our specification. Our model differs from existing models in that we allow the true productivity of AFQT to vary over time for the reasons given above.[26]

The goal of this section of the paper is to study the implications for the speed of employer learning and the true productivity of AFQT over the early career assuming that the observed racial wage differences are driven entirely by statistical discrimination. Assuming that racial wage differences are driven entirely by statistical discrimination obviously rules out taste-based discrimination and other potential explanations for the wage gap. In discussing the results below, we describe how they would change if the wage gap was partially due to these other factors.

The model, which is fully described in the appendix, yields an estimating equation that relates log wages to a linear function of both an individual's own ability (which is initially unobserved), $AFQT$, and the mean ability of his race, $\overline{AFQT}$. A key assumption, common in the statistical discrimination literature, is that average ability for each race is known. The weight placed upon individual $AFQT$ may increase over time for two reasons: (i) employers learn about the individual's ability and (ii) the true productivity of ability may also increase with experience.

[26] Returns to AFQT could also change for reasons that are not captured in our model. If, for example, training is positively correlated with AFQT, the effect of additional training would resemble an increase in the true productivity of AFQT in our model.

We define the weights that the employer places on the individual's own ability and on the average ability of the individual's race at experience level x be given by Θ_x and $(1 - \Theta_x)$, respectively. We also define λ_x to be the true productivity of AFQT at experience level x. In the appendix we show that, under certain assumptions regarding the nature of learning and what employers initially know, log wages follow:

$$w_x = \lambda_x \left\{(1 - \Theta_x)\overline{AFQT} + \Theta_x AFQT\right\} + k_x \tag{2.2}$$

where k_x and experience-specific constant.

This representation of log wages has an intuitive interpretation. Log wages are a function of experience plus a weighted average of mean group ability, $\overline{AFQT}$, and actual ability, $AFQT$. The first source of the weight put on $\overline{AFQT}$ and $AFQT$ comes from employers learning over time. If initially employers do not observe anything that is correlated to $AFQT$, they rely on group averages to set wages. In this case $\Theta_x = 0$, so all the weight is put on $\overline{AFQT}$. As employers observe more signals about the productivity of the worker, the weight will gradually be shifted from the group mean to the individual's ability.[27] We show in the appendix that as experience increases $\Theta_x \to 1$. The rate of this convergence, which is also the speed of learning, will depend on the quality of the signals that employers get every period.

The other part of the weight put on $\overline{AFQT}$ and $AFQT$ comes from the true productive value of ability. This time-varying true productive value is captured by the parameter λ_x. As argued above, suppose ability is not as important for productivity for initial jobs as it is for jobs later in the career. In this case, λ_x will be low initially and increase over time, which means that additional weight will be put on both $\overline{AFQT}$ and $AFQT$ as time passes. If the true productive value of AFQT increases

[27] Similarly, employers distribute some weight on education intitially, which dicreases over time as employers learn more about ability. This education profile is captured in k_x. We do not pay particular attention to this since we are interesed in statistical discrimination on the basis of race and not education.

rapidly enough, the weight on $\overline{AFQT}$ can actually increase over time despite the fact that direct learning would naturally tend to decrease it. As long as λ_x increases faster than the speed of learning such that $\lambda_x(1-\Theta_x) > \lambda_{x-1}(1-\Theta_{x-1})$, more and more weight will be put on group average ability $\overline{AFQT}$.

We are only interested in the case of high school graduates, so education is held constant at 12. This means that we could estimate equation (2.2) directly by regressing log wages on mean AFQT for each race and AFQT for each experience level separately similar to Lange (2007):

$$w_{i,x} = \beta_{x,\overline{AFQT}}\overline{AFQT}_{race} + \beta_{x,AFQT}AFQT \cdot + \beta'_{\Phi}\Phi_{i,t} + \beta_x + \varepsilon_x \tag{2.3}$$

The parameter β_x captures the effect of the variables observed only by the employer previously denoted by k_x. Also $\Phi_{i,t}$ represents the demographic characteristics of a particular worker.

We can rewrite (2.3) as a function of the *Black* indicator variable rather than a function of $\overline{AFQT}_{race}$. In particular, we can rewrite the first term on the right hand side of (2.3) as:

$$\begin{aligned} \beta_{x,\overline{AFQT}}\overline{AFQT}_{race} =& \beta_{x,\overline{AFQT}}\left(\overline{AFQT}_{black} - \overline{AFQT}_{white}\right) Black \\ & + \beta_{x,\overline{AFQT}}\overline{AFQT}_{white} \\ =& \beta_{x,Black}Black + \beta_{x,\overline{AFQT}}\overline{AFQT}_{white} \end{aligned} \tag{2.4}$$

Note that $\overline{AFQT}_{black}$ and $\overline{AFQT}_{white}$ are the same for everyone. Letting

$$\beta^*_x = \beta_x + \beta_{x,\overline{AFQT}}\overline{AFQT}_{white} \tag{2.5}$$

we can write the wage equation as:

$$w_{i,x} = \beta_{x,Black}Black + \beta_{x,AFQT}AFQT + \beta'_{\Phi}\Phi_{i,t} + \beta^*_x + \varepsilon_x \tag{2.6}$$

This means that instead of including $\overline{AFQT}_{race}$ in equation (2.3), we could include a dummy variable that takes value one if the worker is black and zero otherwise and still be able to estimate the parameters $\beta_{x,\overline{AFQT}_{race}}$ and $\beta_{x,AFQT}$. In this case $\beta_{x,AFQT}$ would be unchanged and $\beta_{x,\overline{AFQT}_{race}} = \beta_{x,Black}/\left(\overline{AFQT}_{black} - \overline{AFQT}_{white}\right)$.

This provides a structural interpretation of the coefficient on Black in the regressions presented earlier in the paper. Employers put weight on race for two reasons: the first part $(1 - \Theta_x)$ is related to learning about ability, and the second part λ_x comes from the changing productivity value of this ability. The size and the sign of the coefficient on *Black* depends entirely on the experience profile of λ_x and Θ_x. Empirically, as can be seen in Table 2.1, the difference between the mean of AFQT for whites and blacks is 1.0922 for high school graduates. After estimating equation (2.6) we can then solve for λ_x and Θ_x:

$$\lambda_x = \beta_{x,AFQT} - \beta_{x,Black}/1.0922 \tag{2.7}$$

$$\Theta_x = \frac{\beta_{x,AFQT}}{\beta_{x,AFQT} - \beta_{x,Black}/1.0922} \tag{2.8}$$

We estimate equation 2.6 in one step by interacting Black and AFQT with a cubic in experience instead of estimating it separately for each experience level.[28] The estimation results are presented in Figure 2.3. The first two plots display the estimated coefficients on Black and AFQT for each experience level as well as the 90% confidence interval. The initial racial difference in wages is about 5 percent and in increases to 10 percent in about 10 years. The effect of a one standard deviation increase in AFQT starts at zero initially and increase to about 15 percent after 12 years of experience. These results are very similar to those previously shown in Table

[28] Estimating this equation for each experience level separately and smoothing the results yields almost identical results. We interact Black and AFQT with a cubic in experience instead for ease of presentation.

2.2.

We use these experience profiles to calculate how much of the changes in the returns to race and AFQT can be attributed to employer learning, and how much to changes in the true productive value of AFQT. Sub-figures 3 and 4 of figure 2.3 plot the learning parameter Θ_x and the parameter λ_x, which captures the evolution of the productivity of AFQT over time. The learning parameter starts near zero and by 12 years increases to 0.6, which means employers observe 60 percent of AFQT in about 12 years. The true productivity of AFQT is also increasing with experience. A one standard deviation increase in AFQT leads to a 6 percent increase in productivity initially which increases to about 24 percent in 12 years.

The weight put on race in the wage regression as a result of employer learning is given by $(1 - \Theta)$. This weight starts at 1.0 and declines to 0.4 after 12 years of experience. Initially employers do not observe ability so they rely heavily on the race of the worker to determine wages. As they learn about individual workers' productivities over time their incentives to statistically discriminate decrease and they rely less on race and more on the observed part of AFQT. This, however, does not mean that the actual return on race decreases with experience. Because the true return to AFQT, λ_x, increases over time, employers actually have stronger incentives to statistically discriminate at higher experience levels. Our estimates show that the effect of the increasing productivity of AFQT dominates the effect of learning in determining the coefficient on race early in the life cycle with the effects roughly canceling out after five years.

Part of the reason why blacks earn less than whites can be explained by the fact that they accumulate less labor market experience than whites. We do not model discrimination in the hiring process directly so our model cannot capture this source of inequality. In order to account for differences in actual experience, we include a cubic in actual experience in the estimation equation 2.6 and present the results in

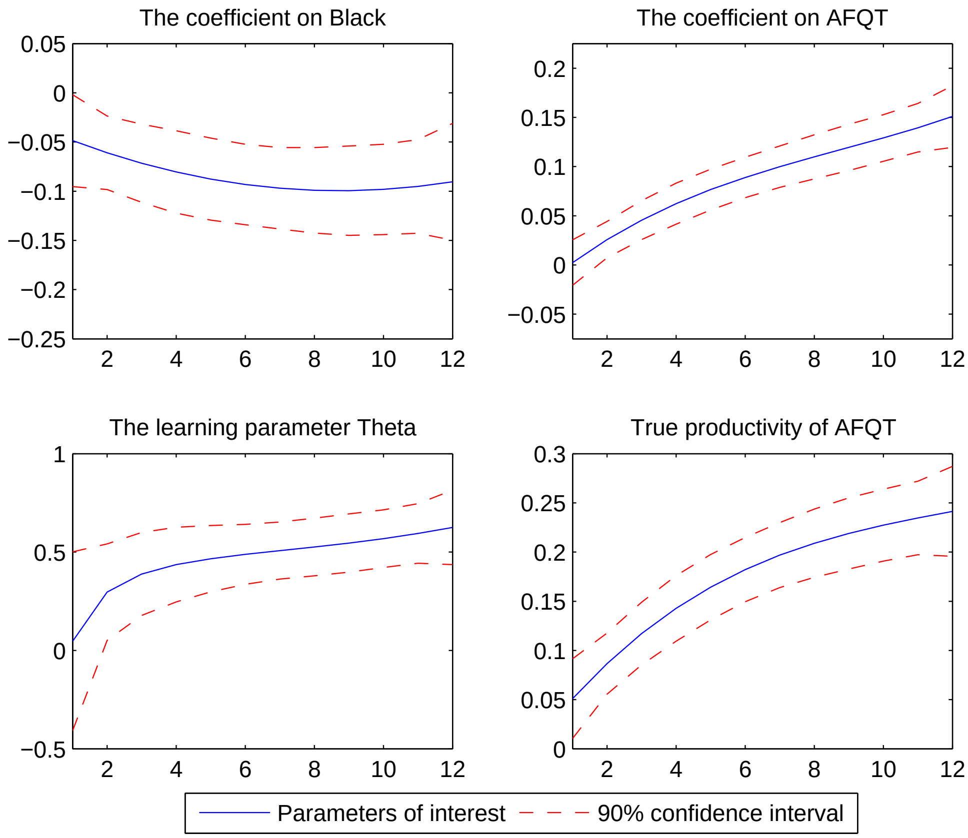

FIGURE 2.3: The Evolution of the Effect of Race and AFQT on Wages, the Learning Parameter Θ_x, and the True Productivity of AFQT, λ_x.

Figure 1 of the online appendix. As expected the coefficient on *Black* does not start as negative and does not fall as much as when we control for actual experience. The coefficients on AFQT and the learning parameter do not seem sensitive to controlling for actual experience. The true productivity of AFQT in subplot 4, however, starts out lower and peaks at about 0.19 as opposed to 0.24 in Figure 2.3. Even after controlling for actual experience our main results from Figure 2.3 remain: employers appear to statistically discriminate on the basis of race and learn about ability over time.

2.4 Robustness Checks

The results so far suggest that AFQT is nearly perfectly revealed in the college market but is only revealed over time in the high school market and that, consistent with statistical discrimination, blacks only receive lower wages in the high school market. In this section we check the robustness of the results along five dimensions. First, we investigate the sensitivity of the results to different assumptions regarding the determination of labor force participation. Second, we investigate whether the first four years of high school labor market experience play a similar role to college in revealing ability by seeing if our results change when we remove wage observations from the first four years after high school. If so, the differences between high school and college that we have documented might more accurately be characterized as age effects. Third, we show specifications that use father's education as a proxy for ability instead of the AFQT. Fourth, we include additional interaction or race and AFQT with year effects to see if time trends affect our results. Lastly, we check if patterns observed in the wage residual are consistent with our learning story.

2.4.1 Controlling for Selection

All of the results presented so far do not account for selection into the labor market. Differences in labor force participation by race can be very important when estimating log wage equations as shown in Butler and Heckman (1977) and Brown (1984). In order to control for selection, we could model the decision to participate in the labor force and estimate a rich structural model of wage offers and labor market entry decisions. This, however, proves to be too complicated for the purpose of this paper. Instead we follow Neal and Johnson (1996) by assigning an arbitrary wage to non-participants and estimate a median regression for the whole sample. If the wage offers that nonparticipants receive lie below the median wage offers participants re-

Table 2.5: The Effects of AFQT on Log Wages Controlling for Selection

	High School		College		College minus HS	
Model	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	.0180	.0261**	.1453**	1205**	0.007	0.020
	(.0138)	(.0157)	(.0511)	(.0416)		
AFQT x exper/10	.1395**	.1262**	.0162	.0485	0.061	0.154
	(.0212)	(.0194)	(.0668)	(.0559)		
Black	-.1196**	-.1007**	.0625	.0765	0.024	0.023
	(.0302)	(.0313)	(.0863)	(.0692)		
Black x exper/10	-.0369	-.0459	-.0811	-.0887	0.643	0.7741
	(.0444)	(.0418)	(.0967)	(.0922)		
Pseudo R^2	0.0464	0.0512	0.0850	0.0950		
No. Observations	13134	13108	4176	4176		
Add. controls	No	Yes	No	Yes	No	Yes
Experience measure:	Years since left school for the first time<13					

Note - We assign a zero log-wage to respondents who are not working at the time of the interview, and then estimate the log-wage equation using a median regression. All specifications control for urban residence, a cubic in experience and year effects. Specifications (2) and (4) also control for region of residence and for part time vs full time jobs. In specification (5) we report the bias-adjusted bootstrapped P-values for the difference in the coefficients of specifications (1) and (3). Similarly specification (6) compares (2) and (4).The standard errors reported are block-bootstraped and control for clustering at the individual level.

* statistical significance at the 90% level

** statistical significance at the 95% level

ceive, these median regression allow us in a crude way to control for selection. This approach of controlling for some form of selection is not, in our opinion, rigorous enough to be used throughout the paper. This method does not deal well with the fact that potential experience overstates actual work experience or that experience is endogenous, including the possibility that employers may not hire workers who they do not expect to be productive.

The results from these median regressions are presented in Table 2.5; these regression results mirror those presented earlier in the paper. Specification (1) estimates our baseline specification for the high school sample. The returns to AFQT are very small initially with a statistically insignificant coefficient of 0.018, but increase sharply in ten years with a statistically and economically significant coefficient of 0.1395. Blacks earn about twelve percent less than whites and this difference increases by an additional four percent in ten years. The same patterns can be seen in specification (2) where we control for region of residence and part time status.

Specifications (3) and (4) repeat the same procedure for the college sample. The returns to AFQT in specification (3) are initially very large and significant with a coefficient of 0.1453 and these returns do not change much with experience: the coefficient on AFQT×exper/10 is statistically insignificant with a magnitude of 0.0162. Similar patterns can be seen in the Black and Black×exper/10 coefficients although their magnitude is smaller and they are not significant. Including additional controls in model (4) does not change the qualitative nature of the results.

In specifications (5) and (6) of Table 2.5 we test if the coefficients presented are significantly different in the college versus the high school market analogously to specifications (5) and (6) of Table 2.2. The results in specification (5) closely resemble those of Table 2.2 in that there are significant statistical and economical differences between the college and the high school samples in the AFQT, AFQT x exper/10 and the Black coefficients. The findings are identical in specification (6) except for the slight statistical insignificance of the differences in the AFQT x exper/10 coefficients.

2.4.2 College Versus the First Four Years of Experience in the HS Market

When looking at the differences in the immediate returns to AFQT across the high school and college markets, one may be concerned that ability is actually revealed in the first four years after high school regardless of whether one attends college. If this is the case, the initial return to AFQT will be higher for college graduates than for high school graduates even if both college attendance and high school labor market experience reveal AFQT equally. In order to test this alternative explanation, we re-estimate the regressions in Table 2.2 and for high school graduates we exclude observations that come from the first four years in the labor market.

The results presented in specifications (1) of Table 2.6 are very similar to those of Table 2.2: the initial returns to AFQT are very low and insignificant at 0.0104, and by ten years this return increases to 0.1201.[29] The same patterns can be seen in specification (2) where we include additional controls. In columns (3) and (4) we test whether the coefficients for high school graduates that have been in the market for four years are different from those of the college graduates. As we can see from the P-values presented, all the coefficients are significantly different between the HS graduate (at experience=4) and the college graduate samples. The results of Table 2.6 confirm that ability is revealed much later in the high school labor market than in the college market.

2.4.3 Father's Education as a Measure of Ability

So far we have provided evidence that graduating from college reveals a single measure of ability AFQT. In this subsection we show that a similar pattern is found in the data for another correlate of ability that is difficult for employers to observe directly: father's education. We estimate the log wage regressions including father's education in Table 2.7. In all specifications father's education is divided by ten, so the coefficients should be interpreted as the return to a ten-year increase in father's education.

Specification (1) shows that, for high school graduates, the effect of father's education on log wages is initially small and statistically insignificant. Analogously to AFQT the returns to father's education increase significantly with experience implying that ten additional years of father's education yields a 15 percent increase in wages ten years after high school. In specification (2) where we also include AFQT and its interaction with experience. Both AFQT and father's education have

[29] Note that we do not change the experience variable implying that the minimum for the experience variable for high school graduates is five. Hence, the returns to AFQT at the minimum is the base return plus 0.5×0.1201.

Table 2.6: College vs. Four Years of Experience After High School

Model	High School (1)	High School (2)	HS at exper=4 vs College P-values (3)	HS at exper=4 vs College P-values (4)
Standard. AFQT	.0104	.0109	0.017	0.023
	(.0200)	(.0199)		
AFQT x exper/10	.1201**	.1126**	0.043	0.076
	(.0230)	(.0226)		
Black	-.1026**	-.0913**	0.001	0.001
	(.0414)	(.0407)		
Black x exper/10	.0074	.0093	0.099	0.097
	(.0466)	(.0459)		
R^2	0.1375	0.1642		
No. Observations	9236	9226		
Add. controls	No	Yes	No	Yes
Experience for HS grad: 4⩽years since left school< 13				

Note - We exlude first 4 years in the labor market for HS graduates. Specification (1) controls for urban residence, a cubic in experience and year effects. Specifications (2) also controls for region of residence and for part time vs full time jobs. In specification (5) we report the P-values for the difference in coefficients between high school at four years of experience, and college. Specification (6) makes the same comparison by controlling for additional variables as in specification (2). The White/Huber standard errors in parenthesis control for correlation at the individual level.

* statistical significance at the 90% level

** statistical significance at the 95% level

small and insignificant intercepts. Although the inclusion of AFQT decreases the magnitude and significance of the coefficients on father's education interaction with experience, this coefficient is still about four times the magnitude of the father's education base coefficient.

We now turn to specification (3), which analyses the effect of father's education on wages for college graduates. The coefficient on father's education is still statistically insignificant, but its magnitude of is quite sizable. The point estimate implies an 8 percent increase in earnings from a ten year increase in father's education. If we compare this to the analogous coefficient in specification (1), we can see that the returns to father's education are initially higher for college graduates than for high

Table 2.7: The Effects of AFQT and Father's education on Log Wages

	High School		College	
	(1)	(2)	(3)	(4)
Model:				
Black	-.0495*	-.0537*	.0339	.1298**
	(.0270)	(.0305)	(.0559)	(.0596)
Father's Education/10	.0361	.0392	.0819	.0575
	(.0386)	(.0402)	(.0648)	(.0646)
Standardized AFQT		-.0042		.1395**
		(.0150)		(.0375)
Black x experience/10	-.1568**	-.0241	-.1497**	-.1448**
	(.0355)	(.0402)	(.0668)	(.0727)
F. Educ/10 x experience/10	.1480**	.0780	-.0219	-.0370
	(.0538)	(.0530)	(.0978)	(.0995)
AFQT x experience/10		.1357**		.0271
		(.0201)		(.0501)
R^2	0.1345	0.1630	0.1446	0.1729
No. Observations	10034	10034	3983	3983
Experience measure: Years since left school for the first time				

Note - All specifications control for urban residence, a cubic in experience and year effects. Potential experience is limited to less than ten and thirteen years for the high school and the college sample respectively. The White/Huber standard errors in parenthesis control for correlation at the individual level.
* statistical significance at the 90% level
** statistical significance at the 95% level

school graduates. The coefficient on father's education times experience enters is negative, small, and insignificant. This last coefficient was large and significant for high school graduates. Similar results hold even after we include AFQT, although this decreases the immediate returns to father's education in the college market.

Taken together, although not statistically significant in all cases, the results for father's education are consistent with our main hypothesis that the ability of high school graduates is revealed gradually, while the ability of college graduates is more or less revealed directly upon entry into the labor market.[30]

2.4.4 *Controlling for Time Trends in Returns to Ability and in Racial Discrimination*

The NLSY79 sample we use contains mostly a single cohort of workers, so there could be concerns that our results reflect time trends in returns to ability and racial discrimination rather then experience effects. In order to address these concerns we include AFQT by year fixed effects and race by year fixed effects interactions to our baseline specifications. We pool all the data for college and high school graduates and interact everything with education dummies except for the AFQT by year fixed effects and race by year fixed effects which are held constant across education groups. The results from this procedure are presented in Table 2.8.

In specification (1) of Table 2.8 we include AFQT by year fixed effects interactions. The results are very similar to those presented in Table 2.2. The returns to AFQT are higher for college than for high school graduates and these returns increase faster for high school than for college graduates although this difference is slightly statistically insignificant. The coefficient on Black and Black times potential experience are also very similar to those in Table 2.2. In specification (2) we add Black by year fixed effects and find that results with respect to AFQT remain unchanged. The coefficient on Black is initially more negative and it decreases faster with experience although it is not statistically different from our previous results. To summarize, we find that time trends in returns to ability and in racial discrimination do not affect of our findings.

Table 2.8: Controlling for Time Changing Effects of Education, Race and AFQT

	(1)			(2)		
	H. School	College	Diff. P-val.	H. School	College	Diff. P-val.
Standard. AFQT	.0180	.1227**	0.003	.0210	.1177**	0.006
	(.0190)	(.0418)		(.0204)	(.0430)	
AFQT x exper/10	.1048**	.0134	0.106	.0935**	.0066	0.128
	(.0378)	(.0609)		(.0435)	(.0645)	
Black	-.0618**	.1097*	0.006	-.1511**	.1005	0.031
	(.0267)	(.0563)		(.0702)	(.0796)	
Black x exper/10	-.0370	-.1315*	0.223	-.0779	-.1563	0.359
	(.0350)	(.0693)		(.0767)	(.0989)	
AFQT x Year F.E.	Yes			Yes*		
Black x Year F.E.	No			Yes**		
R^2	0.3059			0.3065		
No. Observations	15907			15907		
Experience measure: Years since left school for the first time <13						

Note - For each specification we pool the data and estimate a model interacted with education fixed effects. Included but not shown in the table are education interactions with a cubic in experience, urban residence and with year effects. Specification (1) controls for the interaction AFQT x year fixed effects and specification (2) adds Black x year fixed effects. Both these interactions are not allowed to vary with education. The AFQT and Black coefficients are presented for the base year 1979. The White/Huber standard errors in parenthesis control for correlation at the individual level.
* statistical significance at the 90% level
** statistical significance at the 95% level

2.4.5 Analysis of the Wage Residual

If learning is occurring more in the high school market than in the college market, this may be true of abilities not captured by AFQT. If employers are learning about these other abilities, the autocorrelation in the wage regression should be increasing with experience until employers have learned everything they need to know. If our learning hypothesis is true, the autocorrelation in the wage residual should be increasing faster for high school graduates, where learning is important, than for college graduates where most of the learning has already taken place.

Figure 2.4 plots the autocorrelations of the wage residual normalized to zero in the first year of experience. The residual is constructed by estimating the wage equation (2.6) separately for each experience level. Using these residuals we calculate

[30] We also investigated whether similar patterns held for father's education in the PSID. Although not statistically significant, the qualitative findings matched those of Table 2.7.

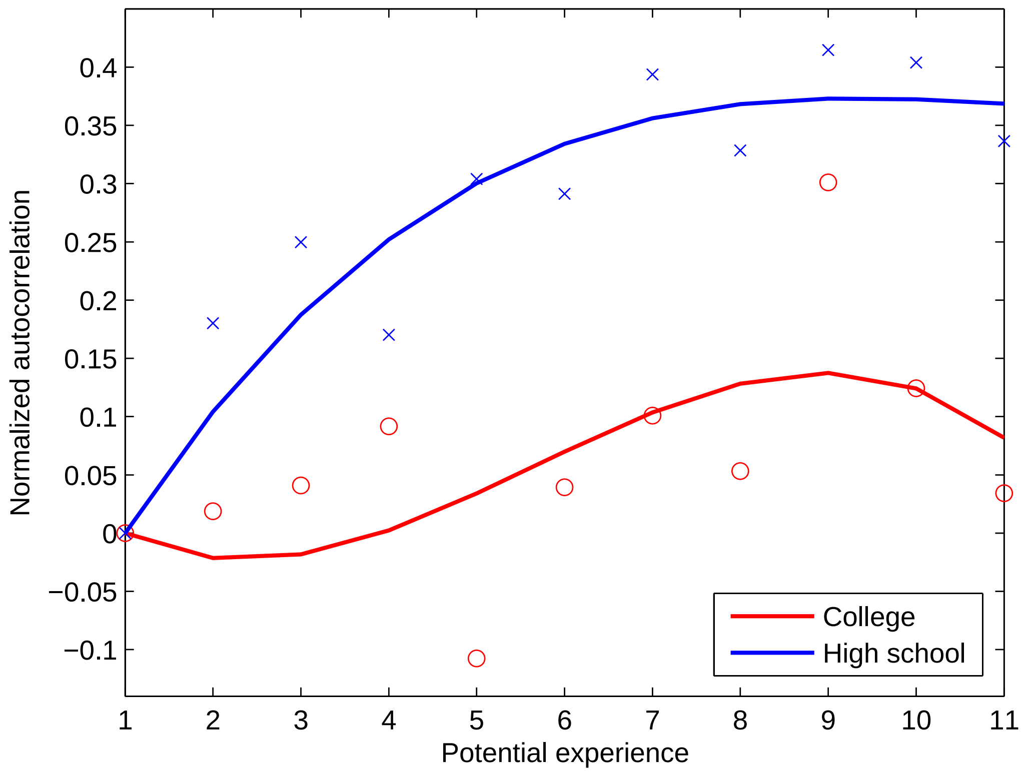

FIGURE 2.4: The Evolution of the Wage Residual Autocorrelation Over Potential Experience

Note: The wage residual is constructed by estimating the wage equation (2.6) separately for each experience level. We normalize the residual autocorrelations by its initial level and fit a cubic in potential experience to it.

the one year autocorrelations at the individual level. We normalize the estimated autocorrelation by its level in the first year of experience and fit a cubic in experience to it. The results confirm that the autocorrelation increases faster for the high school than for the college sample.[31] This finding is consistent with our learning story, although our hypothesis may not be the only driver of the observed patterns.

If learning is important, the weight put on ability should increase over time and this will increase the variance of the observed wages in the population since wages should reflect ability more with experience. We investigate whether the standard

[31] Because of the low number of individuals that have wage observations for two consecutive years we can not draw strong conclusions in terms of statistical significance.

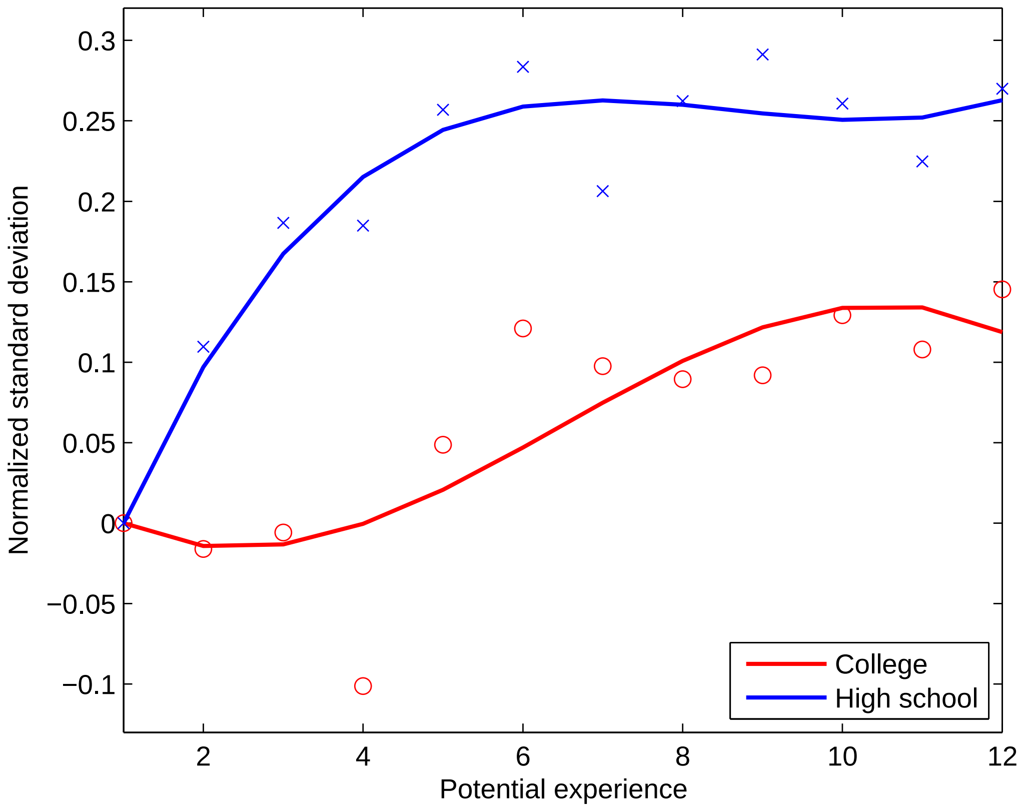

FIGURE 2.5: The Evolution of the Wage Residual Standard Deviation Over Potential Experience

Note: The wage residual is constructed by estimating the wage equation (2.6) separately for each experience level. We normalize the residual standard deviation by its initial level and fit a cubic in potential experience to it.

deviation of the wage residual grows faster for high school graduates than for college graduates as our hypothesis would predict. Figure 2.5 plots the wage residual standard deviation normalized by its initial level and fits a cubic to it. Consistent with our learning story, the standard deviation seems to increase faster for the high school sample than for college graduates.[32]

[32] These differences are statistically significant from zero.

2.5 Conclusion

The main argument in this paper is that education plays more than just a signaling role in the determination of wages. Specifically, we argue that graduation from college allows individuals to directly reveal their ability to potential employers. Using data from the NLSY, we show that the returns to AFQT, our measure of ability, are large for college graduates immediately upon entering the labor market and do not change significantly with labor market experience. In contrast, returns to AFQT for high school graduates are initially very close to zero and rise steeply with experience. These results suggest that ability is observed perfectly for college graduates but is revealed to the labor market more gradually for high school graduates.

Consistent with the notion that ability is nearly perfectly revealed, we find that, if anything, blacks earn more than whites in the college market. The lack of evidence of statistical discrimination in the college market is especially noteworthy given the large difference in the AFQT distribution for college-educated blacks and whites. On the other hand, we provide evidence that blacks earn six percent less than whites initially, and this gap increases with labor market experience in the high school market. We argue that this wage difference in the high school market may arise solely due to statistical discrimination given the information problem that potential employers face. Estimates of a model of employer learning and statistical discrimination are consistent with this explanation.

The combination of discrimination against blacks in the high school market and perfect revelation of ability in the college market is also consistent with the fact that, conditional on AFQT, blacks are more likely to earn a college degree than whites. Facing discrimination in the high school market, blacks on the college-high school margin have a stronger incentive to reveal their ability directly by attending college.

The amount of statistical discrimination that black workers face after high school may be reduced by devising some channel that allows blacks to better signal their ability to the market. One way to bridge the informational gap between workers and employers would be administering some form of an exit examination for high school graduates. Arguments for exit exams have been made before and there is some literature that analyzes and argues for such tests on the grounds that they provide a way for individuals to reveal ability to the labor market - see, for example, Bishop (2006), Bishop (2005), and Bishop and Mane (2001). Exit exams would give employers a clearer signal of workers ability and would reduce their incentives to statistically discriminate.

Appendix A

Appendix for Chapter 1

A.1 Proofs

A.1.1 Proof of Proposition 1

The problem at some time t faced by each household who dies at time T is:

$$maxV\left(X_t, w_t, x_t\right) = \underset{\theta_t, l}{sup} - E_t e^{-\gamma_i \left(X_T + \frac{1}{r}\left(e^{r(T-t)} - 1\right)\beta_i A^l\right)} \tag{A.1}$$

For notational simplicity we drop the i and l subscript from now on. Conditional on being in location l we solve for the value function and for the optimal portfolio decisions. The maximization problem is subject to the wealth evolution equation:

$$\begin{aligned} dX_t &= \theta_t' D_{P_t}^{-1} dP_t + \left(X_t - 1'\theta_t\right) rdt + w_t dt \\ &= \theta_t' \left(\mu dt + \Sigma_P dB_t\right) + \left(X_t - 1'\theta_t\right) rdt + w_t dt \end{aligned}$$

where θ_t is nx1.

The Hamilton-Jacobi-Bellman equation associated with this problem is [1]:

$$
\begin{aligned}
0 &= \dot{V} + V_X\left(\left(X_t - \theta_t' 1\right) r + w_t + \theta_t' \mu_t\right) + V_y s_t + V_s \phi\left(m - s_t\right) \\
&\quad + \frac{1}{2} V_{XX} \theta_t' \Sigma_P \Sigma_P' \theta + \frac{1}{2} V_{ss} \Sigma_s \Sigma_s' + V_{Xs} \theta_t' \Sigma_P \Sigma_x'
\end{aligned}
$$

Now we take FOC and solve for θ_t:

$$
\theta_t = -\frac{1}{V_{XX}} \Sigma_{PP}^{-1}\left[V_X\left(\mu_t - 1r\right) + V_{Xs}\Sigma_{Ps}\right]
$$

where for simplicity we we have defined $\Sigma_{YZ} = \Sigma_Y \Sigma_Z'$. Plugging back in and simplifying we get:

$$
\begin{aligned}
0 &= \dot{V} + V_X\left(Xr + w_t\right) + V_y s_t + V_s \phi\left(m - s_t\right) + \frac{1}{2} V_{ss} \Sigma_{ss} \\
&\quad - \frac{1}{2} \frac{V_X^2 \left(\mu_t' - 1'r\right) \Sigma_{PP}^{-1}\left(\mu_t - 1r\right)}{V_{XX}} \\
&\quad - \frac{1}{2} \frac{V_X V_{Xs}\left[\left(\mu_t' - 1'r\right)\Sigma_{PP}^{-1}\Sigma_{Ps} + \Sigma_{sP}\Sigma_{PP}^{-1}\left(\mu_t - 1r\right)\right]}{V_{XX}} \\
&\quad - \frac{1}{2} \frac{V_{Xs}^2 \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{V_{XX}}
\end{aligned}
$$

Now we can substitute the initial wealth out of the value function. Define

$$
\begin{aligned}
\bar{U} &= e^{r(T-t)} X_t + \int_t^T e^{r(T-u)}\left(\theta_u' D_{P_u}^{-1} dP_u - 1'\theta_u r du + w_u du\right) \\
&\quad + \frac{1}{r}\left(e^{r(T-t)} - 1\right)\beta_i M^l + p_T^l
\end{aligned}
$$

[1] Svensson and Werner (1993) derive a similar HJB for a related problem

Then the value function can be written as:

$$\begin{aligned} V\left(t, X_t, w_t, s_t\right) &= sup_{\theta_s} - E_t e^{\bar{U}} \\ &= -e^{-\gamma_i\left[X_t e^{r(T-t)} + \frac{1}{r}\left(e^{r(T-t)}-1\right)\left(\beta_i A^l + \xi_i + \varepsilon_i^l\right)\right]} g(T-t, y_t, s_t) \end{aligned}$$

with terminal condition $g(0, w, s) = e^{-\gamma[p_T]}$. Finally in the final simplified form:

$$\begin{aligned} 0 &= \dot{g} - g_s\left(\phi\left(m - s_t\right) - \frac{1}{2}\left[\left(\mu_t' - 1'r\right)\Sigma_{PP}^{-1}\Sigma_{sP} + \Sigma_{Ps}\Sigma_{PP}^{-1}\left(\mu_t - 1r\right)\right]\right) \\ &+ g\left(\gamma e^{r(T-t)} y_t + \frac{1}{2}\left(\mu_t' - 1'r\right)\Sigma_{PP}^{-1}\left(\mu_t - 1r\right)\right) \\ &- \frac{1}{2} g_{ss}\Sigma_{ss} - g_y s_t \\ &+ \frac{1}{2}\frac{g_s^2 \Sigma_{Ps}\Sigma_{PP}^{-1}\Sigma_{sp}}{g} \end{aligned}$$

Since this is nonlinear we need to do a change of variables to get it to be linear. Making the substitution:

$$g\left(T-t, y, s\right) = \upsilon\left(T-t, y, s\right)^{\frac{\Sigma_{ss}}{\Sigma_{ss} - \Sigma_{Ps}\Sigma_{PP}^{-1}\Sigma_{sP}}} e^{-\frac{1}{2}\left(\mu_t' - 1'r\right)\Sigma_{PP}^{-1}\left(\mu_t - 1r\right)(T-t)}$$

we get:

$$\begin{aligned} 0 &= \dot{\upsilon} + \gamma e^{r(T-t)} y_t \frac{\Sigma_{ss} - \Sigma_{Ps}\Sigma_{PP}^{-1}\Sigma_{sP}}{\Sigma_{ss}}\upsilon \\ &- \upsilon_s\left(\phi\left(m - s_t\right) - \frac{1}{2}\left[\left(\mu_t' - 1'r\right)\Sigma_{PP}^{-1}\Sigma_{Ps} + \Sigma_{sP}\Sigma_{PP}^{-1}\left(\mu_t - 1r\right)\right]\right) \\ &- \frac{1}{2}\upsilon_{ss}\Sigma_{ss} - \upsilon_y s_t \end{aligned}$$

Making a last substitution $G(t,y,s) = \upsilon(T-t,y,s)$ we get the linear pde:

$$\begin{aligned} 0 &= \dot{G} - \gamma e^{r(T-t)} y_t \frac{\Sigma_{ss} - \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{\Sigma_{ss}} G + \left(\phi(m - s_t) - \Sigma_{Ps}\Sigma_{PP}^{-1}(\mu_t - 1r)\right) G_s \\ &+ \frac{1}{2}\Sigma_s \Sigma_s^T G_{ss} + G_y s_t \end{aligned}$$

with terminal condition:

$$G(T,y,s) = e^{\frac{1}{2}(\mu_t' - 1'r)\Sigma_{PP}^{-1}(\mu_t - 1r)\frac{\Sigma_{ss} - \Sigma_{Ps}\Sigma_{PP}^{-1}\Sigma_{sP}}{\Sigma_{ss}}(T-t) - \gamma \frac{\Sigma_{ss} - \Sigma_{Ps}\Sigma_{PP}^{-1}\Sigma_{sP}}{\Sigma_{ss}} p_T}$$

Now we do a change of measure by multiplying the probability density function by:

$$\xi_T = exp\left[-\frac{1}{2}\left[\frac{\Sigma_{sP}\Sigma_{PP}^{-1}(\mu_t - 1r)}{\sqrt{\Sigma_{ss}}}\right]^2 (T-t) - \left[\frac{\Sigma_{sP}\Sigma_{PP}^{-1}(\mu_t - 1r)}{\sqrt{\Sigma_{ss}}}\right]\frac{\Sigma_s}{\sqrt{\Sigma_{ss}}}(B_T - B_t)\right]$$

Then we can use the Feynman Kac theorem to get:[2]

$$G(t,y,s) = E\left[\xi_T G(T,y,s) e^{\left[-\gamma \frac{\Sigma_{ss} - \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{\Sigma_{ss}} \int_t^T e^{r(T-u)} y_u du\right]}\right]$$

Now we can finally write the value function in a form that does not involve the portfolio decision variable θ:θ:

$$\begin{aligned} V(t,y,s) &= -e^{-\gamma\left[X_t e^{r(T-t)} + \frac{1}{r}\left(e^{rT-t} - 1\right)\left(\beta_i A^l + \xi_i + \varepsilon_i^l\right)\right] - \frac{1}{2}(\mu_t' - 1'r)\Sigma_{PP}^{-1}(\mu_t - 1r)(T-t)} \\ &\cdot [G(t,y,s)]^{\overline{\frac{\Sigma_{ss}}{\Sigma_{ss} - \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}}} \end{aligned}$$

In order to save some notation:

[2] See Duffie (1996), Theorem and Condition 2 on p. 296

$$V\left(t,y,s\right)=-e^{\bar{F}}\left[E_t e^{F_T}\right]^{\frac{\Sigma_{ss}}{\Sigma_{ss}-\Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}}$$

where:

$$\begin{aligned}\overline{F} &= -\frac{1}{2}\left(\mu_t'-1'r\right)\Sigma_{PP}^{-1}\left(\mu_t-1r\right)\left(T-t\right)-\frac{1}{2}\frac{\left(\Sigma_{sP}\Sigma_{PP}^{-1}\left(\mu_t-1r\right)\right)^2}{\left(\Sigma_{ss}-\Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}\right)}\left(T-t\right)\\ &\quad -\gamma\left[X_t e^{r(T-t)}+\frac{1}{r}\left(e^{rT-t}-1\right)\left(\beta_i A^l+\xi_i+\varepsilon_i^l\right)\right]\end{aligned}$$

$$\begin{aligned}F_T &= -\left[\frac{\Sigma_{sP}\Sigma_{PP}^{-1}\left(\mu_t-1r\right)}{\sqrt{\Sigma_{ss}}}\right]\frac{\Sigma_s}{\sqrt{\Sigma_{ss}}}\left(B_T-B_t\right)\\ &\quad -\gamma_i\frac{\Sigma_{ss}-\Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{\Sigma_{ss}}\left(\int_t^T e^{r(T-u)}y_u du+Ay_T+Bs_T+C\right)\end{aligned}$$

Notice that $\overline{F}$ is just a constant while the term F_T depends on the states. Using the assumed processes for y_t and s_t we can write F_T in terms of states at time t. First we make a transformation that transforms the m independent Brownian motions $\mathbb{B}_t$ to a single brownian motion Z_t for each location. To do this we redefine:

$$dZ_t=\frac{\Sigma_s}{\sqrt{\Sigma_{ss}}}d\mathbb{B}_t$$

which means that now we can write the evolution of the drift s_t of the income process as:

$$\begin{aligned}ds_t &= \phi\left(m-s_t\right)dt+\Sigma_s dB_t\\ &= \phi\left(m-s_t\right)dt+\kappa dZ_t\end{aligned}$$

with $\kappa = \sqrt{\Sigma_{ss}}$. This means that s_t is just a simple one dimensional Ornstein-Uhlenbeck process.

We can use the solutions to the processes y_t and s_t to write the components of F_T as:

$$s_T = s_t e^{-\phi(T-t)} + m\left(1 - e^{-\phi(T-t)}\right) + \int_t^T \kappa e^{\phi(u-T)} dZ_u$$

$$\begin{aligned} y_T &= y_t + \int_t^T s_u du \\ &= y_t + s_t \frac{1}{\phi}\left(1 - e^{-\phi(T-t)}\right) + m\left((T-t) - \frac{1}{\phi}\left(1 - e^{-\phi(T-t)}\right)\right) \\ &\quad + \int_t^T \frac{\kappa}{\phi}\left(1 - e^{-\phi(T-s)}\right) dZ_s \end{aligned}$$

$$\begin{aligned} \int_t^T e^{r(T-u)} y_u du &= -y_t \frac{1}{r}\left(1 - e^{r(T-t)}\right) \\ &\quad + s_t \frac{1}{\phi}\left(-\frac{1}{r}\left(1 - e^{r(T-t)}\right) + \frac{1}{r+\phi}\left(e^{-\phi(T-t)} - e^{r(T-t)}\right)\right) \\ &\quad - m\frac{1}{r^2}\left(rT + 1 - e^{r(T-t)}(rt+1)\right) + mt\frac{1}{r}\left(1 - e^{r(T-t)}\right) \\ &\quad - \frac{m}{\phi}\left(-\frac{1}{r}\left(1 - e^{r(T-t)}\right) + \frac{1}{r+\phi}\left(e^{-\phi(T-t)} - e^{r(T-t)}\right)\right) \\ &\quad \int_t^T \frac{\kappa}{\phi}\left[-\frac{1}{r}\left(1 - e^{r(T-s)}\right) + \frac{1}{r+\phi}\left(e^{-\phi(T-s)} - e^{r(T-s)}\right)\right] dZ_t \end{aligned}$$

Rewriting F_T in terms of states at time t we get:

$$F_T = -\gamma_i\left(\tilde{k}_{1t} y_t + \tilde{k}_{2t} s_t + \tilde{k}_{3t}\right) + \int_t^T \left(\tilde{k}_4 - \gamma_i \sqrt{\Sigma_{ss}} \tilde{k}_{2t}\right) dZ_u$$

where $dZ_t = \Sigma_s / \sqrt{\Sigma_{ss}} dB_t$ and:

$$\tilde{k}_{1t} = \frac{\Sigma_{ss} - \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{\Sigma_{ss}} \left(A - \frac{1 - e^{r(T-t)}}{r} \right)$$

$$\begin{aligned} \tilde{k}_{2t} &= \frac{\Sigma_{ss} - \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{\Sigma_{ss}} \left\{ A \frac{1 - e^{-\phi(T-t)}}{\phi} + B e^{-\phi(T-t)} \right. \\ & \left. - \frac{1 - e^{r(T-t)}}{\phi r} + \frac{e^{-\phi(T-t)} - e^{r(T-t)}}{\phi (r + \phi)} \right\} \end{aligned}$$

$$\begin{aligned} \tilde{k}_{3t} &= -m\tilde{k}_2 + \frac{\Sigma_{ss} - \Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}{\Sigma_{ss}} \left\{ Am (T - t) + Bm + \frac{mt \left(1 - e^{r(T-t)} \right)}{r} \right. \\ & \left. - m \frac{1}{r^2} \left(rT + 1 - e^{r(T-t)} (rt + 1) \right) + C \right\} \end{aligned}$$

$$\tilde{k}_{4t} = -\frac{\Sigma_{sP}\Sigma_{PP}^{-1} (\mu - 1r)}{\sqrt{\Sigma_{ss}}}$$

The only random part of F_T now $\int_t^T \left(\tilde{k}_4 - \gamma_i \sqrt{\Sigma_{ss}} \tilde{k}_{2t} \right) dZ_t$, which is an Ito integral over a deterministic function and thus a martingale. Therefore, the distribution of F_T is a normal with mean and variance:[3]

$$E_t (F_T) = -\gamma_i \left(\tilde{k}_{1t} y_t + \tilde{k}_{2t} s_t + \tilde{k}_{3t} \right)$$

$$Var (F_T) = \int_t^T \left(\tilde{k}_4 - \gamma_i \sqrt{\Sigma_{ss}} \tilde{k}_{2t} \right)^2 dt$$

Using the fact that F_T is normally distributed we can write the value function as:

[3] See Shreve (2004) chapter 4.

$$V(t,y,s) = -e^{\overline{F}}\left[e^{E(F_T)+\frac{1}{2}Var(F_T)}\right]^{\frac{\Sigma_{ss}}{\Sigma_{ss}-\Sigma_{sP}\Sigma_{PP}^{-1}\Sigma_{Ps}}}$$

Simplifying the above equation and including the location superscript l, the value function is:

$$V(t,y,s,l) = -e^{U_t^l}$$

with*:*

$$\begin{aligned} U_t^l &= -\gamma_i\left[e^{r(T-t)}X_t + \frac{1}{r}\left(e^{r(T-t)}-1\right)\left(\beta_i M^l + \xi_i + \varepsilon_i^l\right)\right] \\ &\quad -\gamma_i\left[\hat{k}_1^l y_t^l + \hat{k}_2^l s_t^l + \hat{k}_3^l + \int_t^T \hat{k}_4^l \hat{k}_2^l \sqrt{\Sigma_{ss}^l} du\right] \\ &\quad +\frac{1}{2}\gamma_i^2\left(\Sigma_{ss}^l - \Sigma_{sP}^l\Sigma_{PP}^{-1}\Sigma_{Ps}^l\right)\int_t^T\left(\hat{k}_{t2}^l\right)^2 du - \frac{1}{2}\left(\mu_t' - 1'r\right)\Sigma_{PP}^{-1}\left(\mu_t - 1r\right)(T-t) \end{aligned}$$

where we define $\hat{k}_{it}^l = \frac{\Sigma_{ss}^l - \Sigma_{sP}^l\Sigma_{PP}^{-1}\Sigma_{Ps}^l}{\Sigma_{ss}^l}\tilde{k}_{it}^l$ for $i = 1...3$ and $\hat{k}_4^l = \tilde{k}_4^l$. ■

A.1.2 Proof of Proposition 2

For clarity we suppress the city-specific l superscript. In the proof of Proposition 1, I show that the optimal amount invested in stocks for an individual who is $t - t_0$ years old and lives in location l is:

$$\theta_{ilt} = -\frac{1}{V_{XX}}\left(\Sigma_{PP}\right)^{-1}\left[V_X\left(\mu - 1r\right) + V_{Xs}\Sigma_{Ps}^l\right]$$

Taking the derivatives of the value function equation (1.4):

$$V\left(t,y,s,l\right)=-e^{U_t^l}$$

$$V_X=-\gamma e^{r(T-t)}e^{U_t^l}$$

$$V_{XX}=\gamma^2 e^{2r(T-t)}e^{U_t^l}$$

$$V_{Xs}=\gamma^2\hat{k}_2^l e^{r(T-t)}e^{U_t^l}$$

Substituting these derivatives in the equation for θ_{ilt} we get:

$$\theta_{ilt}=\left(\Sigma_{PP}\right)^{-1}\left[\frac{1}{\gamma e^{r(T-t)}}\left(\mu_t-1r\right)+\frac{\hat{k}_{2t}^l}{e^{r(T-t)}}\Sigma_{Ps}^l\right]$$

where:

$$\begin{aligned}\hat{k}_{2t} &= \left\{A\frac{1-e^{-\phi(T-t)}}{\phi}+Be^{-\phi(T-t)}\right.\\ &\quad \left.-\frac{1-e^{r(T-t)}}{\phi r}+\frac{e^{-\phi(T-t)}-e^{r(T-t)}}{\phi\left(r+\phi\right)}\right\}\end{aligned}$$

■

A.1.3 Proof of Proposition 6

For a given A and B, in each period we look for a C^l that clears the housing markets in some given period t. For the generation of agents born in some period t, the problem in equation 1.7 is identical to static horizontal sorting models studied in the urban economics literature. Bayer, McMillan and Rueben (2005) prove that under the assumption that ε_i^l has continuous support there exist a unique vector $\mathbf{C} = \{C^1, C^2 ... C^L\}$ that clears the market and gives the unique sorting of households across space. Their proof applies exactly to this problem so is not reproduced here. The reader is referred to the original article for details.

Given that we now have unique values for C^l, we turn to determining what A and B need to be in order for the markets to be in equilibrium across time. First note that a measure 1 of households is born and looking to buy a home each period. Because the joint distribution of heterogeneity in $Pr\left(\gamma_i, \beta_i, \xi_i, \varepsilon_i^l\right)$ is iid over time, the same set of households will be in the market in every time period. In other words in every period the full support of the joint distribution $Pr\left(\gamma_i, \beta_i, \xi_i, \varepsilon_i^l\right)$ will be realized.

Consider the sorting problem in equation 1.7 at two different time periods t and u where $y_t \neq y_u$ and $s_t = s_u$. If $k_1 \neq 0$ then the equilibrium $\mathbf{C}$ of the sorting equilibrium in period t will have to be different than that in period u. Because we are looking for a vector $\mathbf{C}$ that is constant across time we conclude that it must be the case that $k_1 = 0$ for every location l. A similar argument leads to the conclusion that in order to have $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ to be constant across time it must be that $k_2^l = 0$ for every location l. Setting $k_1^l = k_2^l = 0$ we find the unique solutions be:

$$A^l = \frac{1}{r}$$

$$B^l = \frac{1}{r\left(r + \phi^l\right)}$$

Given the above solutions $\mathbf{A}$, and $\mathbf{B}$ we can now get the unique equilibrium market returns for stocks in equation 1.6. We have therefore found the three vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ that give rise to equilibrium both in the asset market and the housing market. ■

A.1.4 Proof of Proposition 7

The location l superscript will be ommited for the rest of the proof for notational simplicity, and this analysis is the same for any location l. The idea here is to use the solution to the process for y_t to solve for the expectation in closed-form. Recall that the processes that drive income are:

$$\begin{aligned} dy_t &= s_t dt \\ ds_t &= \phi\left(m - s_t\right)dt + \kappa dZ_t \end{aligned}$$

where $dZ_t = \Sigma_s/\sqrt{\Sigma_{ss}}dB_t$ and $\kappa = \sqrt{\Sigma_{ss}}$. This change of variables transforms the s_t process to a simple one dimensional Ornstein-Uhlenbeck process. The solution for this process for some $T > t$ is:

$$s_T = s_t e^{-\phi(T-t)} + m\left(1 - e^{-\phi(T-t)}\right) + \int_t^T \kappa e^{\phi(v-T)} dZ_v$$

For any $T > t$ we can write y_T as:

$$\begin{aligned} y_T &= y_t + \int_t^T s_u du \\ &= y_t + \int_t^T \left(s_t e^{-\phi(u-t)} + m\left(1 - e^{-\phi(u-t)}\right) + \int_t^u \kappa e^{\phi(v-u)} dZ_v\right) du \\ &= y_t + \int_t^T \left(s_t e^{-\phi(u-t)} + m\left(1 - e^{-\phi(u-t)}\right)\right) du + \int_t^T \left(\int_v^T \kappa e^{\phi(v-u)} du\right) dZ_v \\ &= y_t + s_t \frac{1}{\phi}\left(1 - e^{-\phi(T-t)}\right) + m\left((T-t) - \frac{1}{\phi}\left(1 - e^{-\phi(T-t)}\right)\right) \\ &\quad + \int_t^T \frac{\kappa}{\phi}\left(1 - e^{-\phi(T-v)}\right) dZ_v \end{aligned}$$

where in the second equality we substitute in for the solution to s_u, in the third equality we changed the order of integration for the last term, and in the last equality we solve the deterministic integrals.

We can now substitute the solution to y directly in the price equation:

$$
\begin{aligned}
p_t &= E_t \int_t^{\infty} e^{-r(u-t)} y_u du + \pi \\
&= \int_t^{\infty} e^{-r(u-t)} E_t \left(y_u \right) du + \pi \\
&= \int_t^{\infty} e^{-r(u-t)} \left(y_t + s_t \frac{1}{\lambda} \left(1 - e^{-\lambda(u-t)} \right) + m \left(u - t - \frac{1}{\lambda} \left(1 - e^{-\lambda(u-t)} \right) \right) \right) du + \pi \\
&= \frac{1}{r} y_t + \frac{1}{r(r+\phi)} s_t + \frac{\phi}{r^2 (r+\phi)} m + \pi
\end{aligned}
$$

In the second equality we interchange the integral with the expectation.[4] In the third equality we substitute for the expected value of y_u. Here we use the fact that $E\left(\int_t^u \frac{\kappa}{\phi}\left(1 - e^{-\phi(u-v)}\right) dZ_v\right) = 0$ since Ito integrals of deterministic functions are martingales. ■

[4] In order to interchange the integral and the expectation, first notice that the expectation is an integral over some probability measure. We can then use Fubini's theorem that says that the order of integration can be changed as long as $\int_t^{\infty} e^{-r(u-t)} E_t \left(\left| \int_t^u \frac{\kappa}{\phi} \left(1 - e^{-\phi(u-v)}\right) dZ_v \right| \right) du < \infty$. This condition is satisfied for our problem. The intuition is that the expectation of the absolute value of the martingale grows slower that the discounting term $e^{-r(u-t)}$. Broadly speaking the expectation grows linearly over time while the discounting term grows exponentially.

A.2 Data Appendix

Table A.1: The Factor Loadings and the R-squared

Metropolitan Area	HMKT	SMBH	HMLH	R-squared
Akron,OH	3.059	-0.174	0.341	0.107
Albany-Schenectady-Troy,NY	4.554	-18.886	2.885	0.588
Albuquerque,NM	11.580	-1.059	16.847	0.470
Allentown-Bethlehem-Easton,PA-NJ	3.490	-24.722	0.453	0.583
Amarillo,TX	4.884	2.958	0.661	0.208
Anchorage,AK	18.350	3.117	10.947	0.209
AnnArbor,MI	6.879	-8.313	-2.644	0.371
Atlanta-SandySprings-Marietta,GA	7.654	-4.361	-2.651	0.596
Augusta-RichmondCounty,GA-SC	3.913	-0.744	3.810	0.360
Austin-RoundRock,TX	7.783	8.152	3.007	0.131
Bakersfield,CA	11.340	-21.111	22.663	0.872
Baltimore-Towson,MD	7.352	-22.227	16.914	0.846
BarnstableTown,MA	25.493	-41.569	-9.980	0.767
BatonRouge,LA	8.048	6.369	8.294	0.538
Beaumont-PortArthur,TX	6.983	6.474	4.988	0.596
Bellingham,WA	-1.084	-9.358	27.285	0.780
Bethesda-Frederick-Rockville,MD	11.745	-45.196	22.672	0.852
Binghamton,NY	4.753	-7.221	-1.851	0.225
Birmingham-Hoover,AL	4.780	0.441	2.174	0.370
Bloomington-Normal,IL	4.266	0.564	2.004	0.236
BoiseCity-Nampa,ID	3.750	-0.033	14.016	0.506
Boston-Quincy,MA	23.981	-36.247	-13.300	0.695
Boulder,CO	14.178	6.945	-2.888	0.245
Bremerton-Silverdale,WA	4.590	-8.663	22.584	0.672
Bridgeport-Stamford-Norwalk,CT	34.996	-49.136	-8.938	0.600
Buffalo-NiagaraFalls,NY	0.267	-3.779	0.404	0.105
Burlington-SouthBurlington,VT	5.099	-21.748	3.842	0.804
Cambridge-Newton-Framingham,MA	25.008	-37.554	-18.296	0.658
Camden,NJ	4.685	-19.207	6.015	0.718
Canton-Massillon,OH	3.639	2.355	-0.821	0.286
CapeCoral-FortMyers,FL	7.485	-15.013	17.192	0.752
Casper,WY	4.515	9.585	13.116	0.285
CedarRapids,IA	3.024	1.403	1.477	0.213
Charleston-NorthCharleston-Summerville,SC	10.308	-7.175	8.320	0.489
Charlotte-Gastonia-Concord,NC-SC	0.862	0.501	-1.556	0.027
Charlottesville,VA	6.853	-15.302	14.728	0.741
Chattanooga,TN-GA	4.609	-1.685	1.836	0.435
Cheyenne,WY	6.160	0.039	4.876	0.271
Chicago-Naperville-Joliet,IL	5.863	-12.193	4.882	0.783
Chico,CA	12.545	-22.399	25.952	0.785
Cincinnati-Middletown,OH-KY-IN	4.171	-2.410	-0.813	0.465
Cleveland-Elyria-Mentor,OH	5.112	0.422	-0.207	0.282
CollegeStation-Bryan,TX	8.952	5.730	4.084	0.392
ColoradoSprings,CO	11.970	1.701	6.523	0.634
Columbia,SC	2.965	-0.983	0.568	0.241
Columbus,OH	3.280	-1.874	-0.808	0.319
CorpusChristi,TX	10.986	2.339	7.058	0.576
Corvallis,OR	1.374	5.700	21.990	0.670
Dallas-Plano-Irving,TX	7.948	2.233	0.370	0.233
Dalton,GA	2.745	-1.786	1.674	0.130
Davenport-Moline-RockIsland,IA-IL	2.765	3.424	3.542	0.292
Dayton,OH	2.790	-0.443	-0.885	0.133
Deltona-DaytonaBeach-OrmondBeach,FL	9.910	-16.977	17.566	0.857
Denver-Aurora-Broomfield,CO	15.812	4.771	-2.136	0.437

Table Continued: The Factor Loadings and the R-squared

Metropolitan Area	HMKT	SMBH	HMLH	R-squared
DesMoines-WestDesMoines,IA	2.158	0.827	1.303	0.147
Detroit-Livonia-Dearborn,MI	2.960	-2.075	-0.316	0.205
Durham-ChapelHill,NC	3.903	-2.138	-1.892	0.177
EauClaire,WI	2.270	-1.993	-0.148	0.092
Edison-NewBrunswick,NJ	22.317	-43.143	0.507	0.773
ElPaso,TX	2.602	0.270	5.954	0.230
Elkhart-Goshen,IN	1.795	0.807	1.221	0.047
Erie,PA	3.309	-0.228	0.566	0.211
Eugene-Springfield,OR	0.533	-2.695	16.770	0.759
Evansville,IN-KY	4.108	1.509	0.606	0.401
Fayetteville-Springdale-Rogers,AR-MO	3.640	-4.286	6.666	0.488
Flint,MI	6.336	-3.552	-0.652	0.238
FortCollins-Loveland,CO	10.549	2.618	0.364	0.316
FortLauderdale-PompanoBeach-DeerfieldBeach,FL	17.287	-28.214	27.252	0.837
FortWayne,IN	2.076	-1.141	-0.763	0.124
Fresno,CA	12.855	-23.551	25.696	0.805
Gary,IN	1.152	1.936	4.186	0.416
GrandJunction,CO	7.009	6.067	13.060	0.384
GrandRapids-Wyoming,MI	1.467	-1.924	-1.448	0.289
Greeley,CO	12.939	1.786	-0.790	0.282
Greensboro-HighPoint,NC	1.948	-1.273	-3.363	0.311
Harrisburg-Carlisle,PA	3.728	-3.242	5.761	0.506
Hartford-WestHartford-EastHartford,CT	9.397	-32.412	-13.102	0.667
Holland-GrandHaven,MI	1.718	-2.755	-3.192	0.172
Honolulu,HI	-18.653	-51.363	60.250	0.738
Houston-SugarLand-Baytown,TX	8.153	5.778	4.499	0.391
Huntsville,AL	4.965	-0.057	0.752	0.187
Indianapolis-Carmel,IN	1.691	-0.066	-0.672	0.102
Jackson,MS	4.559	-0.186	2.523	0.244
Jacksonville,FL	9.029	-11.038	11.407	0.698
Janesville,WI	1.128	1.005	4.419	0.146
Kalamazoo-Portage,MI	4.265	0.214	1.585	0.204
KansasCity,MO-KS	5.302	-3.242	-1.073	0.548
Kennewick-Pasco-Richland,WA	0.068	2.742	7.110	0.199
Knoxville,TN	4.363	-0.468	4.608	0.515
LaCrosse,WI-MN	4.909	0.216	3.743	0.375
Lafayette,LA	5.760	9.518	9.392	0.690
LakeCounty-KenoshaCounty,IL-WI	2.766	-9.366	0.953	0.458
Lakeland-WinterHaven,FL	7.457	-9.498	16.365	0.564
Lancaster,PA	0.167	-7.425	6.818	0.549
Lansing-EastLansing,MI	4.135	-3.783	-0.958	0.338
LasCruces,NM	5.886	-2.965	8.585	0.641
LasVegas-Paradise,NV	11.720	-20.360	22.298	0.865
Lexington-Fayette,KY	4.485	-2.457	0.756	0.343
Lima,OH	2.825	-1.018	0.448	0.140
Lincoln,NE	2.833	0.498	1.926	0.260
LittleRock-NorthLittleRock-Conway,AR	4.994	0.834	1.884	0.388
Longview,TX	5.003	3.210	4.121	0.539
Longview,WA	2.226	4.943	16.713	0.660
LosAngeles-LongBeach-Glendale,CA	14.561	-54.850	29.291	0.848
Louisville-JeffersonCounty,KY-IN	1.889	0.305	0.683	0.187
Lubbock,TX	5.479	2.047	2.525	0.396
Macon,GA	3.188	-1.321	0.838	0.227
Madera-Chowchilla,CA	11.664	-29.035	35.194	0.813
Madison,WI	3.246	-2.106	5.315	0.210
Manchester-Nashua,NH	15.200	-28.583	-5.439	0.665
Mansfield,OH	3.021	1.412	1.429	0.171
Medford,OR	11.435	-19.949	26.830	0.849
Memphis,TN-MS-AR	0.303	-3.176	-2.724	0.196

Table Continued: The Factor Loadings and the R-squared

Metropolitan Area	HMKT	SMBH	HMLH	R-squared
Merced,CA	10.050	-27.243	21.825	0.737
Miami-MiamiBeach-Kendall,FL	13.639	-18.152	26.455	0.765
Midland,TX	9.405	13.731	8.258	0.433
Milwaukee-Waukesha-WestAllis,WI	5.457	-5.180	6.919	0.558
Minneapolis-St.Paul-Bloomington,MN-WI	9.183	-8.632	2.726	0.575
Mobile,AL	6.034	3.625	4.347	0.185
Modesto,CA	13.173	-34.490	24.795	0.804
Monroe,LA	5.386	3.338	2.515	0.328
Monroe,MI	3.564	-3.929	0.131	0.164
Napa,CA	20.727	-52.064	29.391	0.738
Naples-MarcoIsland,FL	16.584	-33.255	32.508	0.760
Nashville-Davidson-Murfreesboro-Franklin,TN	4.070	-1.678	1.498	0.164
Nassau-Suffolk,NY	25.344	-47.593	-3.155	0.830
NewHaven-Milford,CT	17.303	-31.329	-5.593	0.657
NewOrleans-Metairie-Kenner,LA	9.893	6.286	1.545	0.507
NewYork-WhitePlains-Wayne,NY-NJ	22.818	-45.541	-2.468	0.815
Newark-Union,NJ-PA	24.421	-44.044	-2.095	0.750
Niles-BentonHarbor,MI	1.460	-3.758	0.369	0.310
Oakland-Fremont-Hayward,CA	17.355	-53.726	23.873	0.796
Odessa,TX	2.770	6.388	4.490	0.372
Ogden-Clearfield,UT	6.929	11.549	8.165	0.512
OklahomaCity,OK	8.301	7.904	8.067	0.721
Olympia,WA	2.077	-3.902	22.664	0.741
Omaha-CouncilBluffs,NE-IA	3.561	1.334	0.542	0.394
Orlando-Kissimmee,FL	10.175	-17.405	21.053	0.786
Oxnard-ThousandOaks-Ventura,CA	24.815	-76.045	22.819	0.819
PalmBay-Melbourne-Titusville,FL	8.127	-13.584	14.058	0.864
Peabody,MA	24.809	-33.914	-13.787	0.665
Pensacola-FerryPass-Brent,FL	9.372	-11.313	13.351	0.795
Peoria,IL	1.594	2.570	0.530	0.149
Philadelphia,PA	3.385	-16.903	2.694	0.690
Phoenix-Mesa-Scottsdale,AZ	7.923	-14.578	16.295	0.680
Pittsburgh,PA	3.354	0.944	0.569	0.197
PortSt.Lucie,FL	12.800	-20.766	19.446	0.844
Portland-SouthPortland-Biddeford,ME	9.980	-21.924	-3.937	0.810
Portland-Vancouver-Beaverton,OR-WA	6.510	-1.059	22.181	0.700
Poughkeepsie-Newburgh-Middletown,NY	20.442	-39.836	-4.061	0.747
Providence-NewBedford-FallRiver,RI-MA	13.802	-31.162	-1.481	0.806
Provo-Orem,UT	6.015	11.759	12.309	0.439
Pueblo,CO	8.399	0.135	1.133	0.377
Racine,WI	4.853	-6.015	6.549	0.606
Raleigh-Cary,NC	3.576	0.891	0.982	0.070
Reading,PA	0.805	-10.277	4.342	0.514
Redding,CA	8.762	-21.875	23.831	0.734
Reno-Sparks,NV	12.459	-23.911	23.915	0.774
Richmond,VA	4.555	-9.201	9.631	0.644
Riverside-SanBernardino-Ontario,CA	12.173	-31.933	25.135	0.865
Roanoke,VA	2.537	-2.519	4.227	0.403
Rochester,MN	4.972	-3.118	-1.308	0.238
Rochester,NY	3.017	-4.286	-3.217	0.282
Rockford,IL	1.174	-0.526	4.274	0.394
RockinghamCounty-StraffordCounty,NH	17.781	-27.745	-8.645	0.712
Sacramento-Arden-Arcade-Roseville,CA	11.827	-38.269	23.842	0.688
Saginaw-SaginawTownshipNorth,MI	3.127	-1.026	1.013	0.264
Salem,OR	4.198	3.304	14.215	0.697
Salinas,CA	17.464	-47.843	23.966	0.783
SaltLakeCity,UT	5.335	10.671	13.153	0.530
SanAntonio,TX	10.301	5.155	8.380	0.456
SanDiego-Carlsbad-SanMarcos,CA	21.814	-55.869	22.163	0.727

Table Continued: The Factor Loadings and the R-squared

Metropolitan Area	HMKT	SMBH	HMLH	R-squared
SanFrancisco-SanMateo-RedwoodCity,CA	19.466	-90.385	12.650	0.744
SanJose-Sunnyvale-SantaClara,CA	12.929	-63.083	18.547	0.536
SanLuisObispo-PasoRobles,CA	14.438	-56.354	28.803	0.709
SantaAna-Anaheim-Irvine,CA	28.367	-82.096	38.745	0.837
SantaBarbara-SantaMaria-Goleta,CA	20.126	-50.303	17.341	0.748
SantaCruz-Watsonville,CA	13.819	-65.713	18.778	0.687
SantaFe,NM	10.337	1.302	17.763	0.606
SantaRosa-Petaluma,CA	14.441	-51.070	20.683	0.693
Savannah,GA	7.408	-3.703	7.944	0.648
Scranton-Wilkes-Barre,PA	3.243	-1.967	2.545	0.106
Seattle-Bellevue-Everett,WA	0.932	-14.413	26.672	0.566
Sebastian-VeroBeach,FL	15.148	-17.709	18.043	0.737
Shreveport-BossierCity,LA	7.884	2.726	5.785	0.382
SouthBend-Mishawaka,IN-MI	-0.108	-0.415	0.439	0.026
Spokane,WA	3.372	0.469	15.803	0.639
Springfield,IL	0.849	-0.990	3.784	0.168
Springfield,MA	7.316	-25.214	-4.894	0.772
Springfield,MO	3.873	0.007	3.867	0.394
Springfield,OH	2.514	0.330	-1.375	0.157
St.Louis,MO-IL	4.136	-5.360	0.628	0.666
Stockton,CA	16.603	-41.185	25.143	0.792
Syracuse,NY	2.235	-8.397	-1.345	0.392
Tacoma,WA	5.135	-6.659	19.555	0.628
Tallahassee,FL	4.799	-5.175	11.619	0.714
Tampa-St.Petersburg-Clearwater,FL	11.015	-16.470	16.055	0.853
Toledo,OH	3.403	-1.024	-0.393	0.239
Topeka,KS	4.245	-0.497	3.057	0.477
Trenton-Ewing,NJ	15.577	-33.589	-2.730	0.714
Tucson,AZ	8.111	-12.010	15.598	0.709
Tulsa,OK	8.037	7.470	1.141	0.700
Tyler,TX	5.397	2.634	0.398	0.132
Vallejo-Fairfield,CA	17.277	-37.947	23.233	0.780
VirginiaBeach-Norfolk-NewportNews,VA-NC	5.923	-16.568	14.717	0.788
Visalia-Porterville,CA	7.391	-22.658	24.832	0.825
Warren-Troy-FarmingtonHills,MI	5.288	-4.561	-1.140	0.272
Washington-Arlington-Alexandria,DC-VA-MD-WV	16.022	-47.490	23.720	0.904
Waterloo-CedarFalls,IA	2.769	3.148	5.393	0.181
Wenatchee-EastWenatchee,WA	-1.048	6.139	15.622	0.327
WestPalmBeach-BocaRaton-BoyntonBeach,FL	19.420	-32.849	26.344	0.865
Wilmington,DE-MD-NJ	4.371	-20.335	9.114	0.810
Wilmington,NC	6.809	-5.212	14.656	0.567
Winston-Salem,NC	4.749	-0.863	0.574	0.311
Worcester,MA	17.265	-28.101	-7.750	0.720
York-Hanover,PA	2.545	-9.258	8.270	0.525

Table A.2: The Estimated Risk Premia and Implied Amenities for 2008

Metropolitan Area	Price	Wage	Amenities	Risk Premia
Akron,OH	133978	37893	17114	1488
Albany-Schenectady-Troy,NY	191780	42523	49034	-23928
Albuquerque,NM	292153	35415	144544	-19299
Allentown-Bethlehem-Easton,PA-NJ	229955	38208	100250	-30157
Amarillo,TX	128082	34729	18884	5841
AnnArbor,MI	165966	39107	43073	-1008
Atlanta-SandySprings-Marietta,GA	191194	38336	47676	6041
Augusta-RichmondCounty,GA-SC	120856	33056	12285	-2980
Austin-RoundRock,TX	198370	37362	49877	1987

Table continued: The Estimated Risk Premia and Implied Amenities for 2008

Metropolitan Area	Price	Wage	Amenities	Risk Premia
Bakersfield,CA	143663	30047	75981	-46690
Baltimore-Towson,MD	267441	47881	78048	-43632
BarnstableTown,MA	317123	51194	88890	-15598
BatonRouge,LA	168206	36346	31159	1614
Beaumont-PortArthur,TX	140564	35507	22388	6259
Bellingham,WA	284456	35592	127464	-53727
Binghamton,NY	120595	34367	19784	-3927
Birmingham-Hoover,AL	153568	39886	5495	1669
Bloomington-Normal,IL	152073	38865	29517	1277
BoiseCity-Nampa,ID	163398	35615	24438	-18861
Boston-Quincy,MA	336747	55220	92699	-7879
Boulder,CO	351691	50058	96189	16956
Bridgeport-Stamford-Norwalk,CT	436013	79108	122523	-36123
Buffalo-NiagaraFalls,NY	115580	37647	10385	-6018
Burlington-SouthBurlington,VT	251974	41139	90307	-26295
Cambridge-Newton-Framingham,MA	387527	60093	127056	-4981
Camden,NJ	195479	42626	44754	-27447
Canton-Massillon,OH	122679	32763	16914	6955
CapeCoral-FortMyers,FL	118378	40898	1526	-36267
Casper,WY	194152	52185	-15784	-12044
CedarRapids,IA	142387	38811	18218	1944
Charleston-NorthCharleston-Summerville,SC	212297	35447	48345	-12361
Charlotte-Gastonia-Concord,NC-SC	197281	39621	41230	2221
Charlottesville,VA	282797	43344	88028	-33482
Chattanooga,TN-GA	126886	34784	1780	-265
Cheyenne,WY	176181	44613	17075	-2685
Chicago-Naperville-Joliet,IL	248207	45510	70747	-15429
Chico,CA	229208	32349	119236	-54105
Cincinnati-Middletown,OH-KY-IN	166943	39066	37096	2389
Cleveland-Elyria-Mentor,OH	144077	40118	21777	5183
CollegeStation-Bryan,TX	139690	28176	44535	7269
ColoradoSprings,CO	208343	38221	60144	3406
Columbia,SC	139798	35328	15821	721
Columbus,OH	165190	38741	37593	2023
Corvallis,OR	277148	37755	99086	-26832
Dallas-Plano-Irving,TX	149710	43458	19554	7398
Dalton,GA	118068	28675	18833	-2710
Davenport-Moline-RockIsland,IA-IL	114984	38571	-10157	634
Dayton,OH	122437	35526	18824	2942
Deltona-DaytonaBeach-OrmondBeach,FL	156630	32098	52917	-35650
Denver-Aurora-Broomfield,CO	258493	48010	51834	20281
DesMoines-WestDesMoines,IA	152250	42506	9953	764
Detroit-Livonia-Dearborn,MI	58617	32094	-32071	533
Durham-ChapelHill,NC	188378	40927	35107	3012
EauClaire,WI	153334	33193	25137	-784
Edison-NewBrunswick,NJ	330532	51865	112694	-34683
ElPaso,TX	135003	28071	39736	-7314
Elkhart-Goshen,IN	135147	32263	35436	-204
Erie,PA	95974	32294	-3583	1804
Eugene-Springfield,OR	207351	33522	63215	-27738
Evansville,IN-KY	110415	36329	-4238	4658
Fayetteville-Springdale-Rogers,AR-MO	149133	32537	39790	-11657
Flint,MI	113815	29488	33532	1381
FortCollins-Loveland,CO	223706	38848	58087	10117
FortLauderdale-PompanoBeach-DeerfieldBeach,FL	239411	41974	94247	-57586
FortWayne,IN	93630	34176	-4478	1411
Fresno,CA	184897	30997	99964	-54344
Gary,IN	117579	35922	330	-3063
GrandJunction,CO	231281	36665	44430	-9774

Table continued: The Estimated Risk Premia and Implied Amenities for 2008

Metropolitan Area	Price	Wage	Amenities	Risk Premia
GrandRapids-Wyoming,MI	88277	33582	-16069	1189
Greeley,CO	195624	28402	88787	11137
Greensboro-HighPoint,NC	143837	35405	28144	4960
Harrisburg-Carlisle,PA	167891	39106	31027	-8813
Hartford-WestHartford-EastHartford,CT	240972	50755	72715	-15708
Holland-GrandHaven,MI	160533	33009	50988	2065
Houston-SugarLand-Baytown,TX	159217	45835	1872	6079
Huntsville,AL	157478	38259	31236	2638
Indianapolis-Carmel,IN	122092	39297	-1721	2270
Jackson,MS	140458	36054	21424	-203
Jacksonville,FL	196673	40028	36462	-21340
Janesville,WI	144590	31826	36699	-5532
Kalamazoo-Portage,MI	147363	33685	33436	1337
KansasCity,MO-KS	144052	40396	8017	2919
Kennewick-Pasco-Richland,WA	165756	33040	66283	-9875
Knoxville,TN	149041	34696	21846	-3172
LaCrosse,WI-MN	149199	35263	20591	-861
Lafayette,LA	135445	40182	-2779	1883
LakeCounty-KenoshaCounty,IL-WI	218836	51782	33225	-10426
Lakeland-WinterHaven,FL	156411	32572	44635	-30456
Lansing-EastLansing,MI	115284	33844	10074	594
LasCruces,NM	150517	27855	52528	-10469
LasVegas-Paradise,NV	178390	39920	60215	-44996
Lexington-Fayette,KY	144070	36413	22959	113
Lima,OH	107168	30351	16977	417
LittleRock-NorthLittleRock-Conway,AR	134756	39012	5170	2762
Longview,TX	95354	36046	-20838	2229
Longview,WA	203250	29703	77843	-18158
LosAngeles-LongBeach-Glendale,CA	340842	42265	200361	-97654
Louisville-JeffersonCounty,KY-IN	133968	37995	784	1069
Lubbock,TX	103485	32447	6290	3503
Macon,GA	123826	34147	13047	46
Madera-Chowchilla,CA	233514	26524	160921	-77485
Madison,WI	216103	44172	44371	-8648
Manchester-Nashua,NH	225464	45432	63246	-15225
Mansfield,OH	110216	29719	24949	1903
Medford,OR	257085	34506	111794	-52034
Memphis,TN-MS-AR	118286	38577	-3265	-189
Merced,CA	120258	27871	79226	-57403
Miami-MiamiBeach-Kendall,FL	251186	35887	89587	-49034
Midland,TX	139240	53968	-49126	7712
Milwaukee-Waukesha-WestAllis,WI	230988	42824	67918	-11209
Minneapolis-St.Paul-Bloomington,MN-WI	186015	47653	7047	-5723
Mobile,AL	130544	30567	13332	371
Modesto,CA	159318	31485	105736	-66758
Monroe,LA	118373	32204	5546	4566
Monroe,MI	117573	33397	12661	-2682
Napa,CA	393989	52169	201297	-96628
Naples-MarcoIsland,FL	232223	62559	32878	-76764
Nashville-Davidson-Murfreesboro-Franklin,TN	167671	39768	16522	-1207
Nassau-Suffolk,NY	411170	57617	132911	-30188
NewHaven-Milford,CT	229016	46918	64867	-17634
NewOrleans-Metairie-Kenner,LA	163257	41740	-5321	13291
NewYork-WhitePlains-Wayne,NY-NJ	402722	54540	140388	-31317
Newark-Union,NJ-PA	390079	56655	140479	-31607
Niles-BentonHarbor,MI	144779	33669	19300	-3714
Oakland-Fremont-Hayward,CA	351486	53093	177091	-88480
Odessa,TX	84268	34622	-24432	2141
Ogden-Clearfield,UT	209648	32799	60305	5439

Table continued: The Estimated Risk Premia and Implied Amenities for 2008

Metropolitan Area	Price	Wage	Amenities	Risk Premia
OklahomaCity,OK	128285	38882	6263	4799
Olympia,WA	254566	39988	82811	-37145
Omaha-CouncilBluffs,NE-IA	135631	43012	-7787	4075
Orlando-Kissimmee,FL	192993	35717	70500	-42265
Oxnard-ThousandOaks-Ventura,CA	405815	46787	259377	-111207
PalmBay-Melbourne-Titusville,FL	100877	37035	1245	-28061
Peabody,MA	318301	50895	107415	-4670
Pensacola-FerryPass-Brent,FL	161147	33338	51831	-23593
Peoria,IL	126718	40787	-12718	3105
Philadelphia,PA	190433	47361	17739	-21036
Phoenix-Mesa-Scottsdale,AZ	156698	36156	37786	-34795
Pittsburgh,PA	126616	42104	-15350	3052
PortSt.Lucie,FL	149184	39777	32702	-40431
Portland-SouthPortland-Biddeford,ME	211700	41522	41508	-10423
Portland-Vancouver-Beaverton,OR-WA	276563	39942	85134	-29251
Poughkeepsie-Newburgh-Middletown,NY	303410	40119	138304	-26197
Providence-NewBedford-FallRiver,RI-MA	227287	40887	69347	-22110
Provo-Orem,UT	241458	23814	102756	-3380
Pueblo,CO	133986	30564	25416	5662
Racine,WI	177849	37012	49031	-12039
Raleigh-Cary,NC	213951	39602	65866	921
Reading,PA	153884	36256	45164	-18381
Redding,CA	219005	34527	105379	-54805
Reno-Sparks,NV	194248	46929	62649	-53131
Richmond,VA	214436	42309	52788	-20864
Riverside-SanBernardino-Ontario,CA	191675	30634	115268	-62994
Roanoke,VA	169746	38727	24979	-6874
Rochester,NY	120462	39812	14085	1750
Rockford,IL	114019	32955	17320	-5917
RockinghamCounty-StraffordCounty,NH	239595	45231	57195	-6151
Sacramento-Arden-Arcade-Roseville,CA	235755	41119	137359	-77007
Saginaw-SaginawTownshipNorth,MI	55675	30143	-27700	180
Salem,OR	194581	32016	52384	-13689
Salinas,CA	247201	42857	133215	-80905
SaltLakeCity,UT	221381	38237	43451	-4870
SanAntonio,TX	151913	34937	33157	1326
SanDiego-Carlsbad-SanMarcos,CA	361444	46649	208206	-90688
SanFrancisco-SanMateo-RedwoodCity,CA	714716	76042	382048	-135615
SanJose-Sunnyvale-SantaClara,CA	520378	58531	304677	-130572
SanLuisObispo-PasoRobles,CA	389682	40635	256355	-110656
SantaAna-Anaheim-Irvine,CA	482828	51894	316297	-140048
SantaBarbara-SantaMaria-Goleta,CA	277409	47957	141104	-73889
SantaCruz-Watsonville,CA	495718	51140	291999	-113154
SantaFe,NM	310957	44927	102140	-16706
SantaRosa-Petaluma,CA	354335	47755	202088	-89487
Savannah,GA	176647	39183	19891	-8893
St.Louis,MO-IL	137199	41823	-2587	-2983
Warren-Troy-FarmingtonHills,MI	123421	44488	-13735	314
WestPalmBeach-BocaRaton-BoyntonBeach,FL	243844	58358	49911	-59038
Wilmington,DE-MD-NJ	244404	43643	86808	-32946

Note - All the variables are for year 2008. The risk premia is estimated by the exposure of a metropolitan area to the three factors and to the idiosyncratic variance. A negative risk premium should be interpreted as a cheaper house price due to the exposure of the MSA to risk. The amenities are calculated as the amount left over from prices after removing the effect of wages, expected growth, and risk premia.

Appendix B

Appendix for Chapter 2

B.1 Mathematical Appendix

In this appendix section we present a model of statistical discrimination that we estimate in section 4. Much of this model is based on the standard employer learning model formalized by Farber and Gibbons (1996). A model closely related to ours was formulated by Lange (2007), who estimates the speed of employer learning assuming symmetrical learning and a competitive labor market. We maintain these crucial assumptions in our specification.[1]

We specify the true log-productivity of a worker as:

$$\chi_{i,x} = f(s_i) + \lambda_x(z_i + \eta_i + q_i) + \tilde{H}(x) \tag{B.1}$$

The function $f(s_i)$ captures the effect of schooling on productivity for individual i. The variable q_i represents the information about the ability of the worker that is observed by the employers, but that is not available to the researcher. On the other hand, z_i is a measure of ability observed by

[1] Whether or not employers have private information about workers is an open question in the literature. Our assumption is supported by findings in Schönberg (2007) who reports that for white high school graduates learning appears to be symmetric, meaning firms do not have any private information. Since we don't find evidence of racial differences in the returns to AFQT, learning should be symmetric for high school graduate blacks too.

the researcher but not the employers. In our case this variable is the AFQT score. The part of productivity that is unobserved by both the employer and the researcher is given by η_i. The effect of (z, q, η) on log-productivity is captured by the parameter λ_x.[2] Finally, $\tilde{H}(x)$ denotes a function that captures experience effects on log-productivity. This function is assumed to be independent of education and ability measure z_i. This means that employers focus on predicting productivity based on variables s_i, q_i and signals they get over time.

The first important assumption we make is that $z_i \perp \eta_i, q_i$. This means that the unobserved part of ability and the information that employers have initially cannot be used to predict z_i. The assumption that $z_i \perp \eta_i$ is innocuous, and there is some evidence that $z_i \perp q_i$ in the data.[3] We suppress the subscript i for ease of notation from now on.

Also assume (z, s, q, η) are jointly normally distributed. This means that the expectation of $\eta \mid (s, q)$ is linear in (s, q):

$$\eta = \alpha_1 s + \alpha_1 q + v \tag{B.2}$$

Although employers do not observe z, we assume they observe the average ability of the group the worker belongs to $\bar{z} = E(z|s, x, race)$. Specifically, in our case employers know the average AFQT for each race. Employers then predict z by the linear relation:

$$z = \bar{z} + e \tag{B.3}$$

Substituting equation (B.3) in (B.1) we can write the initial log-productivity at $x = 0$ as:

$$\begin{aligned} \chi &= rs + \lambda_0(\bar{z} + e + \eta + q) + \tilde{H}(0) \qquad \text{(B.4)} \\ &= E(\chi|\bar{z}, q) + \lambda_0(e + \eta) \end{aligned}$$

[2] The lack of separate coefficients for z, η, q is without loss of generality since we can define η and q such that their coefficients are the same as that of z.

[3] In all the specifications of Table 2.2 and Table 2.5, the coefficients on AFQT are almost zero and not statistically significant for high school graduates. Assuming that AFQT matters for productivity initially, this can be interpreted as evidence that the information employers have initially cannot be used to predict AFQT.

So $\lambda_0(e+\eta)$ is the expectation error employers have initially. Over time, as they observe job performance and learn about χ, this expectation error decreases. More specifically, every period x employers get a signal given by:

$$y_x = z + \eta + \varepsilon_x \tag{B.5}$$

where ε_x is independently distributed over time as a normal with a time dependent variance $\sigma^2_{\varepsilon,x}$. We maintain that ε_x is orthogonal to all other variables in the model.

Similar to Lange (2007), the normality assumptions make the structure of employer learning very simple. In the initial period, when $x = 0$, the mean of the prior of employers' beliefs about $(z + \eta)$ is:

$$\mu_0 = \bar{z} + \alpha_1 s + \alpha_1 q \tag{B.6}$$

At some period $x > 0$ the employers get a signal y_x and they update their beliefs. Because of the normality assumption the mean of the posterior is:

$$\mu_x = (1-\theta_x)\mu_{x-1} + \theta_x y_x \tag{B.7}$$

where θ_x is some optimal Bayesian weight that the employers put on the prior mean. This process continues for any amount of experience as long as the worker's performance is observed by the employers.

At time x employers would expect the productivity of a worker to be:

$$E_x(\chi|\bar{z}, q, s, Y^x) = rs + \lambda_x q + \lambda_x\left[(1-\theta_x)\mu_{x-1} + \theta_x y_x\right] + \tilde{H}(x) \tag{B.8}$$

where $Y^x = \{y_1, ...y_x\}$. As employers learn more and more the term $\left[(1-\theta_x)\mu_{x-1} + \theta_x y_x\right]$ converges to $(z + \eta + q)$ so their expectation error collapses to zero.

Similar to the standard employer learning literature, we will maintain the assumption that all employers have access to the same information and that labor markets are competitive. Wages are then set equal to the expected productivity of a worker:

$$W_x = E_x[\exp(\chi)|\bar{z}, q, s, Y^x] \tag{B.9}$$

The normality assumptions above imply that the distribution of χ conditional on (s, q, Y^x) is normal. We can then write log wages as:[4]

$$w_x = \lambda_x\left[(1-\theta_x)\mu_{x-1} + \theta_x y_x\right] + C_x \tag{B.10}$$

where

$$C_x = rs + \lambda_x q + \tilde{H}(x) + \frac{\sigma_x^2}{2}$$

Equation (B.10) gives the wages paid to a worker given $(\bar{z}, q, s, Y^x)$. We cannot observe q and Y^x, so in order be able to estimate the wage equation we need to express log wages as a function of what we observe or $(\bar{z}, z, s, x)$. The fist step to doing this is to define a linear projection of (q, η) :

$$q = \gamma_1 s + u_1 \tag{B.11}$$

$$\eta = \gamma_2 s + u_1$$

This allows us to determine log wages as a function of only $(\bar{z}, z, s, x)$. This linear projection is given by:[5]

$$E^*(w_x|z,s) = \lambda_x\left[(1-\theta_x)E^*(\mu_{x-1}|z,s) + \theta_x E^*(y_x|z,s)\right] + c_x \tag{B.12}$$

where

$$c_x = rs + \lambda_x(\gamma_1 s + u_1) + \tilde{H}(x) + \frac{\sigma_x^2}{2}$$

Substituting in eq. (B.12) for μ_x as given in eq. (B.7), and for q given in eq. (B.11), we can

[4] Using properties of a lognormal distribution $E[\exp(\chi)|\bar{z}, q, s, Y^x] = \exp(E[\chi|\bar{z}, q, s, Y^x] + \frac{\sigma_x^2}{2})$. The expectation error is independent of $(\bar{z}, q, s, Y^x, \eta)$, so $\frac{\sigma_x^2}{2}$ does not vary with $(\bar{z}, q, s, \eta)$.

[5] Here $E^*(X|Y)$ denotes the linear projection of X on Y.

write log wages at $x = 1$ as:

$$w_1 = \lambda_1 \left[(1-\theta_1)\bar{z} + \theta_1 z\right] + k_1 \tag{B.13}$$

where:

$$k_1 = \lambda_1(1-\theta_1)\left[\alpha_1 s + \alpha_1(\gamma_1 s + u_1)\right] + c_1$$

Log wages at period $x = 1$ is a weighted average of the mean group AFQT and of the AFQT score plus a constant. The constant k_1 reflects that employers prior depends not only on mean ability $\bar{z}$, but also on schooling s and information available only to employers q.

Repeating this procedure for some $x > 1$ we can express log wages as:

$$w_x = \lambda_x \left\{ \prod_{i=i}^{x}(1-\theta_i)\bar{z} + \left[1 - \prod_{i=i}^{x}(1-\theta_i)\right] z \right\} + k_x \tag{B.14}$$

where

$$k_x = \lambda_x \prod_{i=1}^{x}(1-\theta_i)\left[\alpha_1 s + \alpha_1(\gamma_1 s + u_1)\right] + c_x$$

In order to give the log-wage equation a form similar to that shown in Lange (2007) we can rewrite it as:

$$w_x = \lambda_x \left\{(1-\Theta_x)\bar{z} + \Theta_x z\right\} + k_x \tag{B.15}$$

which is our estimating equation in section 4. Note that as experience increases the weight on $\bar{z}$ goes to zero and the weight on z to one since, as long as employers are getting new signals every period, $\prod_{i=i}^{x}(1-\theta_i) \to 0$.

B.2 Sample Creation

In this study we use the NLSY dataset for years 1979-2004. We only consider observations after the respondent has left school for the first time. Actual experience is counted as the total number of weeks that the respondent declares s/he has worked since last interview after they leave school for the first time. Potential experience is constructed as years since the respondent left school. Valid observations are kept even if the respondent goes back to school after leaving school for the first time but the additional years of education are subtracted from the experience measures.

Although the respondents report all the jobs held since the last interview, we only use the information of the current job they are holding at the time of the interview (CPS item). In addition, military jobs, jobs at home or jobs without pay are excluded from the construction of experience and from the analysis. The wage variable is the hourly rate of pay at the most recent job from the CPS section of the NSLY. The real wage is created using deflators from the 2006 economic report of the president. All observations with wages less than $1 and more than $100 are dropped. Our education variable is the highest grade completed by the respondent at the time of interview. The AFQT variable is normalized by age since respondents took the AFQT at different ages.

There are 5404 non-hispanic males in the NLSY79 sample. We drop 373 respondents who never left school or do not declare when they first left school. Out of remaining respondents 1489 graduated before 1978. For this group we constructed the work history before 1978 using three set of questions from the 1979 interview as in Altonji and Pierret (2001)(AP hereafter). Out of them, 809 respondents were dropped since their work history could not be constructed.

Next we drop 13 individuals who by the 2002 interview did not have 8 years of education, 145 if the wage was missing, 203 if AFQT was missing, and 83 individuals who at the time of the interview were not working in civilian jobs for pay or whose wages were less than $1 or more than $100. The final sample contains 3778 individuals and 38168 observations.

After keeping only observations when the highest grade completed is 12 or 16 we are left with 2714 respondents and 23732 observations. If we were to construct the sample as AP by keeping observations before year 1993 and dropping the individuals who do not have a first occupation, the sample would contain 2968 individuals and 20753 observations (AP had 2976 individuals and 21058 observations).

Table B.1: The Effects of AFQT and Schooling on Log Wages

	(1)	(2)	(3)
Model:			
Education	.0668**	.0725**	.0831**
	(.0058)	(.0045)	(.0051)
Black	-.0008	-.0244	-.0118
	(.0227)	(.0190)	(.0207)
Standardized AFQT	.0324**	.0602**	.0310**
	(.0116)	(.0010)	(.0107)
Education x experience/10	-.0240**	-.0042	-.0259**
	(.0076)	(.0038)	(.0068)
AFQT x experience/10	.0856**	.0496**	.0954**
	(.0159)	(.0079)	(.0137)
Black x experience/10	-.0735*	-.0639**	-.0737**
	(.0299)	(.0145)	(.0251)
R^2	0.2823	0.3357	0.3044
Sample	Replication of AP Years 1979-1992	Full sample Years 1979-2004	Full sample Experience<13
No. Observations	20617	37918	25726
Experience measure: Years since left school for the first time			

Note - Specification (1) is a replication of the results of AP. We also control year effects, a cubic in experience, a cubic in time with base year 1992, urban residence, and first occupation. Regression (2) uses the whole sample for years 79-04 and doesn't control for first occupation. We see a large coefficient on AFQT initially and a flat profile. Specification (3) limits the potential experience to less than 13 so the fast increase in the AFQT coefficient over time reappears. The White/Huber standard errors in parenthesiscontrol for possible correlation at individual level.

* significance level at the 95% level

** significance level at the 99% level

B.3 Replication of Altonji and Pierret (2001)

In this section we replicate the results reported on Altonji and Pierret (2001) using our sample selection criteria. AP estimate a log earning equation with linear interactions of education, race and AFQT with experience of the form:

$$w_i = \beta_0 + \beta_1 s_i + \beta_2 r_i + \beta_3 z_i + \beta_{s,x}(s_i \times x_i) + \beta_{r,x}(r_i \times x_i) + \beta_{z,x}(z_i \times x_i) + f(x_i) + \beta'_\Phi \Phi_i + \varepsilon_i \quad \text{(B.16)}$$

Log wages w_i of individual i are given as a function of schooling s_i, race r_i, AFQT scores z_i, experience x_i, and other controls Φ_i.The results of the replication are presented in Table B.1. Specification (1) uses the sample selection closest to AP with observations coming from interview years 1979-1992. The coefficients presented here differ slightly from those presented in AP because of few differences in sample construction. First, the construction of potential experience is slightly different. The potential experience measure here is years since first left school, and any years of additional education after entering the labor market are subtracted from the experience measure. This measure seems to capture the time a person actually spends in the labor market better than the experience measure in AP, which is simply age minus education minus seven. Secondly, we do not control for interactions of education and AFQT with time as that makes identification very hard and makes the estimates unstable. Regardless of the slight differences, the main qualitative results of AP are still present in the results presented in Table B.1.

Following AP's interpretation, employers seem to statistically discriminate on the basis of education. The coefficient on education is positive and significant when a worker has no experience and falls as the worker gains more experience. On the other hand, employers initially put little weight on AFQT since it might not be visible to them. As the worker gets more experience the employers slowly learn about their ability so they increase the weight they put on AFQT. The coefficient on black is insignificant and small initially, but it becomes significant and negative over time. AP use these as evidence that there is statistical discrimination on the basis of education but not on the basis of race.

Column (2) uses the same specification for our whole sample for interview years 1979-2004. The results seem similar that the returns to AFQT are greater initially and have a flatter profile with experience. The change in the AFQT coefficients in the longer sample used in Column (2) is driven by a nonlinear relation between log wages and AFQT over experience. In order to keep the interpretation of the coefficients on AFQT simple we focus on the approximately linear part of this relationship, which corresponds to experience levels less than thirteen years. The regression using this criterion is presented in column (3) of Table B.1. Restricting experience to less than thirteen years restores the low intercept and steep profile of AFQT. For the same reason explained above, we constrain the sample in our main analysis to less than thirteen years of experience.

B.4 Sample Weights

Throughout this paper we have used both the nationally representative cross-sectional sample and the supplemental sample, which oversamples blacks and low-income whites, without using sample weights. Because our final sample is not representative of the U.S. population, questions may arise about whether we should be using weights in our estimation or not. There have been examples in the literature where weights have made a difference when using the NLSY79 data. For example, MaCurdy et al. (1998) find differences in estimating the distributions of labor market earnings and hours of work when using weighted versus unweighted NLSY79 data. In order to address this concern, we estimate our key regressions using the sampling weights found on the NLSY79 and present the results in Table B.2.

The results from for all specifications are very close in magnitude and not statistically different from the results previously presented in the unweighted regressions in Tables 2 and Table 5 of the main paper. Because sampling weights do not make any difference in our results, we follow Altonji and Pierret (2001) as well as others in the literature in not including weights in presenting our main results.

Table B.2: Main Regressions Using Sample Weights

	High School		College		College minus HS P-values	
Model	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	-.0047	-.0684	.1319**	.1274**	0.001	0.004
	(.0157)	(.0309)	(.0399)	(.0411)		
AFQT x exper/10	.1222**	.1135**	.0292	.0349	0.116	0.183
	(.0203)	(.0198)	(.0557)	(.0559)		
Black	-.0924**	-.0684**	.0710	.0819	0.011	0.018
	(.0314)	(.0309)	(.0562)	(0556)		
Black x exper/10	-.0254	-.0340	-.0727	-.0747	0.563	0.617
	(.0389)	(.0379)	(.0720)	(.0720)		
R^2	0.1565	0.1784	0.1609	0.1765		
No. Observations	11795	11772	4112	4112		
Add. controls	No	Yes	No	Yes	No	Yes
Experience measure:	Years since left school for the first time <13					

Note - All specifications control for urban residence, a cubic in experience and year effects. Specifications (2) and (4) also control for region of residence and for part time vs full time jobs. In specification (5) we report the P-values for the difference in the coefficients of specifications (1) and (3). Similarly specification (6) compares (2) and (4). The White/Huber standard errors in parenthesis control for correlation at the individual level.

* statistical significance at the 90% level

** statistical significance at the 95% level

B.5 Structural Estimates with Actual Experience

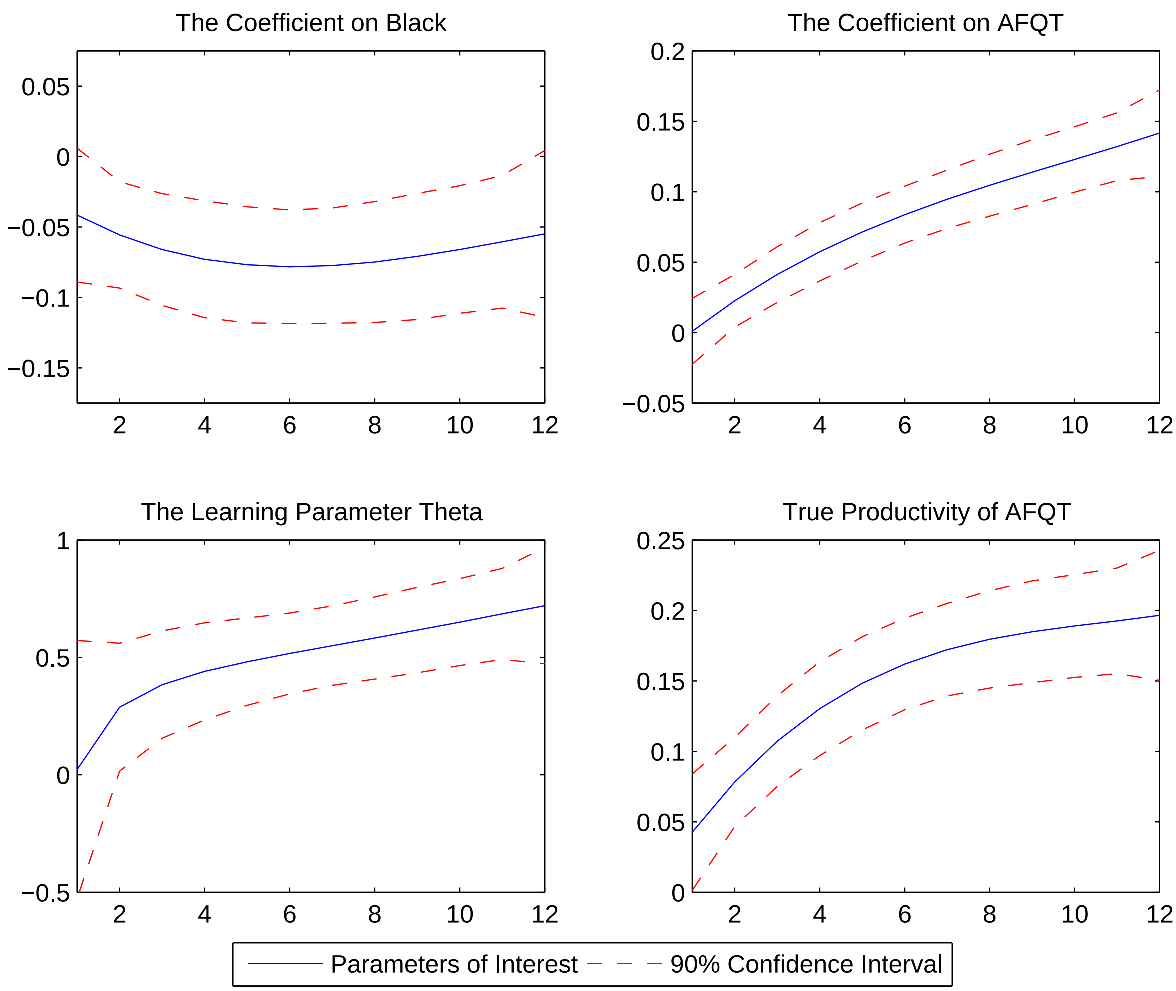

FIGURE B.1: The Evolution of the Effect of Race and AFQT on Wages, the Learning Parameter Θ_x, and the True Productivity of AFQT, λ_x. using actual experience

Bibliography

Altonji, J. G. and Pierret, C. R. (1998), "Employer Learning and Signaling Value of Education," in *Internal Labour Markets, Incentives and Employment*, eds. I. Ohashi and T. Tachibanaki, pp. 159–195, New York: McMillan Press.

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

Arcidiacono, P. (2005), "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings," *Econometrica*, 73, 1477–1524.

Arcidiacono, P., Bayer, P., and Hizmo, A. (2008), "Beyond Signaling and Human Capital: Education and the Revelation of Ability," NBER Working Papers 13951, National Bureau of Economic Research,.

Ashenfelter, O. and Krueger, A. B. (1994), "Estimates of the Economic Returns to Schooling from a New Sample of Twins," *American Economic Review*, 84, 1157–73.

Athanasoulis, S. and Van Wincoop, E. (2000), "Growth uncertainty and risksharing," *Journal of Monetary Economics*, 45, 477–505.

Bauer, T. K. and Haisken-DeNew, J. P. (2001), "Employer Learning and the Returns to Schooling," *Labour Economics*, 8, 161–180.

Bayer, P. and Timmins, C. (2005), "On the equilibrium properties of locational sorting models," *Journal of Urban Economics*, 57, 462–477.

Bayer, P. and Timmins, C. (2007), "Estimating equilibrium models of sorting across locations," *Economic Journal*, 117, 353–374.

Bayer, P. R. M. and Rueben, K. (2009), "An Equilibrium Model of Sorting in an Urban Housing Market," *NBER Working Paper 10865*.

Becker, G. S. (1964), *Human Capital*, New York: Columbia University Press for the National Bureau of Economic Research.

Bernhardt, D. (1995), "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, 315–39.

Berry, S., Levinsohn, J., and Pakes, A. (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–90.

Betts, J. R. and Morell, D. (1999), "The Determinants of Undergraduate Grade Point Average: The Relative Importance of Family Background, High School Resources,and Peer Group Effects," *Journal of Human Resources*, 32, 268–293.

Bishop, J. H. (2005), "High School Exit Examinations: When Do Learning Effects Generalize?" *Yearbook of the National Society for the Study of Education*, 104, 260–288.

Bishop, J. H. (2006), "Why California Needs a High School Exit Examination System: Enrollment+ Motivation+ Engagement=> Learning," *CAHRS Working Paper Series*, p. 453.

Bishop, J. H. and Mane, F. (2001), "The Impacts of Minimum Competency Exam Graduation Requirements on High School Graduation, College Attendance and Early Labor Market Success," *Labour Economics*, 8, 203–222.

Bishop, K. (2008), "A dynamic model of location choice and hedonic valuation," *Unpublished, Washington University in St. Louis.*

Bjerk, D. (2007), "The Differing Nature of Black-White Wage Inequality Across Occupational Sectors," *Journal of Human Resources*, 42.

Brown, C. (1984), "Black-White Earnings Ratios since the Civil Rights Act of 1964: The Importance of Labor Market Dropouts," *The Quarterly Journal of Economics*, 99, 31–44.

Butler, R. and Heckman, J. J. (1977), "The Government's Impact on the Labor Market Status of Black Americans: A Critical Review," in *Equal Rights and Industrial Relations*, ed. F. E. B. et. al., Madison, Wisc: Indus. Relations Res. Asoc.

Campbell, J. and Viceira, L. (2002), *Strategic asset allocation: portfolio choice for long-term investors*, Oxford University Press, USA.

Case, K. E., Cotter, J., and Gabriel, S. A. (2010), "Housing Risk and Return: Evidence From a Housing Asset-Pricing Model," Working Papers 201005, Geary Institute, University College Dublin.

Cocco, J. (2005), "Portfolio choice in the presence of housing," *Review of Financial studies*, 18, 535.

Cocco, J., Gomes, F., and Maenhout, P. (2005), "Consumption and portfolio choice over the life cycle," *Review of financial Studies*, 18, 491.

Cochrane, J. (1991), “A simple test of consumption insurance,” *Journal of Political Economy*, 99, 957–976.

Coulson, E. C. L. and Villupuram, S. (2008), “Urban Economic Base as a Catalyst for Movements in Real Estate Prices,” *Working Paper*.

DeVaro, J. and Waldman, M. (2006), “The signaling role of promotions: Further theory and empirical evidence,” MPRA Paper 1550, University Library of Munich, Germany.

Duffie, D. (1996), *Dynamic Asset Pricing Theory*, Princeton University Press.

Duffie, D. and Jackson, M. O. (1990), “Optimal hedging and equilibrium in a dynamic futures market,” *Journal of Economic Dynamics and Control*, 14, 21 – 33.

Fang, H. (2006), “Disentangling The College Wage Premium: Estimating A Model With Endogenous Education Choices,” *International Economic Review*, 47, 1151–1185.

Farber, H. S. and Gibbons, R. (1996), “Learning and Wage Dynamics,” *The Quarterly Journal of Economics*, 111, 1007–47.

Flavin, M. and Nakagawa, S. (2008), “A model of housing in the presence of adjustment costs: A structural interpretation of habit persistence,” *American Economic Review*, 98, 474–495.

Flavin, M. and Yamashita, T. (2002), “Owner-occupied housing and the composition of the household portfolio,” *The American Economic Review*, 92, 345–362.

Galindo-Rueda, F. (2003), “Employer Learning and Schooling-Related Statistical Discrimination in Britain,” IZA Discussion Papers 778, Institute for the Study of Labor (IZA).

Gibbons, R. and Katz, L. F. (1991), “Layoffs and Lemons,” *Journal of Labor Economics*, 9, 351–80.

Glaeser, E. and Gyourko, J. (2010), “Housing dynamics,” *Working Paper*.

Greenwald, B. C. (1986), “Adverse Selection in the Labour Market,” *Review of Economic Studies*, 53, 325–47.

Hayashi, F., Altonji, J., and Kotlikoff, L. (1996), “Risk-Sharing between and within Families,” *Econometrica: Journal of the Econometric Society*, 64, 261–294.

Heathcote, J., Storesletten, K., and Violante, G. L. (2008), “Insurance and opportunities: A welfare analysis of labor market risk,” *Journal of Monetary Economics*, 55, 501 – 525.

Henderson, V. (2002), “Valuation of claims on nontraded assets using utility maximization,” *Mathematical Finance*, 12, 351–373.

Henderson, V. (2005), "Explicit solutions to an optimal portfolio choice problem with stochastic income," *Journal of Economic Dynamics and Control*, 29, 1237–1266.

Hinkelmann, C. and Swidler, S. (2008), "Trading house price risk with existing futures contracts," *The Journal of Real Estate Finance and Economics*, 36, 37–52.

Hizmo, A. (2010a), "The Common Variation in Housing Price Returns," *Working Paper*.

Hizmo, A. (2010b), "Hedging Housing Risk with Stock Indexes from Local Employers," *Working Paper*.

Kraft, H. and Munk, C. (2010), "Optimal housing, consumption, and investment decisions over the life-cycle," .

Lang, K. (1994), "Does the Human-Capital/Educational-Sorting Debate Matter for Development Policy?" *American Economic Review*, 84, 353–58.

Lang, K. and Manove, M. (2006), "Education and Labor-Market Discrimination," NBER Working Papers 12257, National Bureau of Economic Research, Inc.

Lange, F. (2007), "The Speed of Employer Learning," *Journal of Labor Economics*, 25, 1–35.

MaCurdy, T., Mroz, T., and Gritz, R. M. (1998), "An evaluation of the national longitudinal survey on youth," *Journal of Human Resources*, 33, 345–436.

Merton, R. (1969), "Lifetime portfolio selection under uncertainty: The continuous-time case," *The review of Economics and Statistics*, 51, 247–257.

Merton, R. C. (1971), "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3, 373–413.

Mincer, J. (1974), *Schooling, Experience and Earnings*, New York: Columbia University Press.

Neal, D. A. (2004), "The Measured Black-White Wage Gap among Women Is Too Small," *Journal of Political Economy*, 112, S1–S28.

Neal, D. A. (2006), *Why Has Black-White Skill Convergence Stopped?*, vol. 1 of *Handbook of the Economics of Education*, chap. 9, pp. 511–576, Elsevier.

Neal, D. A. and Johnson, W. R. (1996), "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 104, 869–95.

Ortalo-Magne, F. and Prat, A. (2010), "Spatial Asset Pricing: A First Step," *Working Paper*.

Piazzesi, M., Schneider, M., and Tuzel, S. (2007), "Housing, consumption and asset pricing," *Journal of Financial Economics*, 83, 531–569.

Pijoan-Mas, J. (2006), "Precautionary savings or working longer hours?" *Review of Economic Dynamics*, 9, 326–352.

Pinkston, J. C. (2009), "A Model of Asymmetric Employer Learning with Testable Implications," *Review of Economic Studies*, 76, 367–394.

Roback, J. (1982), "Wages, rents, and the quality of life," *The Journal of Political Economy*, 90, 1257.

Rosen, S. (1979), "Wage-based indexes of urban quality of life," *Current issues in urban economics*, 3.

Schönberg, U. (2007), "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 25, 651–691.

Shiller, R. (1993), *Macro markets*, Clarendon Press.

Shiller, R. (2003), "The new financial order: Risk in the 21st century," *Cato Journal*, 23.

Shreve, S. E. (2004), *Stochastic calculus for finance. II*, Springer Finance, Springer-Verlag, New York.

Spence, A. M. (1973), "Job Market Signaling," *The Quarterly Journal of Economics*, 87, 355–74.

Stiglitz, J. E. (1975), "The Theory of "Screening", Education, and the Distribution of Income," *American Economic Review*, 65, 283–300.

Svensson Ingrid, M. and Lars, E. (1993), "Nontraded assets in incomplete markets* 1:: Pricing and portfolio choice," *European Economic Review*, 37, 1149–1168.

Tepla, L. (2000), "Optimal hedging and valuation of nontraded assets," *European Finance Review*, 4, 231.

Tracy, J., Schneider, H., and Chan, S. (1999), "Are stocks overtaking real estate in household portfolios?" *Current Issues in Economics and Finance*, 5.

Van Hemert, O. (2010), "Household interest rate risk management," *Real Estate Economics*.

Van Nieuwerburgh, S. and Weill, P. (2010), "Why Has House Price Dispersion Gone Up?" *Review of Economic Studies*.

Vigna, E. (2009), "Mean-variance inefficiency of CRRA and CARA utility functions for portfolio selection in defined contribution pension schemes," *CeRP Working Papers*.

Waldman, M. (1984), "Job Assignments, Signalling, and Efficiency," *RAND Journal of Economics*, 15, 255–267.

Weiss, A. (1995), "Human Capital vs. Signalling Explanations of Wages," *Journal of Economic Perspectives*, 9, 133–54.

Wigdor, A. K. and Green Jr., B. F. (1991), *Performance Assessment for the Workplace*, Washington, DC: National Academy Press.

Yao, R. and Zhang, H. (2005), "Optimal consumption and portfolio choices with risky housing and borrowing constraints," *Review of Financial Studies*, 18, 197.

Zeldes, S. (1989), "Consumption and liquidity constraints: an empirical investigation," *The Journal of Political Economy*, 97, 305–346.

Zhang, Y. (2006), "Asymmetric Information, Employer Learning, and the Job Mobility of Young Men," Tech. rep., Working Paper. Indiana University-Purdue University Indianapolis.

Biography

Aurel Hizmo was born on 29th of April, 1982 in Kavaje, Albania. He received a B.S.B.A in Economics, *summa cum laude*, from Saint Louis University in 2005. Before that, he studied Economics at the University of Tirana (2000-2001) and Saint Louis University, Madrid campus (2001-2003). He also received a M.A. in economics from Duke University in 2008. Part of his studies at Duke have been funded by the Distinguished Economics Department Graduate Fellowship and the Summer Dissertation Research Fellowship. Aurel has accepted a position as an Assistant Professor at the Department of Finance at the Leonard N. Stern School of Business at New York University.