### **Stanford University**

#### From the SelectedWorks of John Donohue

October, 2006

## The Economics of Labor and Employment Law

John Donohue, Yale University



Available at: https://works.bepress.com/john\_donohue/31/

The Economics of Labor and Employment Law John J. Donohue III October 3, 2006

Labor law encompasses a broad and amorphous body of rules and regulations that govern an enormous array of features of the working lives and the economic welfare of workers and their families. Around the world and within countries, there is dramatic variation in the rules that govern unionization and the work relationship, as well as the process and permissible bases for hiring and firing workers. On the one hand, many countries in Latin America – Brazil is one good example – regulate the work relationship in minute detail. Judges are permitted to determine whether even minor changes in working conditions are consistent with humanist principles of fairness. The outcome – reminiscent of Grant Gilmore's line that "in hell, there will be nothing but law, and due process will be strictly observed" – has been a disaster for the economic well-being of Brazilian workers that has only been mitigated by a substantial flight to the unregulated black market in labor. On the other hand, countries such as the United States have had relatively less restrictions on labor contracting, and have enjoyed immense growth in their labor markets over time.

In the last decade, a strong theme of writing on labor markets in the OECD and the IMF has been that labor institutions are a major source of unemployment in advanced countries. Inflexible, over-regulated markets are clearly bad, but the goal of policy is to determine optimal levels of regulation, which requires judgment about where in the large spectrum between the mercantile approach of some Latin American countries and the completely laissez faire lies the appropriate degree of regulation. Michael Mussa, former chief economist at the IMF, wrote in 2002 that the collapse of the Argentinean economy at that time could have been avoided with greater flexibility in its economics system, particularly in its labor markets.<sup>1</sup>

At the same time, the International Labor Organization issued a report stating that it

"takes issue with the view that labour market rigidity has been the major cause of unemployment and that greater labour market flexibility is the solution ... jobless rates appear to have risen independently of levels of labour market regulations ...trade union power was reduced in many countries, together with unemployment benefits and in some cases minimum wages, producing little if any positive employment effect." (www.jobsletter.org.nz/jbl05210.htm).<sup>2</sup>

Some also point out that labor regulations and union-imposed standards for work time (as well as high marginal tax rates) enable Europeans to enjoy significantly greater leisure

<sup>&</sup>lt;sup>1</sup> Freeman, Richard (2005), 'Labour Market Institutions without Blinders: The Debate over Flexibility and Labour Market Performance', *International Economic Journal*, 19 (2), 129–145.

<sup>&</sup>lt;sup>2</sup> Quoted by Freeman, at p. 133.

than Americans: the French, Germans, and Italians work about 400 hours less than the 1800 yearly hours worked by the average American worker.<sup>3</sup>

#### I. Overview: A Cross Country Comparison of the Regulation of Labor

I begin the readings with a paper that takes a very comprehensive look at the regulation of labor across 85 countries in order to understand the origins and consequences of varying regulatory approaches -- Botero, Juan C., Simeon Djankov, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer (2004), 'The Regulation of Labor', *Quarterly Journal of Economics*, 119 (4), 1339-1382. Botero et al. focus their inquiry on three theories of institutional design. The efficiency theory posits that governments choose institutions in order to maximize some social welfare function or at least that institutions adjust to that end. The political power theory assumes that the ruling classes make decisions that benefit themselves and their patrons. The legal traditions theory traces the development of institutions to the country's core legal tradition, which is typically inherited from colonial powers. This theory predicts differential outcomes for countries with civil versus common law traditions, and one of the authors' primary aims is to distinguish this effect from purely political forces.

Botero et al. divide the broad expanse of labor law into three categories: employment protection, collective bargaining, and social security. Their empirical agenda begins with the coding of the laws of 85 countries as of 1997 for the three categories, at which point they explore which institutional theories best explain the legal outcomes. Their worker protection index seeks to capture the marginal cost to employers of departing from standard contract provisions by, for example, hiring temporary workers or dismissing employees. Their coding of collective bargaining laws focuses on the power of labor unions and the balance of power between firms and unions enshrined in procedural protections. Their index of social security is based on old age pensions, coverage for illness and unemployment insurance. The authors create their indices so that higher values will reflect greater protections of and benefits for workers. For example, as Table 1 illustrates, Portugal had one of the highest index values in the sample, while the United States and New Zealand ranked nearer the bottom because of their relatively low values for employment and collective bargaining protections. But one could query whether these indices can be taken as having more than ordinal significance. Is the regulation of employment law in Portugal really five times more stringent than that of New Zealand, and what does that really mean?

<sup>&</sup>lt;sup>3</sup> Alesina, Alberto, "Europe," *NBER Reporter*, Summer 2006, p. 8.

(1 of tugal, Officer States, New Zealand)			
	Aggregate Index Values Shown in Bold Ranking within 85 Countries shown in Parentheses		
Countries	Employment Laws	Collective Relations Laws	Social Security Laws
Portugal	0.809	0.649	0.735
	(3rd)	(4th)	(27th)
United States	0.218	0.259	0.646
	(77th)	(76th)	(46th)
New Zealand	0.161	0.25	0.719
	(84th)	(77th)	(32nd)
Mean	0.488	0.445	0.569
Median	0.475	0.455	0.677
N	85	85	85

#### Table 1: Employment Law Index Values and Rankings (Portugal, United States, New Zealand)

In their regression models based on these indices, Botero et al. find, contrary to the efficiency theory, that prosperous countries tend to provide more generous social security systems but that employment regulation is unrelated to income. Civil law countries mandate greater worker protection than common law states in the areas of employment and collective bargaining law. Support for the political theory emerges with evidence that longer experience with centrist or leftist governments and higher levels of union density lead to heavier labor market regulation. When the legal and political determinants are included in the same regression, the authors find that only the former generate robustly significant effects on regulatory regimes.

In order to bolster further the findings in favor of legal tradition, Botero et al. report relatively high correlations between their labor law indices and economic entry regulations, which the authors quantify by counting the number of steps or number of days it takes to start a business in a particular country. The authors conclude from their analysis that stronger labor regulation leads to worse economic outcomes, including lower levels of labor force participation, higher levels of youth unemployment, and a larger underground economy. Thus, their work is deemed to undermine the efficiency theory and support the view that legal tradition substantially influences labor regulation, with the common law leading to the lowest level of regulation.

Obviously, there is a high degree of distillation involved in the attempt to quantify such a large body of law into index numbers. Many questions can be raised about whether the authors have been sufficiently attentive to differences between the law on the books, which is their main focus in coding, and the law in practice, which is obviously harder to identify for so many laws in countries all across the globe. Moreover, their various index scores represent averages of sub-indices, so that two countries with identical values on, say, the social security index might have very different protections for the sub-components of old-age benefits, health benefits, and unemployment compensation benefits. Taking the U.S. as an example, one can imagine ways in which federal labor regulation may sensibly be captured in a single index, but this would mean that the broad variation in regulation across the states would be missed. Finally, critics of Botero et al. have charged that their analysis may simply be picking up correlations that have no causal significance. But even the critics have conceded that this is very provocative work raising important questions that need further exploration, particularly as aging populations in the developed world will put increasing pressure on the old age benefit programs of many countries.

#### II. The Economics of American Labor Law

While Botero et al. try to examine the broad contours of labor regulation across the globe, important papers by Richard Posner, Ken Dau-Schmidt, and Cass Sunstein evaluate the elements of American labor law from the perspectives of the Chicago School, progressive law and economics, and behavioral economics and come to strikingly different conclusions on the value of certain features of the legal landscape. In a characteristically provocative essay -- **Posner, Richard A. (1984), 'Some Economics of Labor Law',** *University of Chicago Law Review*, **51** (4), **988-1011** -- Posner depicts labor unions as cartels whose primary function is to elevate the firm wage beyond competitive levels. Posner argues that the consequences of such anti-competitive pricing tend to be socially harmful, and therefore he laments the fact that federal labor law has encouraged the emergence of cartels.

Posner states that labor law does not passively manage the collective bargaining environment, but rather facilitates the organization of firm employees while narrowly setting the rules of the game. During the interwar period, Posner concedes that labor markets often did not generate competitive wages due to informational asymmetries. In the absence of organizations, owners could exploit workers' lack of outside options or their investment in firm-specific human capital to keep wages artificially low. Therefore, one might consider the Wagner Act and the subsequent unionization movement as a means of reducing wage distortions since it codified the right of employees to redress grievances over pay and working conditions. Without such protection, the classic free rider problem would surely prevent workers from mounting successful resistance to the "predatory" actions of employers.

Posner describes the collective bargaining system from the birth of a union to the strike phase while highlighting the likely motives of employers and employees. He observes that the two sides in a labor dispute effectively find themselves in a situation of bilateral monopoly. The union has only one "customer" for its services, and the firm may only negotiate with the union representatives. In the event of a strike, the union must weigh lost income against future wage concessions, while the employer must balance the reputation gains from standing firm against the cost of lost production. Posner claims that labor law ultimately determines the nature and extent of costs borne by both parties. By specifying the terms under which the firm may hire replacement workers, for example, labor law slightly shifts the cost balance in favor of the employer.

Posner rejects the claims most famously advanced by Freeman and Medoff that unions actually enhance the productivity of the labor force, finding union support for the minimum wage or federal oversight regarding working conditions to be better explained by his cartel theory. He argues that government intervention through the National Labor Relations Act helps employees overcome the large-numbers difficulty that hampers cartelization efforts. Specifically, labor law a) reduces competition between union sympathizers and workers who would accept the competitive wage, b) restricts free riding among the labor force, thereby increasing the wealth and power of the union, and c) establishes picketing as a method for detecting cheating among the cartel's ranks.

A decidedly more optimistic assessment of unions and labor laws underlies the work of Ken Dau-Schmidt -- Dau-Schmidt, Kenneth G. (1992), 'A Bargaining Analysis of American Labor Law and the Search for Bargaining Equity and Industrial Peace', *Michigan Law Review*, 91 (3), 419-514. Dau-Schmidt's paper argues against the view that unions are barriers to efficiency that lead to rising prices, as well as declining output and labor employment, and therefore should not be encouraged by federal labor law. According to Dau-Schmidt's formulation, union wage benefits derive from employer rents and productivity boosts rather than cartelization, and the generated surplus becomes the focus of Coasean collective bargaining. As a result, unions provide wage gains for their employees without imposing attendant inefficiencies in hiring levels or consumer prices by squeezing employer rents and increasing productivity. Dau-Schmidt observes that because bargaining players may resort to strategic behavior such as strikes or lockouts in order to obtain greater shares of the surplus, federal labor law can play a socially beneficial role by establishing the "rules of the game" in order to discourage such strategic behavior.

Dau-Schmidt begins by identifying three sources of union wage increases: 1) rents from the cartelization of labor markets; 2) employer product market rents and quasirents generated by the capital stock and 3) increases in productivity. He then discusses how employers might respond to the union's wage demands regardless of the source of surplus. The first hypothesis posits that employers "move up" their demand curve, i.e. employ fewer workers at the higher wage. Yet, this suboptimal outcome is unlikely to ensue as long as employers and workers can bargain over employer rents. For, under certain assumptions about the employees' utility functions and the employer's profit function, there exists a set of Pareto-improving wage-employment points off the demand curve. Turning to the costs of collective bargaining, Dau-Schmidt distinguishes between the transaction costs associated with strikes and the propensity of both sides to resort to collectively irrational strategic behavior – in accord with the predictions of the wellknown prisoner's dilemma.

Dau-Schmidt's assessment of American labor law conflicts with the monopoly model that conceives union activity as an adverse force on economic activity, one that achieves inequities between organized workers on the one hand and consumers and laid off employees on the other. Dau-Schmidt notes that the goal of federal labor law has been to facilitate collective bargaining and "industrial peace," and he argues that theory and empirical evidence refute the claims that wage increases are obtained solely through cartelization, and that employers do not bargain in response to union demands. If the size of the pie can be increased through bargaining, the parties have an incentive to reach the efficient solution if bargaining costs can be lowered. Consequently, labor law can therefore lower bargaining costs as a way to achieve more efficient outcomes than the cartel model anticipates.

Dau-Schmidt offers his preferred set of assumptions that yield conditions under which unions may actually generate efficient and equitable outcomes, especially when one allows for productivity increases as a source of wage increases. However, mindful of strategic motivations, Dau-Schmidt contends that governments can and should adjust negotiation costs in order to direct parties to the set of efficient bargaining solutions. In his view, the bargaining model also matches the public policy goals of labor law and the empirical activities of employers and unions. Nevertheless, he identifies several suboptimal aspects of labor law that cannot be explained by the promise of the bargaining model. These include the lack of stricter penalties for violating the National Labor Relations Act (which would help align bargaining costs more efficiently), delays in the union certification process, the ability of employers to permanently replace striking workers, and the problem of bad faith bargaining just after organization. Dau-Schmidt also effectively marshals accessible game theory concepts to stress the strategic underpinnings of wage negotiations. One can think of Posner and Dau-Schmidt as offering the polar positions in the law and economics literature on the costs and benefits of unions and American labor law.

Along with his coauthors Christine Jolls and Richard Thaler, Cass Sunstein has been on the forefront of the effort to incorporate concepts of cognitive psychology and behavioral economics into the analysis of law.<sup>4</sup> In a specific application of this broad theme -- Sunstein, Cass (2001), 'Human Behavior and the Law of Work', *Virginia Law Review*, 87 (2), 205-276 – Sunstein analyzes the structure of employment law in the hope of enriching previous law and economics work by being attentive to areas in which actual human behavior seems to depart from the standard neo-classical assumptions.

Sunstein makes two major claims. First, traditional employment law runs astray when it is not based on workers' actual values and behavior.<sup>5</sup> For example, a standard economic analysis might begin with the Coase Theorem, which states that with zero transaction costs the initial allocation of rights does not matter since the parties will negotiate to the same, efficient solution regardless of the initial allocation – what I have referred to as the identity prediction.<sup>6</sup> While Sunstein accepts the theorem's validity with respect to efficiency – thus implicitly accepting Dau Schmidts's point and rejecting Posner's thesis -- he doubts that the solution would be *the same* depending on who gets the right first. His assessment is based on the "endowment effect" a behavioral finding that people tend to value something more highly when it is initially allocated to them rather than allocated to someone else. The endowment effect creates a spread between a person's willingness to buy a right that they lack and their willingness to sell a right that they were allocated. So contrary to the conventional analysis, for Sunstein the initial allocation matters because it determines the optimal outcome. The real question then becomes: who should initially be assigned the right? By ignoring behavioral findings – such as the endowment effect - the traditional analysis overlooks the importance of alternative approaches to employment law.

<sup>&</sup>lt;sup>4</sup> Jolls, Christine, Cass R. Sunstein and Richard Thaler (1998), 'A Behavioral Approach to Law and Economics', *Stanford Law Review*, 50 (5), 1471-1550.

<sup>&</sup>lt;sup>5</sup> Some important points of divergence between the conventional view of workers' values and behavior and the newer behavioral view are that workers are particularly loss averse, they greatly discount the future, and they care more about their relative economic position than their absolute economic position.

<sup>&</sup>lt;sup>6</sup> Donohue, John (1989), 'Diverting the Coasean River: Incentive Schemes to Reduce Unemployment Spells,' *Yale Law Journal*, 99, 549-609.

The second major claim that Sunstein makes is that waivable workers' rights are currently underutilized but are a promising approach for future employment law reforms. More specifically, Sunstein suggests a two-tiered system, in which the first tier is made of non-waivable statutory minimal safeguards, and the second tier is made of waivable worker's rights. These waivable workers' rights would come in two forms, constrained and unconstrained. The constrained rights would be subject to government regulation and possible price floors, while the unconstrained ones would be subject only to market forces and free contracts between employers and employees. The underlying motivation for the waivable workers' rights is that it is often unclear what a "market mimicking" employment law rule would look like so waivable workers' rights are the best way to elicit information. They compel employers to be very specific in contracts about the rights they want the employee to waive and thus allow employees to be better informed about the rights that they have.

This information eliciting requirement is important to Sunstein because, contrary to the premise of conventional analysis, workers often don't know their rights, as Pauline Kim's work has underscored.<sup>7</sup> For example, Kim found in Missouri that a large majority of employees (80% or more) were quite ignorant of the laws concerning wrongful discharge. Many thought it was unlawful to be fired if the employer merely wanted to hire someone else or if he personally disliked them. However, all are lawful grounds for discharge in Missouri. Similar results were found in California and New York. Kim also notes that such worker ignorance is independent of a number of variables, including geography, age, work experience, and union experience. While there are many causes for this lack of understanding, behavioral economics and cognitive psychology provide insight in at least some of them. As Sunstein outlines, people like to align their beliefs about how things are with how they think things should be to reduce cognitive dissonance, thus making themselves subject to self-serving bias. So their understanding of fair laws is mediated by their biased beliefs in fairness, which in turn makes them believe that they can not be fired without cause. Ultimately, the lack of information on the side of the workers undermines the traditional explanation for an at-will doctrine, namely that it reflects the shared understanding of the parties.

Sunstein's analysis sweeps across many important areas of employment law including unionization, occupational safety and health, discrimination, vacation and parental leave time and workers compensation. He is always applying the behavioral insights with the intention of finding criteria to distinguish waivable from non-waivable rights. In his opinion, important criteria for determining whether rights should be nonwaivable are: whether third parties are affected, whether workers have inadequate information (even if information eliciting procedures are in place), and whether the rights involve the establishment of norms that are not acceptable to society at large (e.g.,. the decision to waive discrimination rights). Ultimately, Sunstein prefers waivable rights because they provide less "rigidity, inefficiency and potential harm to workers and consumers" that are created by non-waivable "one-size-fits-all" rights.

Sunstein's article is important in that it: 1) incorporates new behavioral findings into debates over controversial doctrines and institutions, such as employment at will and unionization; and 2) recommends a two-tiered system where there is a minimum floor of

<sup>&</sup>lt;sup>7</sup> Kim, Pauline (1997), 'Bargaining With Imperfect Information: A Study of Worker Perceptions of Legal Protection in an At-Will World,' *Cornell Law Review* 83(1):105-160.

workers rights below which no employer may go and a number of waivable rights above this floor that are subject to negotiation. In the area of the contract at will, he suggests that "legislatures should experiment with waivable for-cause rules, and also that courts should move in this direction by penalizing employers who lead employees to believe that they have protection against at-will discharge."

## **III.** The Impact on Economic Welfare of the Regulation of Labor in the U.S. and the World

Thomas Holmes tries to determine empirically whether state right-to-work laws stimulate business activity -- Holmes, Thomas J. (1998), 'The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders', *Journal of Political Economy*, 106 (4), 667-705. Such laws outlaw the union shop, the hiring structure in which all the firm's employees must belong to the union. Presumably, if unions dampen business activity in the way that Posner's analysis would suggest, then laws that decrease the extent of union penetration should encourage business expansion.

Holmes notes that the cross-section data are not supportive of the "unions tax business" thesis in that the states traditionally known as manufacturing centers in the United States do not have right-to-work laws. But such cross-section data cannot readily reveal the causal impact of the law. Even a panel data analysis that regresses manufacturing growth on a legal indicator variable may not reveal the causal impact of a right-to-work law if other unrelated characteristics of right-to-work states, for which these models do not account, lead to geographic shifts in manufacturing activity, while the passage of right-to-work laws coincidentally correlates with these features.

Holmes develops a theoretical model that illuminates why policy differences at state borders should matter. Holmes describes the manufacturing entrepreneur's decision as one between remaining in the firm's current location versus moving to a more probusiness locale, provided moving costs are not prohibitive. Based on his model, Holmes predicts there will be a discontinuous jump in manufacturing activity at the border but that the effect will dissipate as one moves farther from the border into the pro-business region. Because moving costs are positive, the farther from the border an entrepreneur finds himself, the less state policy matters and the less relocation occurs.

Holmes' empirical analysis assigns to each county a distance variable representing the minimum distance from the population centroid to the border. Holmes then explores the effect of right-to-work laws on two dependent variables: the employment share of the manufacturing sector and the postwar growth rate of manufacturing employment. Straightforward comparisons of means at various distances from the policy border reveal stark differences in both measures. On the anti-business side, the mean employment share of the manufacturing sector is only 21.0 percent with a post-war growth rate in manufacturing employment of only 62.4 percent; on the probusiness side, the mean employment share is 28.6 percent with a growth rate of 100.7 percent. Moreover, the employment shares follow the patterns predicted by Holmes' model when one moves through the various anti- and pro-business "layers."

Holmes' regression analysis estimates a county's manufacturing employment share as a function of the state's business posture, its distance from the border and its position along the border (employing several functional forms for the latter two). Holmes' primary analysis uses data from states bordering right to work states, from which he conducts a county level analysis on those counties whose population centroid is within 100 miles on either side of the border separating the "probusiness" states from the "antibusiness" states. In addition, Holmes conducts a state level analysis looking at pairs of states that lie along the border – that is, he pairs a probusiness state on one side of the border with corresponding antibusiness state on the other side of the border. In this sideby-side comparison of 17 pairs of such states, Holmes measures their mean manufacturing employment share in 1992 and their manufacturing employment growth rate between 1947 and 1992. He finds once again that there are noticeable differences in both variables when crossing from one side of the border to the other. The right-to-work law is found to increase the manufacturing share – which averages about 20 percent – by about 6.6 percentage points, or roughly by one-third. In addition, Holmes finds that manufacturing employment grew substantially faster between 1947 and 1992 in the probusiness counties relative to the antibusiness counties. Ultimately, Holmes has marshaled impressive evidence that a more pro-business industrial policy – proxied by the presence of a state right-to-work law -- increases manufacturing activity.

Shifting focus from the effect on manufacturing of anti-union right-to-work laws in the United States (Holmes 1998) to pro-worker legislation in states across India provides further support for the view that governmental solicitude for workers can come at a price -- **Besley, Timothy and Robin Burgess (2004), 'Can Labor Regulation Hinder Economic Performance? Evidence from India',** *Quarterly Journal of Economics*, **119 (1), 91-134.** At a time when many Asian economies experienced significant GDP increases tied to growth in the manufacturing sector, the share of manufacturing output in India only rose by 5 percentage points between 1960 and 1995. Besley and Burgess query whether state-level variation in labor policy explains the crossstate differences in manufacturing output and employment as well as the overall slow national manufacturing growth. Their primary conclusion is that pro-worker policies are associated with more depressed conditions and the growth of the informal economy.

Pursuant to the Regulation and Development Act of 1951, the central government in India established the national industrial policy. However, beginning in the 1950s, Indian states became more involved with industrial relations by amending another federal statute, the Industrial Disputes Act of 1947. Besley and Burgess code each of the 113 state amendments as pro-worker (1), pro-employer (-1) or neutral (0), which are then used to characterize the 10 treatment states during the period 1958-1992. Initial evidence suggests that the pro-worker states started out with higher productivity relative to proemployer and control states and that this gap shrank considerably by 1990.

Besley and Burgess hypothesize that regulation might affect economic outcomes in two ways. If the price of labor relative to capital rises after regulation, one might expect to observe a substitution effect away from labor. Since regulation only covers registered firms in India, the unregistered sector may grow in size and productivity. If regulation leads to greater union expropriation of investment returns, however, the capital stock may decline in the face of holdup threats.

Their empirical strategy employs a panel data model in which a logged economic outcome measure is regressed on lagged regulation measures, a vector of controls, and state and year fixed effects. Besley and Burgess estimate this model with a variety of dependent variables, including total state, agricultural, non-agricultural and manufacturing output. The negative coefficient on regulation achieves significance only when the latter is used, which lends credence to their hypothesis because the regulations amending the 1947 Industrial Disputes Act were *sector specific* and should only impact the manufacturing sector. If instead the effect in another regression on a different sector such as agricultural output turned out to be negative, this would imply that the labor regulations for manufacturing actually were proxying for other (bad) government policies. Moreover, the estimate becomes more negative and statistically significant when only registered firms are included and becomes positive and significant when unregistered firms are analyzed, consistent with the hypothesized substitution effect to the informal sector.

The sensitivity analysis performed by Besley and Burgess generally supports these baseline findings. Adding several control variables for state infrastructure, health, education and political parties in power has little effect on the regulation coefficient. Only the addition of state-specific trends appears to dilute the effect of regulation on manufacturing output. Turning to alternative economic measures, Besley and Burgess report a negative effect of regulation on manufacturing employment, labor usage and capital formation but no discernible effect on wages. However, they consider the possibility of endogeneity through reverse causation from manufacturing performance to adoption of labor regulation measures. Using instruments such as union membership and patterns of land tenure (which is hypothesized to correlate with development and thus political power), Besley and Burgess again confirm the robustness of their original estimates.

Finally, the authors assess the effect of regulation on poverty, distinguishing between urban and rural effects. In line with their earlier estimates, labor regulation has a significant positive effect on urban poverty since most manufacturing firms are located in cities and regulation has a negative effect on manufacturing output. However, this effect disappears when state trends are added to the model. Besley and Burgess's overall conclusion: pro-worker regulation policies, which are ostensibly designed to improve economic welfare for the worst off, adversely impact the poor.

Further insight into the impact and behavior of unions can be gained by looking at the circumstances in which unions thrive or are undermined. One major event in the U.S. that substantially impacted the labor market was the deregulation of four major American industries that occurred in the 1970s and 1980s. This initiative led to the easing of governmentally imposed barriers to entry and the elimination of rate schedules in the trucking, railroad, airline, and telecommunications industries. A useful examination of the impact on labor markets of this important experiment in deregulation is provided by **Peoples, James (1998), 'Deregulation and the Labor Market', Journal of Economic Perspectives, 12 (3), 111-130.** 

One might predict that the removal of barriers to entry to create an essentially competitive environment would have a dampening effect on unionization and reduce the capacity of unions to drive hard bargains for higher wages. Indeed, data from the trucking sector for the period 1978-1996 indicate a fall in union membership from 46 to 23 percent, an extraordinary increase in the labor force, and a concomitant 28 percent decrease in earnings. The railroad industry, however, did not exhibit much of a decline in

union participation due to its natural barriers to entry. Overall railroad employment did fall considerably, while weekly wages decreased only slightly. Among the airlines union membership fell from 45 to 36 percent, as total employment trended upward. Telecommunications companies exhibited the largest weakening of union power; membership totals fell from 59 percent to 29 percent from 1973 to 1996. Although Peoples attributes this decline to the introduction of labor-saving technologies, he also notes that the skill required to implement them led to an upsurge in employment and wages over the observation period. The declines in union power in these four industries that was prompted by deregulation still left the four industries above the national average in terms of percentage union membership and earnings.

Peoples highlights the relationship between regulatory legislation and industrial organization in the four industries. He observes that the trucking industry best approaches the competitive paradigm in the absence of regulation because of low skill requirements and entry costs. The Motor Carrier Act of 1935, however, restricted entry and set rates for interstate commercial transport. The Teamsters union gained substantial power during this period only to see that strength weakened by the easing of barriers to entry following passage of the 1980 Motor Carrier Act. Regulation of the railroads, which tend to resemble natural oligopolies, began in the 1920s. This move eventually harmed the industry because the setting of rates above their competitive levels exposed it to competition from other shipping modes, namely trucking. The 1926 Railroad Labor Act spurred heavy yet disjointed unionization in this sector that kept crew numbers above their efficient level as technology made their services redundant. Legislation in the 1970s and 1980s allowed carriers to charge competitive rates and amend work rules in a more resourceful manner. The airline experience with respect to pricing behavior and union membership mirrored that of the railroads since regulation seemed a logical response to high startup costs and fears of price wars. Finally, the history of deregulation in the telecommunications industry story is dominated by the initial privileged market position and eventual breakup of AT&T. After a period of extensive union growth, the dissolution of AT&T led to highly fractured -- and weakened -- collective bargaining.

Peoples concludes with some empirical evidence of deregulation's effect on wage premiums. Regressing log earnings on a variety of individual-level covariates and a dummy for employment in a deregulated industry using CPS data, he presents separate graphs of the coefficient on the deregulation variable for each of the four industries. Trucking experienced the most extreme decrease in the wage premium after 1982 (the premium eventually transforms into a loss relative to non-transportation operatives). The railroad industry did not exhibit a noticeable trend in its wage premium, and airline employees experienced a short-lived increase by 1982 but an eventual decline by 1991. The telecommunications sector displays the clearest break in trend following deregulation: after six years of steady wage gains, changes in the collective bargaining environment led to an abrupt reversal of fortune. While deregulation generated many economic gains, it did impose some major costs on certain elements of the unionized workforce.

#### **IV. Mandating Employee Benefits**

#### A. Minimum Wage Laws

One of the most intriguing debates in labor economics was launched by David Card and Alan Krueger, who upended the world of price theory with an empirical paper that purported to show no disemployment effect from an increase in the minimum wage -- Card, David and Alan B. Krueger (1994), 'Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania', *American Economic Review*, 84 (4), 772-793. Card and Krueger's heretical attack was based on their examination of the impact on employment in the low-skill service industry from the passage of a New Jersey bill that increased the state minimum wage (a classical example of a price floor) to \$5.05 – eighty cents above the federal minimum wage level – in the early nineties. They compared employment, wages and prices at fast-food restaurants in New Jersey to restaurants in eastern Pennsylvania, which serve as a natural control group given the economic and demographic similarities between the two areas.

Based on estimates by previous researchers of the elasticity of low-wage employment to the minimum wage, the scheduled 18% increase in the New Jersey minimum wage would be expected to cause the number of workers per fast-food restaurant to decline by from 0.4 to 1. Card and Krueger, however, find no indication that the rise in the minimum wage reduced employment. Indeed, nine months after the effective date of the minimum wage increase in New Jersey, Card and Krueger found that, compared to Pennsylvania restaurants, New Jersey fast-food restaurants had *added* 2.7 full time workers. The basic finding of an *increase* in employment held across specifications and was robust to a slew of sensitivity checks.

While such a finding contradicts the simple price theoretic prediction of an increased minimum wage under perfect competition, other theories are compatible with an increase in employment. Specifically, under monopsony, wages and employment are both suppressed below the competitive level, so a price increase that moves wages toward the competitive wage will expand employment. A similar result occurs in the general job-search model, where workers are actively searching employers for jobs. If the wage increases, that firm should have a larger steady state labor force because it can attract more new workers while losing fewer workers to other firms. In this case, if the minimum wage increased, one might expect employment in the fast food industry to rise. But both of these non-competitive models would predict that output prices should drop in New Jersey relative to Pennsylvania. In fact, prices for fast-food products increased faster in New Jersey than in Pennsylvania, apparently reflecting the employer's ability to pass at least some of the additional cost of the minimum wage onto the consumer. Thus, none of the models seem to lead to predictions that can be fully reconciled with the data.

Not surprisingly, Card and Krueger's findings spurred many critical responses. A mere six months after Card and Krueger's original publication, Neumark and Wascher published a working paper in which they concluded the opposite of Card and Krueger namely, that "...the New Jersey minimum-wage increase led to a relative decline in fast-food employment in New Jersey" if compared to the control group in Eastern Pennsylvania. In the final version of their "comment" on Card's and Krueger's original paper -- Neumark, David and William Wascher (2000), 'Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment', *American Economic Review*, 90 (5), 1362-1396. -- Neumark and Wascher conclude, based on a newly created data set, "that the New Jersey

minimum-wage increase led to a 3.9 percent to 4.0 percent decrease in fast-food employment in New Jersey relative to the Pennsylvania control group..." Neumark and Wascher gathered firm payroll records from fast food restaurants because they deemed the phone-survey data of Card and Krueger unreliable. In particular, Neumark and Wascher point to the high variability of employment changes in the Card and Krueger survey- data which they interpret as sign of extreme measurement error that marred Card and Krueger's estimates.

In a reply to Neumark and Wascher's reevaluation, Card and Krueger re-visited their analysis while replacing their initial survey-data with payroll data provided by the Bureau of Labor Statistics -- Card, David and Alan B. Krueger (2000), 'Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply', *American Economic Review*, 90 (5), 1397-1420. They still find faster employment growth in New Jersey than in the control group – although in most specifications these results are not significant. Furthermore, Card and Krueger raise questions about the data utilized by Neumark and Wascher, which had been supplied to them by the Employment Policies Institute (EPI). Card and Krueger argue that the EPI collected this data from a small and unrepresentative group of restaurants. Overall, after evaluating all the available data, they conclude the following: "The increase in New Jersey's minimum wage probably had no effect on total employment in New Jersey's fast-food industry, and possibly had a small positive effect."

While the question of the true effect of the increase of the minimum wage in New Jersey is still debated, the importance of the original Card and Krueger paper is clear: it has sparked the re-examination of beliefs held by generations of economists and policy makers. Their general finding that a modest increase in the minimum wage does not necessarily reduce overall employment, while still in need of ultimate validation, has had an impact on minimum wage legislation in the U.S. and elsewhere. Of course, Card and Krueger's results can be harmonized with price theory if the increase in minimum wages stimulates the income of the class of individuals who like to patronize fast-food restaurants. In this case, it is the stimulus in the demand for the product that Card and Krueger evaluate that offsets the dampening effect on the quantity demanded of labor from the increase in the minimum wage (an important input in the fast-food industry).

#### **B.** Mandated Maternity Benefits

While the most common benefit that governments mandate for workers is probably a minimum wage, many other benefits can also be required by law. Jonathan Gruber focuses on the impact on wages of mandated maternity benefits. **Gruber**, **Jonathan (1994), 'The Incidence of Mandated Maternity Benefits'**, *American Economic Review*, **84 (3), 622-641.** Specifically, Gruber explores whether adoption of required maternal leave benefits led to a downward adjustment of the female relative wage, which would reflect a transfer of the benefits' cost to the group that values them most. In addition, Gruber estimates whether and how women respond in terms of their labor supply. Exploiting the natural experiments produced by variation in states' adopting mandated maternity benefits as well as the 1978 passage of the federal Pregnancy Discrimination Act (PDA), he finds that nearly all of the mandate's costs are borne by the target group and, therefore, these mandates have virtually no effect on female labor supply decisions.

Gruber estimates that in the late 1970s, nearly half of all women either did not receive maternity benefits or faced differential coverage relative to other health care needs. Gruber calculates the expected cost of giving birth (\$767) as well as the expected cost to a firm of adding maternity benefits (\$984) to establish the significant expense of expanding health care coverage. Gruber argues that if this government mandate does not lead to lower wages, then the mandates can be no more efficient than the distortionary taxes used to finance public benefits.

Relying on a differences-in-differences-in-differences (DDD) estimation strategy, Gruber's empirical model compares the labor market outcomes of "treated" individuals against "control" subjects within states that passed mandated benefits laws and then draws the same comparison between states that did and did not pass such laws. The treatment group is taken to be married women between the ages of 20 and 40, and the control population covers all persons over 40 and unmarried men aged 20-40. This nonparametric approach is supplemented by a finer measure of the mandate's expected cost using the estimates described above. Gruber used only Illinois, New Jersey and New York as the treatment states before adoption of the PDA (which become the controls in the second experiment) because of limitations in the CPS data.

Gruber's DDD estimate of the impact of mandated benefits (controlling for other demographic and experience characteristics as well as fixed and time-varying effects), suggests that the relative wage of married women of child-bearing age fell by a statistically significant 4.3 percent. In addition, Gruber finds that a rise in hours worked, a fall in employment, and small overall changes to labor supply accompanied this wage decrease. When he considers individually estimated mandate costs, the resulting coefficients point to 100 percent cost shifting to wages without any demonstrable effect on net labor input. Finally, estimation of the DDD regression using federal passage of the PDA to identify the effects of mandated benefits confirms the state-level conclusions: wages fall by about 2 percent, hours worked increases, employment decreases and the net effect on labor supply is negligible. Gruber concludes that, despite legislation protecting against differential pay on the basis of sex, maternity benefits drive a wedge between male and female wages.

#### **C. Mandating Accommodations**

The third paper to examine the impact of governmental directives to provide workers with particular benefits is **Jolls, Christine (2000), 'Accommodation Mandates',** *Stanford Law Review*, **53 (2), 223-306.** Jolls tries to provide a comprehensive economic framework that can be used to systematically analyze the distributive effects of mandates on accommodated workers relative to nonaccomodated workers. So, for example, Jolls uses her framework to predict that mandates that accommodate disabled workers will result in unchanged or increased relative wage levels but decreased relative employment. In the case of mandates that accommodate female workers, she predicts a decrease in relative wages along with an ambiguous effect on relative employment levels. Jolls then argues that the empirical evidence roughly supports her predictions. Jolls' analytical framework builds upon Lawrence Summers' labor supply and demand model, which applies to mandates directed at workers as a whole. Jolls extends Summers' framework so that it models two distinct, yet interconnected, classes of workers – those who receive the accommodation and those who do not.

Critical to Jolls' framework is the fact that most groups to which accommodation mandates are directed are simultaneously protected by anti-discrimination laws designed to prevent discriminatory treatment in the payment of wages and in hiring and firing. Consequently, the effects of mandates will depend on how much employers' behavior is constrained by the wage and employment components of anti-discrimination law. When the wage and employment discrimination components are truly binding, mandates' costs are distributed over the entire labor market. This implies that some of the costs of mandates are carried by nonaccommodated workers. Thus, the mandate might still prove distributionally advantageous to accommodated workers are smaller than the costs imposed on the rest of the workforce.

If, however, *only* the equal wage component of antidiscrimination law is binding, the costs of mandates falls fully upon accommodated workers in terms of a decline in their relative employment. Importantly, this decrease in employment occurs regardless of whether a mandate's value exceeds its costs. On the other hand, when wage laws are not binding, the cost of the mandate will be shouldered by accommodated workers primarily through lower relative wages; the employment effects will be ambiguous (depending on the cost-benefit ratio). Jolls notes that the presence of occupational segregation influences the overall effectiveness of antidiscrimination law and therefore can be an important element in using her analytical framework. Specifically, she argues that in the case of female workers, strong occupational segregation prevents the equal wage component of antidiscrimination law from having bite, since equal wages need only be maintained within a narrow labor market. In contrast, disabled workers are generally occupationally integrated, which by itself makes wage anti-discrimination laws more easily enforceable. However, since disabled workers represent a relatively small fraction of the labor force, the small numbers problem makes employment discrimination difficult to prove, undermining the power of antidiscrimination law for these workers.

Having laid out the theoretical model, Jolls proceeds to test the model's predictions about wages and employment against empirical evidence.<sup>8</sup> First, she selects three accommodation mandates: the Americans with Disabilities Act (ADA), state laws that require health insurance plans to provide for maternity-related expenses, and the Family and Medical Leave (FMLA). For each, she identifies the accommodated group targeted by the mandate's provisions, derives her model's predictions for each provision, and compares her predictions to the relevant empirical evidence. For those provisions

<sup>&</sup>lt;sup>8</sup> Jolls notes that the empirical analysis is complicated in that the most readily available data is aggregated across the accommodated groups as a whole, as opposed to accommodated groups *within* individual labor markets, which is what her theoretical framework assesses. This forces Jolls to treat a group as large as female workers as a homogenous entity. For example, her analysis is based on the assumption that female workers are in segregated markets. Although a majority is, a not insignificant percentage of females work in male-dominated and mixed labor markets. For these women, the effects of an accommodation mandate might be more evident in employment levels. Furthermore, the mandates she studies are often overlapping, which further complicates the empirical effort to disentangle their individual effects.

that apply to female workers, her model predicts lower relative wages with unchanged employment levels, which she concludes is largely consistent with the findings of Gruber's 1994 study. In the case of mandates directed at workers with disabilities, Jolls' framework predicts approximately equal wages along with declines in relative employment, which is supported by both Acemoglu and Angrist's (1999) study of the effects of the ADA and Ruhm's (1998) study of European mandates.

#### **D.** Training

Human capital theory is one of the pillars of the neoclassical economic theory of labor. One major insight – first analyzed fully by Gary Becker -- is the distinction between general and specific human capital. General human capital enhances productivity independent of the firm while specific human capital enhances productivity only in the firm in which an employee is currently working. Becker theorized that firms would never invest in general skill training because the worker would reap the complete benefit from his general training through higher wages. This claim follows directly from the assumption of competitive labor markets in which workers' wages are determined by their productivity.

While Becker's theory is broadly true – for example, law firms don't pay for smart undergraduates to go to law school – there are situations where firms seem to be paying for certain programs designed to enhance general human capital. In their "Beyond Becker" paper, Acemoglu and Pischke focus on the areas in which the predictions of conventional training theory based on Becker's general/specific skill model seem to be erroneous -- Acemoglu, Daron and Jörn-Steffen Pischke (1999), 'Beyond Becker: Training in Imperfect Labour Markets', *Economic Journal*, 109 (453), 112-142. In order to explain firm-sponsored general training of employees, they relax Becker's assumption of a competitive labor market. Acemoglu and Pischke argue that noncompetitive models seem better suited to explain apprenticeships, temporary-help training and firms sending their employees to MBA programs.

Acemoglu and Pischke argue that labor market imperfections change what Becker would consider to be general skills into "*de facto* specific skills" and that in several non-competitive scenarios firms are willing to invest in the general training of their employees. The common feature in their examples is that wages are below productivity and the wage structure is compressed, which means that the gap between wages and productivity (firm rent) increases with increased general training.

According to Acemoglu and Pischke, the following factors allow firms to profit by providing general training to their workers: 1) High job search costs, which matchspecific surpluses that allow the firm to obtain – through bargaining - some of the workers' productivity as profit. 2) Asymmetrical information between the firm and outside firms about the employee and asymmetrical information between the firm and its employees. 3) Labor market institutions, such as minimum wages, unions, or progressive unemployment benefits. Minimum wages, for example, give firms incentive to train their sub-minimum wage productivity workers without having to pay them more. 4) If general and specific skills are complements – which according to Acemoglu and Pischke is most often the case - then an increase in general skills will also increase the value of specific skills. While Acemoglu and Pischke concede that they do not yet have the necessary understanding to make sensible policy recommendations, they feel confident enough in their non-competitive models to assert that Becker's recommendation – easing the liquidity constraints of workers – is not sufficient to ensure optimal training. Contrary to Becker and the standard theory built around his model, Acemoglu and Pischke see a potentially positive impact on investment in human capital from subsidies and regulations. In particular, they note that "different training systems may make different labor market regulation regimes optimal." They also conjecture that countries that subsidize the training of low-skilled workers have enabled them to take advantage of new technologies, thereby explaining why wage inequality did not increase in those countries as it did in most others.

#### E. Employment at Will and Wrongful-Discharge Laws

The endowment effect, or status quo bias, refers to a propensity to value something more highly if one possesses it initially than if someone else possesses it. This effect can explain why individuals are loss averse – they are far more unhappy to lose something they currently have than they are pleased to gain something they previously did not have. When it comes to the labor market, it would seem that these psychological effects are much more potent in Europe than in the U.S. and may explain the striking contrast between the U.S. and Europe with respect to the issue of job security. Europeans tend to have much greater protections against discharge than American workers, who traditionally could be fired for any reason or no reason under the doctrine of employment at will. But the dramatic protections of European workers come at a price: Europeans who have jobs have great security but those who don't have a much lower chance of securing employment than, say, unemployed Americans would.

In the U.S., however, the trend to provide protections against certain unfair discharges has been growing for the last quarter century in the form of a series of exceptions to the doctrine of employment at will. Richard Epstein has been unhappy with this direction of the law, and has attempted to revive support for the traditional U.S. employment relationship -- Epstein, Richard A. (1984), 'In Defense of the Contract at Will', *University of Chicago Law Review*, 51 (4), 947-982. Epstein reasons that the intended benefits ascribed to these exceptions actually impose more disorder in the legal system through increased complexity and litigation. Epstein argues that employment at will contracts actually serve the interests of both employers and employees rather than promote exploitation.

Concerning fairness, Epstein argues that both parties to a contract should be free to enter into an at-will agreement if they deem it to be the optimal contract. This libertarian argument rejects the need for intrusive governmental mandates over what types of economic arrangements individuals may devise. Citing the high frequency of atwill contracting, Epstein concludes that agents must be acting rationally in choosing such contracts. Moreover, he expresses faith in the ability of agents -- especially employees -to protect themselves from predatory action, and assigns scant weight to reports of fraudulent or coercive behavior in litigation claims.

Epstein is persuaded that utility is enhanced by the employment at will doctrine as confirmed by its widespread adoption of at-will contracts when labor markets are free.

He discounts cases in which the naïvete of one or both parties actually cuts against claims of rational behavior. Seeking to explain the empirical fact that agents freely enter into atwill contracts, he draws an analogy with simple partnerships and the mechanisms used by either side to prevent abuse of the relationship and rent seeking. The thrust of the comparison is that partners -- like employers and employees -- retain a bilateral monopoly over their contributions to the business venture. Consequently, the shared threat of withdrawal, which at-will contracts permit, serves as a check against exploitative behavior. With respect to the problem of bilateral monopoly, he maintains that the fluidity of at-will contracts offsets the "hold-up" that ensues when either party attempts to take the employment relationship as a hostage. In addition, employees are free over the life-cycle of employment to diversify their labor supply when free entry and exit from jobs exist.

Finally, Epstein claims that issues of redistribution should not (or cannot) inform the debate over at-will contracts. Epstein argues that the increased litigation engendered by the exceptions to employment at will soaks up social and firm-specific resources, overall levels of resources diminish and these losses are spread among both firm owners and their employees. In light of this consideration and in the absence of any clear gains from redistribution, Epstein counsels against relying on the abolition of at-will contracts to achieve such ends.

While Epstein defends the idea that employment at will should be the default rule, Stewart Schwab offers a justification for what he calls the current intermediate position of the courts that favors neither the at-will or for-cause default rule -- Schwab, Stewart J. (1993), 'Life-Cycle Justice: Accommodating Just Cause and Employment At Will', *Michigan Law Review*, 92 (4), 8–62. Schwab terms this background presumption the life-cycle default rule. Focusing primarily on the case of the career employee, Schwab argues that this flexible stance is the optimal one for courts to take today as it minimizes the risk of opportunism for both employers and employees by considering how the incentives for opportunistic firing and shirking vary in magnitude over the career relationship.<sup>9</sup>

According to Schwab, at-will employment allows employers to easily fire bad or unproductive workers, but if fails to protect workers from opportunistic firings, such as being discharged before receiving a commission or before a pension vests. On the other hand, a just-cause employment rule, such as those specified in many union contracts, protects the employee from arbitrary firings, but does not do much to prevent shirking. In contrast to these polar, unvarying rules, the life-cycle default rule affords greater flexibility to accommodate the changing risks of opportunistic behavior throughout the career relationship. A just cause background rule would seem most appropriate in the early- and especially late-career stage when the risk of opportunistic firing is arguably the greatest. On the other hand, an at-will default rule would work well in the mid-career

<sup>&</sup>lt;sup>9</sup> Schwab notes that career employment has been a relatively recent phenomenon, becoming much more common after World War II and the introduction of pensions. As such, he does not dispute Epstein's claim that at-will employment was the optimum for much of the previous century when it was not common for workers to stay with a single employer for their entire working career. However, he argues that as the employment relationship has moved toward career employment, the common law has changed along with it, approaching the "life-cycle doctrine in employment law."

stage when the risk of employees shirking on the job is highest. Schwab finds that courts have done this in practice: intervening to protect employees when the danger of employer opportunism is high, while upholding the presumption of at will-employment when the risk of an employee shirking is high.

Schwab argues that these interventions by the courts are essential for policing the opportunistic incentives created by factors such as efficiency wages, high monitoring and training costs, and vague contracts. For example, a firm might pay a worker a higher efficiency wage to induce higher levels of effort. However, as the worker ages and the value of his wage is perceived to exceed his productivity, the employer has an incentive to let him go. On the other hand, high monitoring and training costs prohibit employers from easily replacing current employees, making it easy for workers to shirk. Nevertheless, supporters of a simple default rule of either at will or for cause point out that there are a number of internal mechanisms and social norms that keep such opportunism in check. They argue that an employer's desire for a good reputation both within and outside the firm would keep him from firing a productive worker. Schwab acknowledges that such mechanisms do exist, but notes they are far from perfect. Young or new workers at a firm may have difficulty observing how an employer treats more senior workers. Additionally, while higher level managers may want to keep productive employees, "low level supervisors" could become involved in petty disputes that could lead to the discharge of a productive worker.

Schwab's examination of common law cases supports his hypothesis of a lifecycle default rule. He cites several general examples of courts upholding the duty of good faith to protect employees against opportunistic behavior by an employer. For example, in the case of Fortune v. National Cash Register Co., the courts imposed a duty of good faith on the employer to prevent the firing of a salesman who was about to receive a commission for equipment installed in his territory.<sup>10</sup> Schwab claims that courts have moved in the direction of a life-cycle default rule by protecting early- and late-career employees from employer opportunism while generally not interfering in atwill cases involving midcareer employees. Schwab cites the case of Grouse v. Group Health Plan, Inc. as an example of the courts protecting a beginning-career employee from employer opportunism. A pharmacist had quit his job on the promise of a job offer from a health clinic, but upon his arrival, the clinic told him they had filled the position. The court determined that the pharmacist had "reasonably relied on the job offer," and thus upheld his claim. Notably, despite such evidence, Schwab also acknowledges the ambivalence of courts in protecting early-career employees. The courts must consider both the new employee who incurs substantial moving costs and an employer who requires time to determine whether or not new worker is actually a good hire.

Court interventions in late-career employment terminations seem more common as the employer has greater incentive to discharge an employee. In addition, the Age Discrimination in Employment Act or ADEA helps to protect employees from

<sup>&</sup>lt;sup>10</sup> Epstein has called this decision "wrong in principle," and argues that this was not a case of employer opportunism as the money from the commission was instead used to pay an installations employee who installed the equipment. Schwab counters by saying that the Fortune decision may be "wrong in application," but it is correct in principle, because it provides a precedent that "courts should scrutinize opportunistic firings in which the employee has largely performed his side of the bargain but has yet to reap his reward."

opportunistic discharge based on the employee's age. Cases such as Foley v. Community Oil Co., where the court upheld the claim of an employee who was fired after 30 years of service, is one example where courts take into account the considerable personal and economic sacrifices an employee makes when he commits to a career relationship with a single firm.

Schwab's hypothesis that courts have indeed upheld the at-will doctrine during the midcareer stage is buttressed by his finding that "midcareer employees have made the fewest contributions to the doctrinal erosion of at-will employment."

In the end, Schwab concludes that the life-cycle default rule is the best rule for both parties. As its flexibility minimizes opportunistic risks on both sides of the bargaining table, it allows employers and employees to invest more in the career relationship, making it more productive. He also notes that it is easier to bargain away from the life cycle rule (i.e. to an at-will or a for-cause contract), than to move toward it. However, as Schwab acknowledges, an important limitation to his analysis is that it is focused on the life-cycle of *men* who generally have the greatest job security in their midcareer years. In contrast, women often face the greatest risk of opportunism mid-career when they may take time off of work to have or raise children.

While Epstein marshals a strong theoretical argument against the trend against the doctrine of employment at will, it is important to examine the empirical evidence on the impact of the adoption of exceptions to this doctrine. Autor, Donohue, and Schwab do just that in Autor, David H., John J. Donohue and Stewart J. Schwab (2006), 'The Costs of Wrongful-Discharge Laws', The Review of Economics and Statistics, 88 (2), **211-231.** This paper notes that the exceptions to the doctrine of employment at will fall into three categories. The public policy exception, which has been adopted in 43 states, prohibits discharges that undermine an explicit public policy of the state. For example, an employee who was discharged because he or she would not commit perjury on behalf of the employer can sue for this violation of the public policy exception. The good faith exception, adopted by 13 states, is designed to protect an employee from being deprived of a major benefit by the bad faith conduct of an employer. For example, an employer who fires a worker just before a pension will vest purely to deprive the worker of this benefit violates the good faith exception. Finally, the implied contract exception, adopted in 43 states, says that language in employee handbooks and manuals can, under certain circumstances, create a contractual right to protect against discharges that are without just cause.

Autor, Donohue, and Schwab analyze a panel of CPS data showing monthly employment rates (the ratio of employment to population) for all fifty states for the period from 1978 – 1999. Epstein's criticism of the implied contract doctrine does seem to have some force in that state employment rates appears to drop by between eighttenths of 1 percent and 1.7 percent when this exception is adopted. The initial impact is largest for female, younger, and less educated workers – all of whom tend to change jobs frequently. The somewhat longer-term effect is greater for older and more-educated workers, who are the ones most likely to litigate. Epstein's fears about the other exceptions may be misplaced, as the adoption of the good faith and public policy exceptions, which admittedly have a more narrow scope than the implied contract exception, appear not to have any dampening effect on employment. Autor, Donohue, and Schwab find that despite the apparent backward shift in the demand curve that results from the adoption of the implied contract exception, there is no evidence of a drop in wages. Indeed, if workers value this exception, one would expect an outward shift in the supply curve that would further dampen wages. The lack of any wage effect may suggest that the employment protection that results from the implied contract exception may increase the bargaining power of incumbent workers, thereby offsetting any dampening effect on wages from a backward demand shift.

#### V. Employment Discrimination

One important attribute of competitive markets is that they are supposed to ensure the rough equation of prices and value. Because capital markets are highly competitive and have relatively low-transaction costs, the efficient capital markets hypothesis posits that stock prices will reflect all publicly available information that bears on the value of the firm. As Donohue (1994) discusses, labor markets have far higher transaction costs than capital markets, and the pressures moving wages to value are far less potent.<sup>11</sup> When one contemplates the history of labor markets in the U.S. - and of course many other countries – discrimination has been an important factor in preventing workers from achieving wages that reflect their inherent productivity. The original goal of employment discrimination law in the United States was to eliminate this disparity by increasing the earnings of certain disadvantaged groups whose employment prospects were hampered by discrimination. Today, some argue that the goal of mimicking the outcome of perfectly competitive labor markets is insufficient and that employment discrimination law should more aggressively pursue broader goals of social fairness that will enhance the economic status of disadvantaged groups beyond what a perfect market would provide.

#### **A. Racial Discrimination**

The massive and indisputable employment discrimination against blacks and women became unlawful throughout the country in 1964 with the adoption of Title VII of the Civil Rights Act of 1964. Congress later broadened the coverage of this statute when it enacted the Equal Employment Opportunity Act (EEOA) of 1972, and then further expanded federal antidiscrimination law (primarily by providing greater damage remedies for successful sex discrimination plaintiffs and workers discharged because of their race) in passing the Civil Rights Act of 1991.

The 1964 Act has received the most scholarly attention for it was clearly the most momentous piece of antidiscrimination law ever enacted. Milton Friedman had been a strong opponent of such antidiscrimination law, arguing in part on the basis of Gary Becker's work that such laws would not be needed since competitive markets drive discriminators – and others that fail to maximize profits -- out of the market. But when the federal law was passed, it appeared that Friedman was wrong: blacks enjoyed

<sup>&</sup>lt;sup>11</sup> Donohue, John (1994), 'Employment Discrimination Law in Perspective: Three Concepts of Equality', *Michigan Law Review*, 92, 2583.

substantial economic gains, particularly in the South. Initially, James Smith and Finis Welch try to carry Friedman's banner by arguing that the Civil Rights Act of 1964 was not responsible for these gains in black economic welfare.<sup>12</sup> Instead, they contended, the gains were all the result of human capital enhancement, not of demand side stimulation resulting from decreased discrimination: blacks had been adding to their low skill levels and modest levels of education, and as they secured more human capital their wages rose appropriately. Smith and Welch argued that the economic gains of blacks were no different during the period from 1940 through 1960 than they were in the following two decades. They took this as evidence against the view that Title VII generated any benefits for black workers.

The major response to Smith and Welch came in **Donohue**, John J. and James Heckman (1991), 'Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks', *Journal of Economic Literature*, 29 (4), 1603-1643. Donohue and Heckman argued that Title VII did indeed generate a decade of economic gains for blacks. As Donohue and Heckman note:

"the evidence of sustained economic advance for blacks over the period 1965-1975 is not inconsistent with the fact that the racial wage gap declined by similar amounts in the two decades following 1940 as in the two decades following 1960. The long-term picture from at least 1920-1990 has been one of black relative stagnation with the exception of two periods – that around World War II and that following the passage of the 1964 Civil Rights Act."

It is now widely accepted that in helping to break down the extreme discriminatory patterns of the Jim Crow South, Title VII did considerably increase the demand for black labor, leading to both greater levels of employment and higher wages in the decade after its adoption.<sup>13</sup>

Ken Chay wrote an important paper attempting to determine whether the EEOA, which broadened the coverage of Title VII in 1972, provided additional independent stimulus beyond that provided by the initial Civil Rights Act of 1964.<sup>14</sup> Chay used the fact that the EEOA had a predictably different impact across industries and between the South and the non-South as a way to estimate the economic consequences for blacks of this strengthening in the federal antidiscrimination law. Prior to 1972, Title VII's prohibition against employment discrimination only applied to firms with 25 or more employees. The Equal Employment Opportunity Act (EEOA) of 1972 lowered this

<sup>&</sup>lt;sup>12</sup> Smith, James P. and Welch, Finis. (1989) 'Black Economic Progress After Myrdal', *Journal of Economic Literature*, 27 (2), 519-64.

<sup>&</sup>lt;sup>13</sup> See also, Freeman, R.B., R.A. Gordon, D. Bell and R.E. Hall (1973), 'Changes in the Labor Market for Black Americans, 1948-72', *Brookings Papers on Economic Activity*, 1, 67-131; Conroy, M. (1994), *Faded Dreams: The Politics and Economics of Race in America*, (Cambridge University Press, Cambridge); and Orfield, G. and C. Ashkinaze (1991), *The Closing Door: Conservative Policy and Black Opportunity*, (University of Chicago Press, Chicago).

<sup>&</sup>lt;sup>14</sup> Chay, Kenneth. (1998) 'The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Employemnt Opportunity Act of 1972', *Industrial and Labor Relations Review* 51(4):608-32.

threshold to include employers with 15 to 24 employees. Moreover, many states already had fair employment practice (FEP) laws that covered these employers, so if the legal prohibition in these states was as effective as the federal prohibition, then the EEOA would be redundant in those states. Of the nine states that did *not* have FEP laws before 1972, eight were in the South. Based on his careful empirical work, Chay concludes that the EEOA increased the demand for black workers among small employers not previously covered by FEP laws.

But if federal antidiscrimination laws adopted in 1964 and 1972 yielded important economic gains for blacks, the same cannot be said about the final expansion that occurred in 1991. In a series of papers, Paul Oyer and Scott Schaefer demonstrate that there is little support for the view that the strengthening of federal antidiscrimination law in 1991 stimulated black or female employment.<sup>15</sup> Indeed, James Heckman who was a major figure opposing the view that labor markets in the 1960s were fully protecting workers against discrimination has emphasized that the labor market is doing a much better job today at rewarding skills than it did a half century ago. Heckman no longer believes that market discrimination substantially contributes to the black-white wage gap (as it once clearly did), and therefore he doubts that at present racial discrimination in the labor market is a first-order problem in the United States.<sup>16</sup> Rather, Heckman looks to other factors (i.e., those that promote skill formation) to explain the black-white earnings gap – a theme that he builds on in Carneiro, Heckman, and Masterov (2005).<sup>17</sup>

An important paper that informs Heckman's analysis of the current reasons for the black-white wage gap is **Neal, Derek A. and William R. Johnson (1996), 'The Role of Premarket Factors in Black-White Wage Differences', Journal of Political Economy, 104 (5), 869-895.** If factors that exist prior to workers' entry into the labor market largely explain the black-white wage gap, then the contribution of racial discrimination to this wage gap is presumably small. Neal and Johnson note that many studies have examined the black-white wage gap and found that it could not be explained with standard measures such as age, years of education, marital status, etc., creating the inference that the contribution of discrimination was sizeable. Neal and Johnson note that years of education may exaggerate the true skill level attained by blacks, given the poorer quality schools that many blacks attend. They argue that scores on the Armed Forces Qualification Test (AFQT) are a better measure of acquired skill (rather than innate ability) that one brings to the labor market.

The authors use a log-linear model that regresses the log of hourly wages on a number of demographic and educational variables. The *unadjusted* wage gap between black and whites is -24.4% for black men and -18.5% for black women. A significant

<sup>&</sup>lt;sup>15</sup> See Oyer, P. and S. Schaefer (2000), 'Layoffs and Litigation', *Rand Journal of Economics* 31(2):345-358; Oyer, P. and S. Schaefer (2002a), 'Litigation Costs and Returns to Experience', *American Economic Review* 92(3):683-705; and Oyer, P. and S. Schaefer (2002b) 'Sorting, Quotas, and the Civil Rights Act of 1991: Who Hires When It's Hard to Fire?', *Journal of Law and Economics* 45(1):41-68.

<sup>&</sup>lt;sup>16</sup> Heckman, James J. (1998), 'Detecting Discrimination', Journal of Economic Perspectives 12(2):101-16.

<sup>&</sup>lt;sup>17</sup> Carneiro, Pedro, James Heckman and Dimitriy Masterov (2005). 'Labor Market Discrimination and Racial Differences in Premarket Factors', IZA Discussion Paper no. 1453. ftp://ftp.iza.org/dps/dp1453.pdf

proportion of the respondents to the National Longitudinal Surveys of Youth (NLSY) took the AFQT in 1980. Using this data for those who took the exam before entering the labor market, Neal and Johnson found that the unexplained wage gap in their regressions containing controls for race, age, and AFQT score is -7.2% for black men and +3.5% (although insignificant) for black women. In other words, the AFQT test score can explain a very large portion of the black-white wage gap for men, and all of the gap for women. One source of continuing debate in the literature is whether these wage regressions should include controls for years of education as well as AFQT score. Neal and Johnson say it should not since the test better captures ability, and so they exclude the education measure from their regressions. Others have included years of education and find that the unexplained wage gap re-emerges when this control is added.

Of course, Neal and Johnson recognize that their analysis would be compromised if the AFQT were racially biased, but they cite a National Academy of Sciences report to negate this charge. Another potential problem with their conclusion that the impact of discrimination is small is the possibility that statistical discrimination could lead to black underinvestment in human capital. Neal and Johnson instead find that the return to higher AFQT scores is significantly higher for black men (although not for black women), so the incentive to invest in developing human capital seems to be high enough to undermine the argument based on statistical discrimination.

#### **B. Sex Discrimination**

As the now-former President of Harvard University, Larry Summers, learned, few issues are as sensitive as the issue of sex discrimination in employment. After a largely (although not entirely) nuanced and sophisticated address to an NBER Conference on Diversifying the Science and Engineering Workforce (January 14, 2005), Summers closed with the following controversial summary about why one sees an under-representation of women in the most elite academic science and engineering positions:

So my best guess, to provoke you, of what's behind all of this is that the largest phenomenon, by far, is the general clash between people's legitimate family desires and employers' current desire for high power and high intensity, that in the special case of science and engineering, there are issues of intrinsic aptitude, and particularly of the variability of aptitude, and that those considerations are reinforced by what are in fact lesser factors involving socialization and continuing discrimination. I would like nothing better than to be proved wrong, because I would like nothing better than for these problems to be addressable simply by everybody understanding what they are, and working very hard to address them.

While Summers was criticizes for his expressed opinion that sex discrimination was not the primary factor explaining the shortfall of women in science at elite institutions, he certainly would not have disputed that substantial discrimination against women was once widespread. As with the issue of race discrimination, however, there is greater debate about the extent of the problem today. Discrimination is always difficult to prove, but one landmark study of a design that provides credible evidence of sex discrimination in employment twenty to thirty years ago is **Goldin, Claudia and Cecilia** 

## Rouse (2000), 'Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians', *American Economic Review*, 90 (4), 715-741.

Goldin and Rouse examine labor market discrimination in the context of auditions and hiring of musicians for the major U.S. orchestras. To test for sex discrimination in the hiring process, they exploit the changes in the audition process introduced by all major U.S. orchestras in the 1970s and 1980s. Of particular interest for their study was the change to "blind" auditions, which effectively hid the identity and gender of the applicant from the hiring committee for certain rounds of the audition process. Using audition and roster data spanning several decades and employing an individual fixed effect strategy, they find that the likelihood of female hiring and advancement is increased by the introduction of blind auditions.

More specifically, using audition data from the late 1950s through 1995, Goldin and Rouse found that in blind audition rounds women were as much as 50 percent more likely to advance from preliminary to final rounds. Furthermore, the likelihood of women winning the finals increased by 33 percentage points if the final round was blind. Using official roster data from 1970 to 1996, they find that completely blind auditions – defined as auditions in which all rounds are conducted with a screen hiding the gender of the applicant – increased the likelihood of a women being hired by 25 percent. Based on the roster data, blind auditions explain 30 percent of the increase in female hiring and 25 percent of the increase in overall female representation in the orchestras. There are, however, some caveats with respect to these findings: first, some estimates are associated with relatively large standard errors that render them statistically insignificant; second there is one scenario – auditions with blind semifinals – in which the effect on females is persistently strongly negative.

This latter finding is potentially troubling to the Goldin-Rouse thesis. Some auditions in the study had a semifinal round, often held on the same day as the preliminary round. This gave judges another chance to hear the contestants before making the decision to advance them to the final round. In the study, having a blind audition in the semifinal round was found to have a strong *negative* effect on the probability of a female musician advancing to the final round. The authors offer one possible explanation: the non-blind semifinals may provide an opportunity for some form of affirmative action. If audition committees "actively seek to increase the presence of women in the final round," and only do so if the woman is above a certain level of quality, then a blind semifinal round could actually have a negative effect on the probability of a female musician advancing to the finals.

To implement the fixed effect strategy, Goldin and Rouse limit the original audition sample to musicians that competed more than once and entered both blind and not-blind audition rounds. The most extensive specification controls for automatic placement, the number of previous auditions, years since last audition, total number of musicians competing in the round, proportion female in the round and type of the position. Also included are year and instrument fixed effects and a dummy for auditioning for one of the "Big Five" orchestras.

Goldin and Rouse address several potential biases that arise from their fixed effect strategy. First, they include time-varying individual covariates to deal with the fact that female musicians who improve over time faster than male musicians seem to be switching from not-blind to blind auditions. Second, the fixed-effect strategy excludes musicians that are hired (or discouraged) after their first audition. Goldin and Rouse point out that this is not an issue because so few musicians are actually hired in a given year. Nonetheless, they control for the number of auditions and show that the estimates are not significantly different when the sample is limited to musicians that compete in at least three auditions. Third, their results could potentially be biased because orchestras that introduce blind auditions may be intrinsically less discriminatory. To address this, Goldin and Rouse establish that orchestra fixed effects do not change previously established estimates. Fourth, potential bias introduced by measurement error due to sex misclassification is assessed. Goldin and Rouse document that their results remain intact when they rerun their estimation and employ the census probability distribution on the gender of names instead of their subjective labeling.

While the data is highly imperfect, another factor that undermines women in the labor market is the large percentage of working women who experience sex harassment on the job.<sup>18</sup> The result of such harassment is reduced job satisfaction, higher absenteeism, adverse health outcomes, increased job turnover, and lower productivity. A recent study by Antecol and Cobb-Clark criticizes the existing literature for being not systematic enough and therefore making comparison of findings and general inferences difficult.<sup>19</sup> For example, there does not seem to exist an agreed upon definition of sexual harassment and most studies rely on small, non-representative samples. Antecol and Cobb-Clark try to rectify the sampling shortcomings by using a large scale data set spanning 15 years and various federal agencies. They find unwanted sexual behavior is increasingly likely to be considered sexual harassment and attribute this change to structural changes in attitudes about what constitutes sexual harassment. Some other broad conclusions that Antecol and Cobb-Clark mention in their literature review include: 1) Sexual harassment is common across employment sectors and observed in many countries. 2) "[T]he incidence of sexual harassment is related both to demographic characteristics and to the nature of one's employment," in particular organizational factors seem to be able to facilitate or inhibit sexual harassment. And 3) Sexual harassment seems to be widely underreported – less then 5% of sex harassment incidents are reported to anyone in authority. Formal complaints are even less frequent. Overall, there is much room for additional research on the extent of work-related sexual harassment, and whether law has played a role in dampening such conduct.

But while the critics of Larry Summers argued that all the problems of women in the labor market came from discrimination in all its forms, a growing body of literature is focusing on attributes of the women themselves. For example, the issue of gender differences in aptitude, specifically aptitude in competitive environments, is explored in

<sup>&</sup>lt;sup>18</sup> Fitzgerald, Louise F., and Alayne J. Ormerod. 1993. 'Breaking Silence: The Sexual Harassment of Women in Academia and the Workplace', in Florence L. Denmark and Michele A. Paludi, eds., *Psychology of Women: A Handbook of Issues and Theories*. Westport, Conn.: Greenwood Press, pp. 553-82. Schneider, Kimberly T., Suzanne Swan, and Louise F. Fitzgerald (1997), 'Job-Related and Psychological Effects of Sexual Harassment in the Workplace: Empirical Evidence from Two Organizations', *Journal of Applied Psychology*, 82 (3), August, 401-515.

<sup>&</sup>lt;sup>19</sup> Antecol, Heather, and Deborah Cobb-Clark. 2004 "The Changing Nature of Employment-Related Sexual Harassment: Evidence from the US Federal Government (1978-1994)." *Industrial & Labor Relations Review*, 0019-7939, April 1, 2004, Vol. 57, Issue 3

# an article by Uri Gneezy, Muriel Niederle, and Aldo Rustichini -- Gneezy, U., M. Niederle and A. Rustichini (2003), 'Performance in Competitive Environments: Gender Differences', *Quarterly Journal of Economics*, 118 (3), 1049-1074.

The authors seek to understand the relative dearth of women in high profile jobs as a major factor in the gender gap in earnings by looking at the performance of women and men in competitive environments. Unlike previous studies that tried to explain the gender gap either through occupational self-selection due to differences in abilities and preference or through employer discrimination, Gneezy et al explore the possibility of gender differentiated performance in competition. As the authors point out, such a difference in competitive performance could "reduce the chance of success for women when they compete for new jobs, promotions, etc." In a series of controlled experiments Gneezy et al. examine the performance of men and women in a computerized maze game as they vary the incentive schemes and group composition for different treatments. They find that while men receive a significant performance boost in competitive environments such as tournaments, the response of women in competitive environments is more nuanced: they do not significantly change their performance in mixed tournaments, but they do increase their performance in single-sex competitions.

In their experimental setup, the authors use the number of mazes solved as their measure of performance. Their subject pool is composed of male and female students at the Technion, a competitive engineering school in Israel. In their basic design, a group of six students, three men and three women, are each given fifteen minutes to solve as many mazes as they can. Varying the payment schemes and the gender composition of the groups, the authors conduct five different treatments, each replicated ten times with different participants. In total, the authors conduct 51 sessions with 324 participants and compare the performance distributions of different treatment groups.

In their benchmark noncompetitive treatment, Gneezy et al. administer a piece rate payment scheme on a group of three men and three women. They find that while men perform slightly better than women on average, there is no significant gender difference in performance. The mean performance of men in the piece rate treatment is 11.23 while that of women is 9.73, resulting in a mean gender gap of 1.4. However, when the authors introduce their main competitive treatment of mixed tournaments they find that men increase their performance significantly, while women's performance remains relatively unchanged. The mean performance of men in mixed tournaments increases to 15 while that of women barely changes to 10.8. While this increase for men is highly significant (p = 0.001), there is no significantly significant difference in female performance under the piece rate and mixed tournament setups (p = 0.62). The increase in the gender gap when moving from the piece rate treatment to the mixed tournament treatment – a jump from 1.4 to 4.2 – is also significant (p = 0.034).

Refining their analysis, the authors also examine additional reasons why a woman who solves the same number of mazes as a man in a noncompetitive environment may not receive the same performance boost as the man would in mixed tournaments. Specifically, the authors test for gender differences in risk aversion, the competitiveness of women in a single-sex environment, and for different self-perceptions of competence across genders.<sup>20</sup> In the case of risk aversion, the authors recognize that the tournament

<sup>&</sup>lt;sup>20</sup> Gneezy et al. also create hypothetical distributions of women's expected performance in mixed tournaments based on their performance in the noncompetitive treatment by comparing them to similarly

structure has two big differences compared to piece rate payment. First, payment is uncertain, and second, the final outcome only depends on *relative* performance. If women are more risk averse than men and effort is costly, then women might not expend as much effort in a mixed tournament as men would. Using a random payment scheme, where a single winner is chosen at random and paid according to his output, Gneezy et al. introduce a noncompetitive game with uncertainty. The performance of men and women in the random payment scheme is not significantly different from their performances under the piece rate structure, leading the authors to conclude that the gender gap is *not* due to differences in risk aversion.

Next, the authors introduce single-sex tournaments to test whether women dislike competition in general or if they simply dislike competing against men. They find that women are indeed competitive, at least in the single-sex tournaments, experiencing a significant performance boost when compared to the noncompetitive treatments (p = 0.0148 when comparing it to piece rate and 0.0469 when comparing it to random payment). Men also experienced increased performance when competing in single-sex tournaments, but this increase was not significantly different from their increased performance in mixed tournaments.

Finally, the authors look at a few explanations of why women and men of similar ability might perform differently under competition.<sup>21</sup> Specifically, they test the hypothesis that that women and men might not feel equally competent when presented with the task of solving mazes, which could affect performance in competition and contribute to the gender gap. Indeed, when the authors allow men and women to choose their level of difficulty (payment per maze solved increases with choice of difficulty), they find that men choose a significantly higher level than women do. The mean choice out of five levels of difficulty is 3.4 for males and 2.6 for females – a highly significant difference (p = 0.0065). Because the psychology literature has established clear links between "task choice and feelings of competence," it would appear that men feel more competent than women when faced with the task of solving mazes.

This paper is significant in that it implies that a gender gap in wages may not only be caused by discrimination or individual differences in ability and preferences, but by differences in performance between men and women resulting from the nature of the competitive environment. In addition, their findings have been offered to provide support

ranked men in the noncompetitive treatment. For example, if a man solved 15 mazes in the noncompetitive treatment and ranked 12-17th out of 60 (both piece rate and random pay treatments are combined here), then in the mixed tournament with 30 observations he would be expected to place between 6-9, and solve a corresponding 17-19 mazes. If women received a similar performance boost, a woman who solved 15 mazes in the noncompetitive treatment should be able to solve at least 17 mazes in the competitive treatment. But, in fact, the actual performance of women was much lower than projected, a difference that is significant at p = 0.04.

<sup>21</sup> The authors also note the additional explanation that men and women actually face different sets of competitors which could be drawn from slightly different ability distributions and affect participants' choice of effort: men face three women and two men, while women face three men and two women. As noted previously, men consistently perform slightly better than women on average in both the noncompetitive treatments and single-sex tournament and significantly outperform women in the mixed tournament. While they are not able to directly test this effect directly, authors calculate that a man a mixed group of three men and three women has a 0.07 to 0.1 greater probability of winning than a woman in the same group.

for the practice of single-sex schooling, as the women in this study demonstrate significant increases in performance in competitive, single-sex environments. A recent report by Goodman, Cunningham, and Lachapelle (2002) examining reasons for female attrition rates in engineering programs has found that women do not necessarily drop out because of poor performance, but rather many women cite "negative aspects of their school's climate such as competition, lack of support, and discouraging faculty and peers" Gneezy et al suggest that changing the nature of the environment has the potential to improve the performance of female workers and students.

#### **C. Statistical Discrimination**

A number of theoretical articles have explored whether statistical discrimination can play an important role in explaining the black-white earnings gap. One reason for skepticism about such an effect is that if, say, blacks are on average treated as their productivity would warrant, then as a class there should be no earnings shortfall, apart from the issue of underinvestment that was discussed above with the Neal and Johnson paper. An informative and insightful paper that explores the impact on the hiring and productivity of minority workers of moving from a more informal worker selection process to one based on standardized testing is **Autor, David and David Scarborough** (2004), 'Will Job Testing Harm Minority Workers?' NBER Working Paper No. 10763. Many have speculated that job testing presents an intrinsic "equity-efficiency trade-off": testing produces productivity gains but at the cost of adverse hiring effects on minorities. Given that minorities and underprivileged groups on average score lower on standardized tests, the potential of this equity-efficiency tradeoff exists

To explore this issue, Autor and Scarborough use employment data from a large, nationwide retail firm that instituted a standardized testing system in 1999. Before June 1999, the company used informal, paper applications to select candidates for line positions (entry level positions). Starting June 1999, the firm, throughout its outlets, began instituting a computer-based application system that included a personality test which the firm uses to select compatible and potentially productive candidates. Autor and Scarborough's sample of 34,257 observations contains information on test scores, worker demographics, termination date and termination reason (if applicable) for hires made between January 1999 and May 2000 in all of the firm's outlets. The question they address is how the introduction of testing and the ensuing improvement in the firms' applicant selection procedure affected minority hiring and productivity.

Autor and Scarborough illustrate that *pre*-testing hiring practices determine the effects of job testing. In other words, whether job testing has a negative impact on minority hiring depends on how hires where made in the absence of standardized testing. They argue that testing leads to reduced minority hiring if the pre-testing procedure is: 1) random and unsystematic or 2) the firm uses a systematic selection criterion *not* based on demographic characteristics. However, if employers statistically discriminate before the test is introduced—that is, if they already use demographic characteristics as a signal for expected productivity of the candidate—then adding testing to the model does not hurt minority hiring but still increases the average productivity of both minority and non-minority workers. The empirical evidence supports this last scenario—uniform increased productivity along with no negative effects on minority hiring.

Using various specifications and attempting to control for endogeneity concerns, Autor and Scarborough test for changes in productivity using two proxies, tenure length and the reason for job termination. The article finds a *uniform* increase in the productivity of all hires across demographic groups. Furthermore, although the data clearly shows that minorities' test scores were significantly lower, the authors find no statistically significant drop in the firm's hiring of minorities. Using a conditional logit model that controls for constant store-specific effects, the authors regress the probability of a hire being black (or, alternatively, being Hispanic) on a dummy variable that equals 1 if the hire was tested (in addition to several other controls). They find that, in all specifications, the coefficient on the job testing dummy is statistically insignificant, leading them to reject the hypothesis that testing reduces the odds of minority hiring.

In addition to the conditional logit model, the article further supports its conclusions by looking at the relationship between a store's hiring and the store's neighborhood demographics. The authors find that, for each store, there exists a close link between minority hiring patterns and minority presence in the surrounding neighborhood. Using a pooled cross section (across tested vs. non-tested applicants) with several different specifications, they conclude that the neighborhood-store relationship was not significantly changed by the introduction of testing. Importantly, these empirical results imply that before testing was introduced, employees must have lready statistically discriminated in their informal screening based on visible demographic characteristics. Nevertheless, since testing measurably increased overall productivity, the authors conclude that testing raised productivity by "improving selection *within* observable race groups."

As the authors recognize, "standard human capital variables such as age, education and earnings" are not included in their analysis. Nevertheless, they point out that applicants for line positions—the focus of their study—tend to be young, have little schooling, and are usually paid minimum wage, such that the exclusion of these variables would not significantly alter their results. Even if one accepts this explanation, an important question remains: Can the article's conclusions be extrapolated to higher-end labor markets where education, age and wages significantly vary? In any event, finding the presence of statistical discrimination in a major retail employer provides evidence that employers do in fact take race into account in making their hiring decisions. While Autor and Scarborough suggest that this may have been an efficient hiring practice, such conscious race-based decision making clearly violates federal antidiscrimination law.