Effort and Comparison Income: Experimental and Survey Evidence

Andrew E. Clark
CNRS and DELTA, Andrew.Clark@ens.fr

David Masclet
CNRS, david.masclet@univ-rennes1.fr

Marie Claire Villeval
CNRS, University of Lyon, IZA, villeval@gate.cnrs.fr
Effort and Comparison Income: Experimental and Survey Evidence

This paper considers the effect of status or relative income on work effort, combining experimental evidence from a gift-exchange game with the analysis of multi-country ISSP survey data. We find a consistent negative effect of others’ incomes on individual effort in both datasets. The individual’s rank in the income distribution is a stronger determinant of effort than is others’ average income, suggesting that comparisons are more ordinal than cardinal. In the experiment, effort is also affected by comparisons over time: those who received higher income offers or enjoyed higher income rank in the past exert lower levels of effort for a given current income and rank.

Effort, Comparison Income, Rank, Peak-End, Experiments

This article is available in Industrial & Labor Relations Review: http://digitalcommons.ilr.cornell.edu/ilrreview/vol63/iss3/3
EFFORT AND COMPARISON INCOME:
EXPERIMENTAL AND SURVEY EVIDENCE

ANDREW E. CLARK, DAVID MASCLET, and MARIE CLAIRE VILLEVAL

The authors test the hypothesis that individual effort on the job depends both on one’s own income and on the individual’s position in the relevant income distribution. Combining experimental evidence from a gift-exchange game with multi-country ISSP survey data, they analyze the extent to which relative income affects an individual’s effort, finding that an individual’s rank in the income distribution more strongly determines effort than does others’ average income, which suggests that comparisons are more ordinal than cardinal. Their experiment also reveals that comparisons over time affect effort: individuals who received higher income offers or enjoyed higher income rank in the past exerted lower levels of effort for a given current income and rank.

A growing body of literature in economics is devoted to the role of social comparisons in explaining various phenomena, including financial market behavior, criminal activity, and subjective well being. One specific area of this literature has focused on income comparisons and labor market outcomes. For example, quits are negatively correlated with a reference wage given by the average wage in the firm for similar workers; women’s labor force participation is influenced by income comparisons; and rank in the local income distribution is a good predictor of migration. These behaviors mostly concern job choice. Little is known, however, about the impact of relative income on how hard employees actually work on the job, even though efficiency wage theories are built on the concept of income comparisons, and relative concerns are appealed to as an explanation of wage compression (Frank 1984).

* Andrew Clark is a research professor in economics at the Paris School of Economics, and a research fellow of the Institute for the Study of Labor (IZA) and the Center for Corporate Performance (CCP); David Masclet is a research associate professor in economics at the National Center for Scientific Research (CNRS), affiliated with CREM (Centre de Recherche en Economie et Management) at the University of Rennes 1 and associate research fellow at CIRANO (Centre Interuniversitaire de Recherche en Analyse des Organisations); Marie Claire Villeval is a research professor in economics at the National Center for Scientific Research (CNRS), affiliated with the GATE (Groupe d’Analyse et de Théorie Economique) research institute at the University of Lyon 2, and a research fellow at IZA and CCP.

The authors are grateful to Charles Bellemare, David de la Croix, Fabrice Etilé, David Fairris, Lata Gangadharan, Alan Krueger, Sabine Krüger, George Loewenstein, Andrew Oswald and the seminar participants at the AISE seminar of the Centre d’Etudes de l’Emploi; the 2005 ESA North American Meeting in Tucson, Arizona; the EALE Conference in Prague; the Utility and Inequality Conference at PSE; and the Universities of Laval, Lyon, Paris 1, and Reading for useful comments. The ISSP data were collected by independent institutions in each country, and documented and made available by the ZENTRALARCHIV FUER EMPIRISCHE SOZIALFORSCHUNG, Köln. They authors are grateful to R. Zeiliger for programming the experiment presented in this paper and to the MiReDREES (French Ministry of Social Affairs) for a grant to support this research. Copies of the computer programs used to generate the results presented in the paper are available on request by writing to David Masclet (david.masclet@univ-rennes1.fr).

0019-7939/00/6303 $05.00
In this paper we attempt to fill this gap by analyzing the influence of income comparisons on effort using both experimental and survey data. We suggest that such income comparisons may explain why some of the empirical evidence on the wage–effort relationship is mixed. Though it is commonplace to assume that wages have incentive effects, or that higher wages make up for higher effort in a compensating differential set-up, higher wages are not always associated with greater effort in empirical work. This has variously been explained by a crowding-out effect of monetary rewards on intrinsic motivation (Frey and Oberholzer-Gee 1997), supra-optimal motivation generating choking under pressure (Ariely et al. 2009), or an earnings target that bounds effort at some threshold (Camerer et al. 1997). Here we test an alternative hypothesis—that individual effort depends on both one’s own income and the individual’s position in the relevant income distribution. In light of their comparisons with others, those who are paid relatively well work harder. We test whether others’ incomes matter, and ask, given one’s own income, which of relative income and income rank in the reference group is most important in determining effort. That is, are social comparisons cardinal or ordinal? We further ask whether income comparisons are not only horizontal (that is, comparable to other individuals at the same point in time) but also intertemporal, so that the time profile of individual income or rank helps to explain current effort at work.

Conclusive empirical proof of the existence of social comparisons is elusive since it is difficult to know to whom individuals compare themselves, and because individuals’ behavior may be correlated within a group—not because they compare to each other, but because they are exposed to common unobserved environmental factors, or because they share similar characteristics. The experimental approach adopted here has the double advantage of defining a priori the reference group and limiting any contextual effects. In addition, our approach relies on actual and costly decisions rather than on subjectively reported behavior. Unlike experimental data, survey data offer the advantage of larger sample sizes and mitigate the criticism that laboratory experiments are to an extent unrealistic. The combined use of both survey and experimental data is still relatively recent (Fehr et al. 2003; Carpenter and Seki 2010; Brown et al. 2008) and can be interpreted as a joint test of robustness. If consistent patterns emerge in both types of data, we can be more confident in the external validity of our laboratory experiments.

Our experiment extends the standard gift-exchange game between an employer and an employee by introducing income comparisons between employees from different firms. The reference group for employees consists of employees in other firms participating in the same experimental session. In this between-firm comparison design, employees are presumed to be similar and are thus expected to receive the same equilibrium wage. In the first stage of this game, the employer offers a wage contract and in the second, employees who accept the contract decide on their effort level. In one treatment we can identify income comparisons as we inform employees, before they choose their effort level, about the wages offered by a sub-set of other employers in the labor market. Testing the robustness of our experimental results required a dataset with information on discretionary effort that closely resembled the experimental design. The survey data come from the 1997 wave of the ISSP (International Social Survey Program), which includes information on both earnings and self-reported discretionary effort. The questions asked in this survey are extraordinarily close to the context of the gift-exchange game, in that the employee’s reported effort is explicitly oriented towards improving the firm’s outcome.

Literature

The existing literature on social comparisons can be broadly divided up into two strands: that on behavior and that on utility. This division can be illustrated by a direct utility function:

\[
U_i = U(a_i, a_j, \ldots) \text{ for } j \neq 1
\]

in which individual i's utility, \(U_i\), depends on both her own actions and those of relevant
others, $a_i$ and $a_j$. This utility function most often gives rise to a decision rule for $i$'s utility-maximizing behavior $a_i^*$ as:

$$a_i^* = f(a_j, \ldots).$$

The behavior and utility approaches to social interactions attempt to find empirical counterparts to (2) and (1) respectively. There are a number of drawbacks to the behavioral approach, the first of which is that data on behavior are not always particularly accurate. Second, behavior often reflects the intersection of supply and demand, whereas we are interested here in individuals’ preferences. Third, under separability conditions, others’ behavior can affect one’s own utility, but not one’s behavior. There are equally problems with the utility approach, via equation (1). In particular, we do not necessarily know how to best measure individual utility, $U_i$.

Interactions in behavior have been widely modeled econometrically, despite the identification problems emphasized by Manski (1993). Many of these studies have concluded that social interactions do indeed influence behavior, in the sense that “if you do more of something, then I am likely to do more of it as well.” One interpretation of this maxim is that this correlation reflects a concern for status or relative standing. Another is that this correlation reflects the process by which individuals may be learning from each other about how pleasant or dangerous goods or activities are (so that their behaviors are correlated), rather than caring about their status. Rival explanations emphasize the potentially key role of common omitted variables such as contextual effects; however, much of what is published in the empirical literature is very careful to try to account for these.

An alternative approach to identifying interactions appeals to proxy measures of utility, such as life satisfaction, job satisfaction, and happiness (Ferrer-i-Carbonell, 2005; Clark, Frijters, and Shields 2008). Perhaps because of a scarcity of surveys that measure both proxy utility and behavior adequately, most attention has been concentrated on the role of income comparisons in the utility function. Empirical estimation has thus mostly been based on the indirect utility function, $V_i$, testing specifications such as

$$V_i = V(y_i, y_j, \ldots)$$

in which utility depends on one’s own and relevant others’ incomes, $y_i$ and $y_j$ respectively, rather than its direct counterpart (1) above.

Both the behavior and utility approaches require that the reference group be identified—to whom does the individual compare? There are a number of potential candidates, including the individual’s peer group (those who share the same characteristics), others in the same household, spouse/partner, friends, neighbors, work colleagues, and the individual him- or herself in the past. An approach to modeling social comparisons which combines both of the above would be to consider $i$’s behavior, $a_i$ as a function of both his or her own income and that of his or her reference group, $y_i$ and $y_j$:

$$a_i = a(y_i, y_j, \ldots).$$

This can be operationalized empirically as

$$a_i = A_0 + b y_i + f y_j + g' X_i + \epsilon_i.$$

The $b$ coefficient in equation (5) shows the extent to which individual behavior, $a_i$, depends on the individual’s own income. If only one’s own income matters in explaining $i$’s behavior, then the estimate of $\beta$ will be insignificant. Alternatively, if income comparisons matter for behavior then both $b$ and $\phi$ will be significant. If action $a$ is normal then we expect $\beta > 0$ and $\phi < 0$. This is the empirical approach that we take in this paper.

The behavior we consider here is effort expended at work. We ask whether workers’ effort, $e_i$, depends on how much others earn, modeling

$$e_i = e(y_i, y', \ldots)$$

where we expect $e_1 > 0$ but $e_2 < 0$. Here $y'$ is some transformation of the income vector.

---

1Formally, $\frac{d^2 U_i}{da da} = 0$.

2Recent contributions in this vein have analyzed saving (Duflo and Saez 2002), tax evasion (Fortin, Lacroix, and Villeval 2007), labor supply (Aronsson, Blomquist, and Sacklén 1999), and students’ success at school (Arcidiacono and Nicholson 2005; Sacerdote 2001).
of other people who are in individual i’s reference group. The idea is that individual i has a comparison or reference person or group j and reduces his or her own input or effort as reference income rises, given his or her own wage.

Much of the efficiency wage literature is based on this idea of the comparison of one’s own wage to those of co-workers (Akerlof and Yellen 1990) or to that of workers in other firms (Summers 1988; Johansen and Strøm 2001). However, to the extent of our knowledge, empirical evidence that workers’ current effort does in fact depend on relative income or on one’s own past income remains slight.4

A recent literature in experimental economics has sought evidence of social comparisons among co-workers. In the first stage of the game proposed by Charness and Kuhn (2007), a principal can offer different wages to his or her two employees. These employees may possess either homogeneous or heterogeneous ability levels, but they know neither the direction nor the magnitude of these differences. In the second stage of the game, employees choose their level of effort, and according to the treatment, wages are either public or private. Income comparisons are shown to affect employees’ behavior only weakly whereas firms reduce income differentials between co-workers for fear of low-effort retaliation from the lower-paid employee. In other words, firms anticipate a negative effect from income comparisons on effort that is not actually observed in workers’ behaviors. This wage compression effect was also found by Güth et al. (2001) in a game in which information about the contracts offered to each employee was manipulated. They demonstrated that principals tend to reduce the income differential between agents when contract information is made public.

In these experiments, productivity differences are introduced between co-workers in order to motivate firms to vary their wage offers. The weak reaction by employees to subsequent income comparisons may show that workers consider productivity differences to be a fair source of income differentials. In our experiment, however, all employees exhibit the same productivity, each firm employs only one worker, and income differences result from firms’ various choices (and not from any skill differences between workers). Gächter and Thoeni (2009) provided another experimental test using the strategy method. Subjects were asked to report their effort decision in reaction to various hypothetical income distributions. They identified a large subset of individuals who reduced their effort when faced with income inequality. In our experiment, incomes are actually chosen by real firm-subjects, and we infer the influence of income comparisons from individuals’ observed effort decisions.

Empirical Strategy

Our empirical strategy is based on the joint use of experimental data produced in the laboratory and on survey data. The survey data analysis helps to check the external validity of the experimental evidence. When we evoke income comparisons here, we define income as the wage offered by the employer to the employee; that is, we do not take into account the cost of effort that will depend on the level of effort chosen by the employee.

Experimental Design

The game. We identify the impact of income comparisons on effort using a version of the standard gift-exchange game (Fehr, Kirchsteiger, and Riedl 1993). Each session involves twenty subjects who are divided into two groups, ten in the role of firms and ten in
the role of employees. Roles are attributed at random and are kept constant throughout the session. All employees have the same characteristics, in contrast to Güth et al. (2001) and Charness and Kuhn (2007). Workers do not differ in ability and thus do not have to form beliefs about the relationship between other employees’ productivity and their incomes. A Benchmark Treatment and an Information Treatment were implemented.

The Benchmark Treatment consists of the standard gift-exchange game. The use of this standard game ensures that our results can be directly compared to those from previous experiments before we introduce a new element in the Information Treatment. In every period of the game, each firm is matched randomly with an employee. There are ten periods and each period consists of two stages. In the first stage, the firm offers a contract consisting of a wage \( w \in [20, 120] \) to its employee. In the second stage, the employee decides whether to accept or reject the contract. If the contract is rejected, both the firm and the employee receive nothing. Upon acceptance, the employee must choose his or her level of effort, \( e \in [0.1, 1] \). The greater the employee’s effort, the higher the firm’s profits; at the same time, the greater the effort cost \( c(e) \) borne by the employee. This effort cost is convex, as shown in Table 1.

In the standard gift-exchange game, the employer’s payoff is

\[
\pi^v = (v - w)e
\]

where \( v \) is an exogenous redemption value; in our experiment, \( v = 120 \). This expression guarantees that the firm does not incur a loss even if the employee chooses the minimum level of effort. The employee’s payoff is

\[
\pi^A = w - c(e) - 20
\]

with a fixed labor market participation cost of 20 (corresponding to travel costs, say).

These payoff functions are common knowledge for all participants. At the end of the period, the firm is informed about the level of effort chosen by the employee, and both the firm and the employee are informed about their respective payoffs. In each new period, the pairs of firms and employees are randomly reshuffled. We implement a perfect stranger matching protocol in the sense that no agent is matched with any principal more than once, which, again, is common knowledge in that it is printed in the instructions. This allows us to rule out any reputation-building behavior (Gächter and Falk 2002).

The Information Treatment is similar in structure except that at the end of the first stage, after the firm’s income offer is revealed, the employee is told about the income offers that four other employees received from their firms in the same period. Employees can thus compare their own income to the income offered to other, ostensibly similar, employees on the labor market (but not co-workers) before rejecting or accepting the

---

\( \text{Table 1. The Cost of Effort in the Experiment} \)

<table>
<thead>
<tr>
<th>Effort ( e )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost ( c(e) )</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>15</td>
<td>18</td>
</tr>
</tbody>
</table>

---

\( ^5 \)It is important to separate these two stages in order to respect the standard gift-exchange game and to check the comparability of our results in the Benchmark Treatment to those in previous work. As others have argued in the literature on gift-exchange games, this design allows us to disentangle the effort decision from the job entry decision and to analyze whether income comparisons exert the same influence, if any, on both decisions. Combining the decision to accept an offer and the choice of effort in one single stage would simplify the design, but this requires that the decision to accept a contract offer and effort choice be determined in the same way, or that rejecting a contract is equivalent to completely shirking one’s responsibilities on the job. As the results will show, we reject this hypothesis. In any case, there is no guarantee that an agent would choose to exert zero effort were he or she not allowed to reject a contract offer.

\( ^6 \)This fixed cost, which must be covered by firm’s wage offer, exists only to avoid having a minimum wage of zero in the set of firm’s actions; it has no implications for the theoretical predictions of the game.
contract, and thus before choosing effort. We choose to display only partial information about other income offers (in each period, four other randomly-chosen income offers, instead of the whole distribution) to produce a greater variety of income distributions within the reference group. In addition, this procedure allows the relative income effect to be identified separately from any period effect. In contrast to the employee, the firm is not informed about the other firms’ income policies, reducing the likelihood that firms will behave differently in the two treatments. This decision is justified by the fact that we are mainly interested in the comparison between employees.

Equilibrium of the game. The equilibrium of this game with selfish and rational players is a minimum wage–minimum effort pair of decisions, \( w = 20, e = 0.1 \). The minimum wage contract should be accepted since the employee has no better alternative. Equally, the employee should accept the contract and choose the same (minimum) effort level in both treatments since the incomes offered by other firms do not enter into the standard individual utility function. Firms should thus offer the same (minimum) income in both treatments. It is possible, however, that in both treatments, income and effort will rise above their theoretically predicted levels. Existing research has demonstrated that employees typically reciprocate high income offers by choosing high effort levels that increase the firm’s payoff; conversely, they reciprocate low income offers by choosing low effort levels that decrease the firm’s payoff (Fehr, Gächter, and Kirchsteiger 1997).7

In addition, information about the income distribution may affect effort in the Information Treatment if individuals are sufficiently sensitive to income comparisons. If subjects make horizontal comparisons (that is, compare themselves to other employees), we may expect effort to be positively correlated with both relative income and income rank. And yet, since firms in this game are never informed about the income distribution, there is no reason why their behavior should differ across treatments.8

Procedures. The experiment was conducted in the experimental laboratory at GATE, Lyon, France, using the Regate software (Zeiliger 2000). A total of 120 undergraduate students from three local Engineering and Business schools participated in one of the six sessions organized. Two of these concerned the Benchmark Treatment and four the Information Treatment. No one participated in more than one session. Upon arrival, the subjects drew a label from an envelope, indicating the name of their computer. The instructions (see the Appendix) were distributed and read aloud. The subjects then filled out a questionnaire that allowed us to check their understanding of the rules of the game. Questions were answered in private.

The program paired firms and employees randomly and anonymously, and the subjects subsequently discovered their role (firm or employee). Since the game was repeated 10 times under a perfect stranger matching protocol, each firm made an income offer to each of the employees. This yielded a total of 200 wage offers in the Benchmark Treatment and 400 wage offers in the Information Treatment. Each employee made 10 contract-acceptance decisions and, if the contract was accepted, chose an effort level. In the next section we analyze these effort decisions.

Each session lasted one hour on average, including the payment that was carried out in a separate room. The experiment was conducted in experimental currency units with 100 points equal to 4 Euros. Each subject earned on average €14 from the experiment, including a show-up fee of €4. This indicates that in each period, the subjects earned on average €1.

Compared with survey data, this experimental approach presents many advantag-

---

7One might argue that individuals may also reciprocate higher income rank and higher relative income with higher effort in the Information Treatment. However, in our experiment, firms were never informed about the income distribution. As a consequence, ranking cannot be considered as intentional on the part of the employer. In any case, employees would presumably only reciprocate relative income if it were valuable to them, which is exactly what we want to demonstrate.

8A firm cannot know whether a lack of reciprocity is due to comparisons or to the employee’s selfishness.
es—income is perfectly measured, effort is observed directly instead of being self-reported, and the reference group is controlled. However, the artificiality of the laboratory may cast doubts on the external validity of the experimental results. For these reasons, we complement our experimental analysis with survey evidence on income and effort.

Survey Data on Work Effort

The survey data, multi-country and cross-section, come from the 1997 Work Orientations module of the International Social Survey Programme, the ISSP (the data and documentation are freely available from http://www.issp.org). The key variables in our empirical analysis are effort, earnings, and hours of work. Income is measured as individual yearly labor market earnings, converted to U.S. dollars using Purchasing Power Parities from the Organization for Economic Cooperation and Development (OECD). Hours of work are measured at the weekly level.

The variable we wish to explain is effort at work, which is almost never observed directly in survey data. To compare our experimental evidence to (larger-scale) survey data, we require a survey in which employees report their willingness to exert effort in order to improve their employer’s outcome, as in the experimental set-up where the employee provides extra effort at his or her own cost in order to increase the employer’s earnings. The question we appeal to in the ISSP data is crafted to measure discretionary effort and is thus arguably well suited to our analysis. All those employed are asked to indicate the extent to which they agree with a number of statements. One of these is, “I am willing to work harder than I have to in order to help the firm or organization I work for to succeed.”

This question is remarkably close to the context of the gift exchange game, in which any level of effort above the minimum is voluntary and increases the firm’s profit but decreases the employee’s payoff. It is reasonable to assume that “helping the firm to succeed” in the survey data is analogous to improving its payoff, and conversely that increasing the firm’s earnings in the experiment is comparable to helping it to succeed.

The weighted distribution of the five possible responses to this question in the 1997 ISSP is shown below in percentages.

<table>
<thead>
<tr>
<th>Response</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>16.7%</td>
</tr>
<tr>
<td>Agree</td>
<td>42.4%</td>
</tr>
<tr>
<td>Neither agree nor disagree</td>
<td>24.2%</td>
</tr>
<tr>
<td>Disagree</td>
<td>12.0%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

Keeping only full- or part-time employees aged 16–65 yields a sample of 12,000 observations over 17 countries (considering the two Germanies separately). Missing values on earnings, hours of work, and effort finally produce a regression sample of around 10,000 observations.

We are interested in describing differences in the response to the effort question between individuals, and we first examine the cross-country pattern in discretionary effort. To do so, we allocate a value of 5 to “strongly agree” through to a value of 1 for “strongly disagree” to the question described above. Table 2 illustrates the number of observations and mean effort, ranked by country from the lowest to the highest average effort. The countries in our study vary in the degree of social reciprocity that workers exhibited. Workers in Mediterranean countries are ranked broadly towards the bottom whereas workers in Anglo-Saxon countries are on average more willing to work hard to help their firm or organization. Portugal is an exception to this general rule, appearing toward the top of the ranking.

9This is similar to variables used in management to capture organizational commitment (see Jaworski and Kohli (1993)). It could also, of course, reflect the employee’s effort to prevent the firm going bankrupt; however, even in this context, employees provide costly effort in order to improve the firm’s outcome.

10There is a strong correlation between unemployment and the mean of this effort variable. The average OECD standardized unemployment rate in 1997 of the lowest seven countries in the ranking was 12.3%, opposed to 5.9% for the ten highest-ranked. Both the Pearson and Spearman correlations between mean effort and the unemployment rate are significant at higher than the 2% level. One interpretation is that social reciprocity allows firms and employees to attain Pareto-superior employment outcomes.
The main thrust of our paper is, conditional on country, to see how workers’ discretionary effort is related to individual demographic and job characteristics. We are especially interested in the role of income on work effort. We control not only for individual income and hours of work but also for reference group income. This latter is defined in a similar way to that in the Leyden school—by calculating average values of income over fairly broad demographic groups\textsuperscript{11}—in this case country, gender, education, and age. There are three education groups (up to 10 years of education, 11 to 13 years of education, and more than 13 years of education), and three age groups (16–29, 30–44, and 45–65). There are thus 17 (country) * 2 (sex) * 3 (education) * 3 (age) = 306 reference groups. These average income measures are called comparison income in the regression tables and correspond to $y^*$ in equation (6) above.\textsuperscript{12} Comparison income for individual $i$ in cell $j$ is calculated excluding $i$’s own income.

**Results**

Effort may be influenced by own income, by relative income or income rank, or by the income the individual received in the past if there are intertemporal comparisons. We estimate discretionary effort equations on both experimental and survey data to determine whether income comparisons affect individual effort behavior.

**Effort and Comparisons to Others**

The average wage offered by firms in the experiment was 53.51 (s.d 19.7) in the Benchmark Treatment and 53.09 (s.d. 20.0) in the Information Treatment (where wage offers were constrained to lie between 20 and 120). Both average figures are clearly above the equilibrium wage of 20 (one-tailed t-test, p<.0001) but are not significantly different from each other. Firms do anticipate reciprocity from their employees, but they do not change their wage offers when income comparisons appear.

Two different specifications of comparison income are used, the first being normalized rank, defined as rank in cell or group/number of observations in cell or group, with a correction for ties. This is a measure of how the individual’s income is ranked relative to the other members of the reference group in the Information Treatment and in the ISSP survey data (and not relative to all individuals). In the experimental data, the rank determines the position of the subject relative to the four other group members for whom wage information was revealed. This measure is bounded between just over zero for the bottom-ranked income in the cell to one for the top-ranked income. The second comparison measure is average reference group income, excluding the individual’s own income. Average and individual earnings levels are expressed in experimental

\begin{table}[h!]
\centering
\begin{tabular}{lll}
\hline
\textbf{Country} & \textbf{Employees interviewed} & \textbf{Mean Effort} \\
No. & \% & \\
\hline
USA & 775 & 6.47 & 3.93 \\
Canada & 546 & 4.55 & 3.75 \\
Portugal & 843 & 7.03 & 3.71 \\
Switzerland & 1,727 & 14.41 & 3.65 \\
Denmark & 600 & 5.01 & 3.64 \\
Great Britain & 545 & 4.55 & 3.63 \\
Japan & 607 & 5.06 & 3.62 \\
Hungary & 626 & 5.22 & 3.60 \\
Czech Republic & 526 & 4.39 & 3.60 \\
Norway & 1,366 & 11.40 & 3.59 \\
East Germany & 261 & 2.18 & 3.59 \\
West Germany & 648 & 5.41 & 3.52 \\
Sweden & 793 & 6.62 & 3.42 \\
Spain & 387 & 3.23 & 3.35 \\
Poland & 564 & 4.71 & 3.26 \\
Italy & 475 & 3.96 & 2.96 \\
France & 698 & 5.82 & 2.85 \\
\hline
Total & 11,987 & 100.00 & 3.55 \\
\hline
\end{tabular}
\caption{Mean Discretionary Effort by Country: ISSP 1997}
\end{table}

\textsuperscript{11}See for example van de Stadt, Kapteyn, and van de Geer (1985).
\textsuperscript{12}This cell-average approach does not suffer from the identification problems that occur when $y^*$ is predicted in a regression framework since the cell-average income is not a linear function of the $X$ variables (the variables that define the cells—country, age, gender and education).
currency units in the experimental data, and in thousands of U.S. dollars per month in the ISSP data.

We estimate the influence of income comparisons on effort in the experimental data via random effects tobits. The use of tobit models is justified by the number of left-censored observations in the sample. Table 3 displays the distribution of effort levels and mean income per effort level, illustrating that minimum effort (0.1) was chosen 98 times out of 180 in accepted contracts (54.4%) in the Benchmark Treatment, and 214 times out of 378 in accepted contracts (56.6%) in the Information Treatment. Not taking this data censoring into account would likely bias the coefficients.

Table 3 indicates a positive relationship between income and effort in both experimental treatments, which is typically observed in the gift-exchange game (Fehr, Gächter, and Kirchsteiger 1997) and is consistent with social motivations leading to reciprocity. Though the income–effort relationship looks somewhat steeper in the Information Treatment, the joint presence of income and comparison income makes such bivariate conclusions untrustworthy. The main effort regression results using the experimental data are shown in Table 4, and those based on the ISSP survey data are illustrated in Table 5. Table 4 consists of two panels. The left panel displays the results of six regressions in which the dependent variable is the effort choice of subjects who accepted a contract offer. The right panel, which we will discuss below, presents the results of alternative specifications that check the robustness of the results. Most regressions are estimated as tobits, which account for both left- and right-censoring (the first of which is endemic in our data). In addition, since each subject is observed a number of times (up to 10 times if the subject accepts all the contract offers), we appeal to panel data methods and estimate all of the regressions in the left-hand panel with random effects. In the Benchmark (Information) Treatment, 20 (22) contracts were rejected. Our left-hand panel sample thus consists of 180 effort decisions in the Benchmark Treatment and 378 in the Information Treatment.

Regressions (1) and (3) consider the role of one’s own income in the Benchmark and Information Treatments, respectively. Regressions (2) and (4) add normalized income rank as an explanatory variable: Higher values of this rank variable correspond to higher positions in the reference group income distribution. Since subjects are not informed about their income rank in the Benchmark Treatment, this “placebo” variable should be insignificant there, unless income and rank are strongly collinear. The last two regressions in the left-hand panel refer to the Information Treatment only. Regression (5)

<table>
<thead>
<tr>
<th>Effort level</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number obs.</td>
<td>98</td>
<td>22</td>
<td>16</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(%)</td>
<td>(54)</td>
<td>(12)</td>
<td>(9)</td>
<td>(7)</td>
<td>(6)</td>
<td>(5)</td>
<td>(2)</td>
<td>(3)</td>
<td>(1)</td>
<td>(1)</td>
</tr>
<tr>
<td>Mean income</td>
<td>50.9</td>
<td>50.1</td>
<td>61.5</td>
<td>64.1</td>
<td>69.7</td>
<td>71.1</td>
<td>71.3</td>
<td>80.0</td>
<td>95.0</td>
<td>60.0</td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number obs.</td>
<td>214</td>
<td>45</td>
<td>32</td>
<td>29</td>
<td>13</td>
<td>18</td>
<td>13</td>
<td>8</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>(%)</td>
<td>(57)</td>
<td>(12)</td>
<td>(8)</td>
<td>(8)</td>
<td>(3)</td>
<td>(5)</td>
<td>(3)</td>
<td>(2)</td>
<td>(0)</td>
<td>(2)</td>
</tr>
<tr>
<td>Mean income</td>
<td>44.4</td>
<td>59.2</td>
<td>65.4</td>
<td>64.0</td>
<td>69.6</td>
<td>75.6</td>
<td>80.8</td>
<td>79.4</td>
<td>0</td>
<td>93.3</td>
</tr>
</tbody>
</table>

If we consider individuals instead of decisions, we observe that only a minority of subjects behave selfishly. Defining as selfish individuals those subjects who choose the minimum effort in at least 8 periods out of 10, 35% of the people in the Benchmark and 27.5% in the Information Treatment are considered selfish. We cannot, however, determine whether this difference is inherent in the nature of the subjects involved in the two treatments or if it is attributable to the dissemination of income information. If it is the latter, some fraction of minimum effort decisions are motivated by social comparisons rather than by selfishness.
replaces income rank by average reference group earnings (excluding own income), and regression (6) includes both income rank and average group earnings. All of the experimental effort regressions control for both gender and number of post-baccalaureate years of education.

Table 5 reports the results of four analogous estimations on the ISSP survey data, in which the dependent variable represents the degree of willingness of workers to work harder to help the company or organization to succeed. Each individual is observed only once and we have 9,854 observations. Ordered probit regressions are estimated since the survey effort question allows five ordered responses. We follow the same logic as for the experimental data—regression (1) includes one’s own income only; regression (2) adds normalized income rank; regression (3) replaces rank by comparison income; and regression (4) estimates the joint influence of own income, rank and average reference group income. We also control for hours of work, age, gender, education and marital status, and we include country dummies. The standard errors in this Table are clustered at the reference group level.  

The results in Tables 4 and 5 show that effort is strongly correlated with one’s own absolute income at the 1% level in both treatments of the experimental data and in the survey data. Regressions (4) and (2) in Tables 4 and 5, respectively, illustrate the influence of others’ income. That is, normalized income rank attracts a positive and significant coefficient conditional on one’s own income. For the same number of dollars/experimental units earned, individuals are willing to work harder the higher their position is in the reference group income distribution. Unsurprisingly, normalized income rank is insignificant in the Benchmark Treatment (see regression (2) in Table 4), where individuals are unaware of their rank. In the experiment (column 4), a rise in rank of one position (out of five) increases effort by 0.57 ( = 0.20*2.87), which is equivalent to a wage increase of 6.52 for given rank. Compared to average income per period (53.09), this latter represents a wage rise of 12.3%. The rank/income elasticity is thus 0.614 ( = 12.29/20). In the survey data, a 20% rank increase has the same effect on effort as an extra $623 per month, which is 33% of average income, yielding a rank/income elasticity of 1.6. This higher elasticity may reflect the wider distribution of income in the survey data, the fact that rank matters more “in real life,” or that rank is more important when reputation-building is possible.

The experimental evidence thus points to income position within the reference group as being an important determinant of how much discretionary effort workers provide, over and above the actual income they receive, the latter of which has been the focus of the literature to date. This is confirmed by the survey data analysis.  

In regressions (5) and (3) in Tables 4 and 5, respectively, average income in the reference group attracts a negative coefficient, which is significant only for the experimental data. If we include both normalized rank and reference earnings in the same regression (column 6 in Table 4 and column 4 in Table 5), this marginally significant effect disappears, whereas the coefficients associated with rank remain positive and significant. Our second key result is therefore that ordinal comparisons, as measured by normalized rank in the income distribution, are more powerful predictors of employee behavior than are cardinal comparisons, that is, from others’ earnings expressed in currency units.  

Income is entered in levels. Entering all of the cardinal income variables as logs produces similar results but is not preferred by the data (the log likelihood is lower).
### Table 4. Effort, Rank and Comparison Income in the Experimental Data

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Regression models</th>
<th>Robustness tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effort in accepted contracts</td>
<td>Effort for all offers</td>
</tr>
<tr>
<td></td>
<td>Tobit*</td>
<td>Tobit</td>
</tr>
<tr>
<td><strong>Dependent Variables</strong></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Own Income</td>
<td>0.106***</td>
<td>0.085***</td>
</tr>
<tr>
<td>(0.017)</td>
<td>(0.026)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Income Rank</td>
<td>1.349</td>
<td>2.871***</td>
</tr>
<tr>
<td>(1.396)</td>
<td>(1.938)</td>
<td>(1.143)</td>
</tr>
<tr>
<td>Comparison Income</td>
<td>–0.034**</td>
<td>–0.019</td>
</tr>
<tr>
<td>(1.963)</td>
<td>(2.014)</td>
<td>(1.377)</td>
</tr>
<tr>
<td>Male*Income</td>
<td>0.0313</td>
<td>0.024</td>
</tr>
<tr>
<td>(0.024)</td>
<td>(0.035)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>(2.326)</td>
<td>(1.977)</td>
<td>(1.986)</td>
</tr>
<tr>
<td>Years of Education</td>
<td>1.284</td>
<td>1.277</td>
</tr>
<tr>
<td>(0.850)</td>
<td>(0.851)</td>
<td>(0.403)</td>
</tr>
<tr>
<td>Inverse Mills Ratio</td>
<td>0.128</td>
<td>0.177</td>
</tr>
<tr>
<td>(3.687)</td>
<td>(3.684)</td>
<td>(1.934)</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>Left-cens obs.</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>Right-cens obs.</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>Wald χ²</td>
<td>215.692</td>
<td>214.151</td>
</tr>
<tr>
<td>(108.31)</td>
<td>(110.44)</td>
<td>(226.48)</td>
</tr>
</tbody>
</table>

* RE Tobit=random effects tobit

b RE GLS=random effects generalized least squares

*** Significant at the 0.01 level; ** at the 0.05 level; * at the 0.1 level. All of the regressions in this table include period dummies; all of the regressions except (11) include session dummies.
Table 5. Effort, Rank, and Comparison Income in the Survey Data: Ordered Probits

<table>
<thead>
<tr>
<th></th>
<th>Willingness to work harder for the firm to succeed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Own Income</td>
<td>0.052***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
</tr>
<tr>
<td>Income Rank</td>
<td>0.109**</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
</tr>
<tr>
<td>Comparison Income</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours per Week</td>
<td>0.010***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Male</td>
<td>0.056**</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
</tr>
<tr>
<td>Age</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Married</td>
<td>0.068**</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
</tr>
<tr>
<td>Years of Education</td>
<td>0.009**</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td>Country dummies</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>9854</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>–13441.2</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors in parentheses.
*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

Other results in Table 4 show that in the experimental data, gender and education have a marginally significant negative effect on effort in the Benchmark Treatment but have no significant impact in the Information Treatment. In the ISSP data, controlling for rank or average income, effort is higher for men, the married and the higher-educated. The difference between the experimental and the ISSP data may reflect the far smaller variance in the demographic variables in the student subject-pool than in the ISSP data. Last, the estimates on the country dummies in the ISSP regressions (not shown) largely reproduce the effort ranking in Table 2.

Robustness Checks

To check the robustness of our experimental results, we have considered a number of alternative specifications, some of which are reported in the right-hand panel of Table 4. For the sake of simplicity, we only report the estimations that include both own wage and normalized rank. First, columns 7 and 8 reproduce columns 4 and 6 yet allow for a less restrictive form of correlation between error terms at the individual level than random effects. The estimation method here is a tobit with clustered standard errors at the individual level. Similar estimations with clustered standard errors have been carried out for each of the previous models. The primary significance of these regressions is that the results in the left-hand panel of Table 4 are unaffected by this clustering. Clustering increases the standard errors, but both own income and rank remain significant in columns 7 and 8.

The main results reported above were based only on those subjects who accepted a contract (and consequently reported an effort level). Alternatively we can include those who rejected the contract, imagining that they would have provided zero effort. In this case, no observations are excluded. We thus estimate in column 9 a random effects tobit model in which the left-censoring is set at effort level 0 rather than 0.1; column (10) shows the equivalent estimates from random effects GLS estimation. Both regressions use all 400 observations, as opposed to 378 previ-
ously. We find that, controlling for absolute income, rank continues to exert a significant effect on effort.

These regressions are based on the strong assumption that rank affects the decision to reject an offer and the choice to exert minimum (but positive) effort to the same degree. To test this hypothesis, we next estimate a random effects probit for the decision to accept an offer, with the same explanatory variables as previously discussed: The results appear in column 11. The probability of accepting an offer depends on the absolute wage offered but is not affected by income rank. A potential explanation is that contract acceptance is a blunt decision, whereas there is more latitude in effort choice. It is therefore important to respect the sequential structure in the gift-exchange game, separating offer acceptance from the choice of effort. This also explains why treating offer rejection as the choice of zero effort reduces the significance of rank (from the 1% to the 5% level).

Bearing this in mind, we proceed to an alternative two-step estimation procedure that respects the sequential nature of the game in order to correct for any selection bias from the exclusion of the observations corresponding to the rejected contracts. We first consider the random effects probit estimated in column 11 as a selection equation, producing the inverse Mills ratio (IMR). We then explain effort, conditional on contract acceptance, corrected for selection bias via the introduction of the IMR as an explanatory variable. This second equation is estimated as a random effects Generalized Least Squares (GLS) with clustered standard errors in column 12, and as random effects tobit (which we prefer, given the importance of left-censoring) in column 13. Both specifications show that rank continues to affect effort (at the 5% significance level). The results from GLS estimation suggest that a rise in income rank of one place (for example 4th to 3rd), which corresponds to a rise in the rank variable of 0.2, will increase effort by two to three ticks on the ten-point (0.1 – 1) scale, as $0.2 \times 1.235 = 0.25$.

The robustness checks therefore all deliver the same conclusion: regardless of the form of the correlation between the error terms at the individual level (random effects or clustered), and regardless of the way in which contract rejection is treated, individual effort is sensitive to income rank.\footnote{We have also estimated models using the Chamberlain procedure (results available upon request). More specifically, we add $\bar{X}$ (the average individual rank of the individual in all previous periods) to the random effects tobit model. Our results remain unchanged. Equally, GLS with fixed effects yields similar conclusions.}

**Effort and Comparison Income Across Groups**

Our main results in Tables 4 and 5 concern average effects over all individuals in the sample. However, we may suspect that certain groups react to relative income in different ways. In particular, based on recent experimental evidence on the impact of gender on competition or social preferences, we consider whether the impact of rank on effort is different for men and women in both the experimental and the survey data.

The experimental results in Table 4 include interactions between Male and both one’s own income and income rank. The estimated coefficients on these interactions are always insignificant, showing that men and women react to income similarly in determining their effort choice. Table 6 carries out the same type of analysis on the more heterogeneous ISSP survey sample, where a number of different scenarios can be tested. In addition to gender, we may consider a potential role for the environment in which wages and effort are decided, investigating interactions by union membership, sector (public vs. private), and managerial status. We have re-estimated the regressions in Table 5, allowing for interactions between the income variables and these different groups. The estimated coefficients on the income terms and the interactions are shown in turn in Table 6.

First, as in the experimental data, there are no sharp differences between men and women. Income is more important for men, but the interactions with both income rank and comparison income and men are insignificant. The second panel considers union status, and here differences do arise. Effort

---

**Effort and Comparison Income Across Groups**

Our main results in Tables 4 and 5 concern average effects over all individuals in the sample. However, we may suspect that certain groups react to relative income in different ways. In particular, based on recent experimental evidence on the impact of gender on competition or social preferences, we consider whether the impact of rank on effort is different for men and women in both the experimental and the survey data.

The experimental results in Table 4 include interactions between Male and both one’s own income and income rank. The estimated coefficients on these interactions are always insignificant, showing that men and women react to income similarly in determining their effort choice. Table 6 carries out the same type of analysis on the more heterogeneous ISSP survey sample, where a number of different scenarios can be tested. In addition to gender, we may consider a potential role for the environment in which wages and effort are decided, investigating interactions by union membership, sector (public vs. private), and managerial status. We have re-estimated the regressions in Table 5, allowing for interactions between the income variables and these different groups. The estimated coefficients on the income terms and the interactions are shown in turn in Table 6.

First, as in the experimental data, there are no sharp differences between men and women. Income is more important for men, but the interactions with both income rank and comparison income and men are insignificant. The second panel considers union status, and here differences do arise. Effort

---

\footnote{We have also estimated models using the Chamberlain procedure (results available upon request). More specifically, we add $\bar{X}$ (the average individual rank of the individual in all previous periods) to the random effects tobit model. Our results remain unchanged. Equally, GLS with fixed effects yields similar conclusions.}
Table 6. Effort, Rank and Comparison Income by Subgroups: ISSP Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Willingness to work harder for the firm to succeed (ISSP)</th>
<th>Interactions by sex</th>
<th>Interactions by union</th>
<th>Interactions by private sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Income</td>
<td>0.025</td>
<td>0.032**</td>
<td>(0.018)</td>
<td>(0.013)</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>0.039**</td>
<td>(0.022)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Own Income*Men</td>
<td>0.016</td>
<td>0.039**</td>
<td>(0.022)</td>
<td>(0.019)</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>0.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Rank</td>
<td>0.075</td>
<td>(0.082)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Rank*Men</td>
<td>0.069</td>
<td>(0.100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison Income</td>
<td>–0.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison Income*Men</td>
<td>–0.041</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own Income</td>
<td>0.035**</td>
<td>0.042***</td>
<td>(0.015)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Own Income*Union</td>
<td>–0.014</td>
<td>0.041</td>
<td>(0.022)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Income Rank</td>
<td>0.044</td>
<td>(0.066)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Rank*Union</td>
<td>0.257***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison Income</td>
<td>–0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison Income*Union</td>
<td>0.060</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own Income</td>
<td>–0.006</td>
<td>0.037*</td>
<td>(0.024)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Own Income*Private</td>
<td>0.087***</td>
<td>0.061**</td>
<td>(0.025)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Income Rank</td>
<td>0.260**</td>
<td>0.107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Rank*Private</td>
<td>–0.185</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison Income</td>
<td>–0.037</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison Income*Private</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors in parentheses. Other control variables are the same as those in Tables 4 and 5 for the experimental and survey results, respectively. The experimental results come from random effect tobits and the survey results come from ordered probits with robust standard errors.

*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

Continued

Table 6. Continued.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Willingness to work harder for the firm to succeed (ISSP)</th>
<th>Interactions by managerial status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Income</td>
<td>0.041*</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Own Income*Manager</td>
<td>–0.015</td>
<td>0.041*</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Income Rank</td>
<td>–0.044</td>
<td>(0.072)</td>
</tr>
<tr>
<td>Income Rank*Manager</td>
<td>0.285***</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Comparison Income</td>
<td>–0.003</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Comparison Income*Manager</td>
<td>–0.080**</td>
<td>(0.037)</td>
</tr>
</tbody>
</table>

The results that we have so far discussed have concerned the relationship between others’ income and the individual’s own effort. We now turn to comparisons to the income that the individual him- or herself received in the past. Broadly speaking, we theorize that past exposure to higher incomes may reduce the utility associated with current incomes private sector and income rank or comparison income. Last, the effort of workers in non-supervisory positions is only affected by their own income. Workers with managerial responsibilities, however, are sensitive to income comparisons, particularly in terms of income rank.18

Effort and Comparisons Over Time

for non-union workers is related to one’s own income only, with no role for income comparisons. Effort for union workers is very sensitive to income rank, perhaps indicating the key role of wage fairness in union negotiations. The third panel demonstrates that one’s own income is more strongly related to effort for private-sector workers, but that there are no significant interactions between

18We are aware that the use of interactions in non-linear models leads to problems with the interpretation of the coefficients (see Ai and Norton 2003). As a check, we also ran separate regressions by the different subgroups identified in Table 6. The comparison of the resulting coefficients on the income variables yielded the same qualitative conclusions. These results are available on request.
and thus decrease the current level of effort. This hypothesis has been tested with measures of satisfaction in panel data (see Clark 1999; Weinzierl 2005) but, as far as we know, not with measures of behavior such as effort. Similarly, a separate body of literature has developed on time-inseparability in behaviors such as consumption and labor supply.

One difficulty expressed in the existing research has been to ensure that \textit{ceteris paribus} holds over the long time periods between waves of survey data. Experimental data are ideally suited to testing models of habituation since we impose the same environment over time, especially in the perfect-stranger framework where there is no role for reputation building. We therefore investigate the role of previous income in determining current levels of effort by estimating random effects tobit models on just the experimental data. The dependent variable is the choice of effort conditional on contract acceptance. Our governing assumption is that higher past income will reduce current effort since past income acts as a benchmark.

We pick up the effect of past income by including running maximum income and running minimum income as additional explanatory variables. We thus ask whether effort at time $t$ depends on the highest (lowest) income the individual had been offered up to and including time $t$. We carry out an analogous analysis with respect to rank to determine whether effort is influenced more by past income or by past income rank. This running maximum/minimum specification is inspired by the peak–end transformation, which has been used to model how a flow of pain is converted into a final global evaluation (Redelmeier and Kahneman 1996).\footnote{Data from period 1 are dropped as income (income rank) and running maximum/minimum income (income rank) necessarily coincide in this period. The period dummies therefore refer to periods 3 to 10.} The period dummies in this regression pick up the fact that the running minimum (maximum) mechanically weakly decreases (increases) over time and avoids any spurious correlation between both running minimum and maximum and the dependent variable.

The usual demographic variables are also included. The results appear in Table 7.

Table 7 illustrates that the past matters. For a given income and a given income rank, effort is significantly lower the higher the most generous income offer is in the past (regression (1)); likewise, lower effort corresponds to a higher income rank achieved in the past (regression (2)). In contrast to these figures, neither running minimum income nor running minimum rank influence the current level of effort. This suggests that high past income and income rank are used as benchmarks with which to evaluate the current offer’s generosity and thus the degree of reciprocity. Regression (3) compares the influence of the two past income measures. The best past rank in the income distribution (significant at the 2\% level) matters more than best past absolute income, which is itself borderline significant (12\%). The insignificance of the interaction between gender and rank demonstrates that, as above, men and women react to rank in the same way.

**Discussion and Conclusion**

Evidence for the role of status or comparisons in determining behavior remains elusive. In this paper we have looked for effects of income comparisons on discretionary work effort in experimental data and have then compared the experimental findings to results from large-scale survey data. Below we discuss three key findings.

First, effort at work depends on the individual’s own income as well as on what others earn, both in the experimental and in the survey data. Our results thus contribute to the still small body of literature showing that comparisons among workers affect behavior in terms of actual costly decisions and not just self-reported well being. We believe ours to be one of the first papers to combine experimental and survey data to document this. Second, income rank (that is, first, second, and so on, in the relevant distribution) is a better predictor of effort decisions than is average reference group income. As such, comparisons are ordinal rather than cardinal. Last, in the experimen-
tal data, the income profile over time matters in and of itself. Those who received higher income or higher income rank in the past exert less effort in the present, at a given current income and income rank. This result is potentially important for understanding, for example, the frequent failure of mergers. Whereas existing research has concentrated on the role of income, mergers may involve substantial changes in rank as well; we have demonstrated the latter to be a strong determinant of motivation.

There are a number of explanations of the rank sensitivity of effort. We have presented our results in terms of income comparisons and concern for status. Alternatively, effort choice may derive from inequality aversion (see, for example, Fehr and Schmidt 1999); that is, those who earn a high income increase their effort so as to reduce the difference between their own earnings (income minus effort cost) and those of lower (and particularly the lowest) income workers. Though it is difficult to distinguish cleanly between theories, we note that inequality aversion would predict a stronger effort role for others’ incomes than for income rank, whereas in both experimental and survey data we find the opposite to be the case. Furthermore, inequality aversion does not clarify the role of past income and income rank in explaining current effort, whereas income comparisons do.

Another way to interpret our results is to say that workers learn what the “fair income” is in the group. In this case, their effort does not depend on within-period comparisons as such but rather on the search for the norm. Workers learn progressively how their current firm’s behavior compares to that of other firms, which would also explain why past wages negatively affect current effort, everything else being equal. As such, our regressions might capture a comparison ef-

<table>
<thead>
<tr>
<th>Dependent variable Models</th>
<th>Effort Level in Accepted Contracts RE Tobit (1)*</th>
<th>Effort Level in Accepted Contracts RE Tobit (2)</th>
<th>Effort Level in Accepted Contracts RE Tobit (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RE Tobit (1)*</td>
<td>RE Tobit (2)</td>
<td>RE Tobit (3)</td>
</tr>
<tr>
<td></td>
<td><strong>RE Tobit=random error tobit.</strong></td>
<td><strong>RE Tobit=random error tobit.</strong></td>
<td><strong>RE Tobit=random error tobit.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Other Notes: Standard errors in parentheses. The demographic and session variables are the same as those in Table 4.</strong></td>
<td><strong>Other Notes: Standard errors in parentheses. The demographic and session variables are the same as those in Table 4.</strong></td>
<td><strong>Other Notes: Standard errors in parentheses. The demographic and session variables are the same as those in Table 4.</strong></td>
</tr>
<tr>
<td></td>
<td>**Statistically significant at the .10 level; **at the .05 level; *<strong>at the .01 level.</strong></td>
<td>**Statistically significant at the .10 level; **at the .05 level; *<strong>at the .01 level.</strong></td>
<td>**Statistically significant at the .10 level; **at the .05 level; *<strong>at the .01 level.</strong></td>
</tr>
<tr>
<td>Income</td>
<td>0.166***</td>
<td>0.098***</td>
<td>0.107***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Normalized Income Rank</td>
<td>2.368***</td>
<td>3.034***</td>
<td>2.844***</td>
</tr>
<tr>
<td></td>
<td>(0.864)</td>
<td>(0.868)</td>
<td>(0.896)</td>
</tr>
<tr>
<td>Running Minimum Income</td>
<td>- 0.009</td>
<td>0.639</td>
<td>- 0.038</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.904)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Running Maximum Income</td>
<td>- 0.022*</td>
<td>0.00</td>
<td>- 0.024*</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.013)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Running Minimum Rank</td>
<td>- 4.259***</td>
<td>4.172</td>
<td>- 3.396**</td>
</tr>
<tr>
<td></td>
<td>(1.417)</td>
<td>(1.453)</td>
<td>(1.453)</td>
</tr>
<tr>
<td>Demographic variables</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Period dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Session dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant</td>
<td>- 6.421***</td>
<td>- 5.296***</td>
<td>- 5.144***</td>
</tr>
<tr>
<td></td>
<td>(1.171)</td>
<td>(1.307)</td>
<td>(1.308)</td>
</tr>
<tr>
<td>Observations</td>
<td>338</td>
<td>338</td>
<td>338</td>
</tr>
<tr>
<td>Left-Censored obs</td>
<td>197</td>
<td>197</td>
<td>197</td>
</tr>
<tr>
<td>Right-Censored obs</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-351.655</td>
<td>-349.642</td>
<td>-349.446</td>
</tr>
<tr>
<td>Wald $\chi^2$</td>
<td>332.93</td>
<td>332.84</td>
<td>349.72</td>
</tr>
<tr>
<td>$p &gt; \chi^2$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
fert based on learning. Although the subjects likely do learn the average wage over time, we do not believe that this learning entirely replaces the rank effect, for a number of reasons. First, if we were observing learning in the experiment, employees should reject more offers over time as they learn what the fair income is, and they should reject more contracts in the Information Treatment than in the Benchmark Treatment. Neither of these predictions holds. Second, if only learning is present, income rank should be insignificant, or it should at least be less important than the reference income within the group. However, reference income in the experiment is less significant than rank is, and when we include both variables in the regression at the same time, only rank remains significant. In the survey data, reference income is never significant. Last, in the experiment, the employees should also care about both their own best and worst wages in the past, which is not the case. As such, we believe that an interpretation in terms of rank and status-seeking is the most consistent with all of our experimental and survey findings.

One general implication of our work is that combining experiments in a controlled environment and survey analysis, based on subjective data, serves as a validation exercise. Although both approaches have been criticized for separate reasons, here they produce remarkably similar and consistent results about the importance of income rank on effort decisions. Another validation procedure would consist of asking the experimental subjects to perform a real effort task instead of picking numbers from a table and would constitute a natural extension of this paper.

More than 20 years ago, Bob Frank (1985) suggested that firms could exchange status for wages. In the context of between-firm comparisons, this paper has shown that these two are indeed substitutes in terms of inciting worker effort. Worker effort is lower in the face of both absolutely and relatively low incomes, where this relativity concerns both others in the same period and oneself in previous periods. This may explain why firms favor income secrecy and also why the same income at a point in time might produce different effort levels. The results also demonstrate the concrete advantage accruing to firms paying rising income profiles. More generally, income comparisons, both to others and to oneself in the past, seem to be a pervasive element of economic life.
General Information

You are going to participate in an experiment on the labor market for the MiRE – Ministry of Social Affairs. If you read these instructions carefully, you can earn a considerable amount of money. The amount of your earnings depends not only on your decisions but also on the decisions of the other participants with whom you will interact. During this session, your earnings will be calculated in points, with 100 points = 4 Euros.

At the end of the session, all the profits you have made in each period will be added up and converted into Euros. In addition, you will receive a show-up fee of 4 Euros. Your earnings will be paid to you in cash in a separate room in order to preserve confidentiality.

At the beginning of the session, each of the 20 participants will be assigned one of two roles: 10 participants will be “employees” and 10 participants will be “firms.” Your computer screen will inform you about your role. You will keep the same role throughout the session. You will never be informed of the identity of the participants you will interact with.

The labor market consists of 10 periods.

Decision-Making in each Period

Each period consists of two stages. In the first stage, each firm is paired randomly and anonymously with an employee. Each firm makes an income offer to its employee. The employee is informed of the income offer made by the firm and is also informed of the income offers made by 4 other firms randomly chosen in the room.

The employee can agree to work for the income offered by the firm or not to accept the firm’s offer. If the employee accepts the offer, he or she proceeds to the second stage.

In the second stage, the employees who have accepted an offer must decide on their quantity of work. The details of the procedure are explained below.

Please note that in each new period, firm-employee pairs are reshuffled. You are sure not to interact more than once with the same firm or with the same employee if you are a firm.

Information about the Labor Market in each Period

At the beginning of the period, the firm makes an offer to the employee. This income ranges between 20 and 120 points. Information about this income offer will be communicated to 4 other employees.

The employee is informed about both the income offer made by his or her firm and the income offers made by 4 other firms to their employees. These firms are chosen randomly.

The employee can accept the offer from the firm and work, or reject the offer and, in this case, not work; in the latter case both the employee and the firm earn nothing for the current period. Only the firm is informed about the acceptance or the rejection of the offer by the employee.

If the employee accepts the income offer, he or she receives the income and must decide on the quantity of work. The firm is informed about this quantity of work but neither other firms nor other employees are informed about it. The employee must bear a transportation cost of 20 points.

How are Payoffs in each Period Determined?

The Employee’s Payoff

If the employee has accepted the firm’s offer, the payoff is zero for the period.

If the employee has rejected the firm’s offer, the employee receives an income. He or she must subtract from this income both a transportation cost of 20 points and the cost associated with the quantity of work he or she has chosen.

The employee determines the quantity of work in choosing a number in between .1 and 1, as indicated in the table below. The smallest quantity of work is .1 and the largest is 1. The higher the number chosen, the greater the quantity of work, and the higher the firm’s payoff.

The greater the quantity of work chosen, the higher is the associated cost to the employee. The table below shows how costs vary with the quantity of work.

In the case that the income offer is accepted, the employee’s payoff in points is determined as follows:

Employee’s Payoff in Points in each Period = Income – Cost of the Quantity of Work – Transportation Cost

Transportation cost = 20 points

Relationship between the quantity of work and the associated cost

<table>
<thead>
<tr>
<th>Quantity of work</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
<th>.6</th>
<th>.7</th>
<th>.8</th>
<th>.9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associated cost</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>15</td>
<td>18</td>
</tr>
</tbody>
</table>

The Employer’s Payoff

At the beginning of each period, the firm receives 120 coupons from the experimenter that can be used to pay the income of the current period. If the firm offers an income of 120 points to its employee and if this offer is accepted, then the firm has no coupons left. If the firm offers an income of 20 points to its employee and if this employee accepts this offer, then the firm retains 100 coupons. More generally, the firm keeps

120 coupons – the income paid to the employee.

How are the remaining coupons converted into points? The number of coupons kept by the firm is multiplied by the quantity of work chosen by the employee. The result indicates the firm’s payoff in points for the current period. Then,

Firm’s payoff in points in each period = (number of coupons – income) * quantity of work

If the employee does not accept the firm’s offer,
the firm loses its coupons and its payoff is zero for the current period.

At the end of the period, the firm and its employee are informed about their respective payoffs. At the end of each period, the next starts automatically. The firms and the employees are re-matched randomly to form new pairs.

Throughout the entire session, you are not allowed to talk if not invited to do so. Any violation of this rule will result in being excluded from the session and not receiving payment. If you have any questions regarding these instructions, please raise your hand. Your questions will be answered in private.

REFERENCES


Debate among labor economists on the pros and cons of a minimum wage law has come to focus on whether labor markets are competitive or monopsonistic. Using principles and concepts of institutional economics, the author argues that this perspective on minimum wages is too narrow. In particular, he uses institutional theory to develop four theoretical rationales for minimum wage legislation: setting a floor on wages to offset imperfect competition and inequality of bargaining power; promote macroeconomic stabilization and full employment; contribute to long-term efficiency and growth; and incorporate labor market externalities and social costs of labor. One revisionist implication is that a minimum wage under plausible conditions may increase economic efficiency even in a purely competitive labor market.

The minimum wage once had strong support among labor economists, most of whom were affiliated with or sympathetic to the institutional school. Among this group were self-professed institutionalists, such as John R. Commons and his colleagues of the Wisconsin School; more mainstream but institutionally sympathetic labor economists of the 1930s and 1940s, such as Chicago labor economists Paul Douglas and Harry Millis; and post–World War II “industrial relations” labor economists, such as Dunlop, Kerr, Lester, and Reynolds (Reder 1982; Prasch 1998; Kaufman 1988, 1994). Institutionalism also provided the economic rationale for the quasi-revolution in labor policy during the New Deal period of the 1930s, including enactment of the National Labor Relations Act, Social Security Act, and Fair Labor Standards Act. Central to the important role played by Wisconsin institutionalism in these developments were the writings of Commons, whom Boulding (1957:7) described as “the intellectual…origin of the New Deal, of labor legislation, of social security, of the whole movement in the country toward a welfare state.”

Debate on the minimum wage continues unabated, but the contribution and influence of institutionalism have sharply declined, particularly in mainstream journal literature. One consequence, I argue, is that the theoretical and policy debate in labor economics over the pros and cons of minimum wage laws has become unduly narrow and, as a consequence, negatively biased against a minimum wage. The purpose of this paper is to take a first step in correcting this problem by using concepts and principles of institutional economics to further develop the positive case for minimum wage legislation. In particular, the paper outlines the core theoretical principles of institutional economics as articulated by Commons and Coase; applies these principles to the analysis of a minimum wage law; and demonstrates why many institutional economists believe a minimum wage is on balance good policy on
both efficiency and equity grounds. I note that the discussion that follows is confined almost entirely to issues of theory and is necessarily a summary of key points and ideas; also, while the discussion has a “history of thought” element, the theory and policy discussion applies directly to today’s economy.

The State of the Minimum Wage Debate

Academic debate in the United States on minimum wage legislation goes back a century (Lees Smith 1907; Webb 1912; Prasch, 1998; Neumark and Wascher 2008). At the start, the chief protagonists tended to divide into two groups. On the con side were adherents of orthodox economics (with some prominent if qualified exceptions, such as J.B. Clark, A. Pigou, and F. Taussig), meaning neoclassical Marshallian/Walrasian price theorists, as well as conservative adherents of the legal doctrines of freedom of contract and substantive due process; on the other side were institutional economists, law scholars from the Legal Realism movement, and various social reformers. The latter group, heterodox to varying degrees, formed the first law and economics movement (Hovenkamp 1990; Fried 1998; Kaufman 2009a).

During the first five decades of the twentieth century the intellectual tide waxed and waned but on balance favored the institutionalists. Their greatest moment of achievement was passage of the New Deal labor program in the mid-1930s, including the minimum wage provision in the Fair Labor Standards Act (FLSA). During the 1950s labor economics was largely under the sway of the neo-institutionalists, and the minimum wage continued to find widespread support. One well-known advocate was Richard Lester (1964).

The tide turned in the 1960s, when neoclassical economics (NE) reclaimed the intellectual high ground in labor economics (Boyer and Smith 2001). The center of the counter-revolution was the University of Chicago, where Milton Friedman and George Stigler led the charge (Reder 1982; Freedman 2008; Kaufman 2010). They and their colleagues and students were strikingly successful in resurrecting and broadening neoclassical price theory and adapting it to labor market analysis. The Chicago School also spawned what Hovenkamp (1990) called the second law and economics movement, led by Coase and Posner.1 Neoclassical economics and modern law and economics now include a wide range of models and cover many diverse market and non-market subjects. Most people (for example, Knesner and Goldsmith 1987; Manning 2003; Cahuc and Zylberberg 2004) agree, however, that the paradigm’s core remains the model of a self-regulating market, exemplified by the theory of perfect competition and the demand/supply (D/S) diagram.2 One area in which this model finds continued wide application is in the analysis of a legal minimum wage (Neumark and Wascher 2008).

It seems safe to say the more one believes in the competitive model as an approximation of the actual operation of labor markets the more one is likely to oppose employment regulation, such as a minimum wage law.

1 Part of Coase’s work, such as the Coase Theorem (named and articulated by Stigler), is congenial to neoclassical economics and modern law and economics; another part, centered on positive transaction cost, provides the foundation for “new” institutional economics (NIE). Coase distanced his work from the old institutional economics, but Medema (1996) demonstrated the numerous points of overlap.

2 Neoclassical economics is defined here as the theory of a (mostly) competitive market economy with Walrasian general equilibrium theory as its exemplar representation. Consistent with this interpretation, Boyer and Smith (2001: 212) described NE as “a sparse model of maximizing behavior in the face of competition and constraints.” In the last two-three decades economic theory has expanded beyond NE (as historically defined) to “mainstream” economics (ME). ME is broader because it moves the competitive model to deep background, or abandons it altogether, and instead makes rational choice, incentives, the law of demand, and formal model building the core components of economics. I focus on NE because the competitive model continues to form the heart of modern analysis and critique of the minimum wage (see Neumark and Wascher 2008); ME, on the other hand, is so eclectic in its assumptions and models that it falls on both sides of the minimum wage debate. Although less central to modern economics, the competitive model is hardly dead; Nobel laureate Robert Lucas (quoted in Samuelson and Barnett (2007:64)) asserts, for example, “About 99% of all successful applied economics is still based on the idea of a competitive equilibrium.”
Stigler certainly exemplifies this generalization, for in an early article he (Stigler 1946) used the competitive model to analyze and critique the minimum wage and then later offered this adverse assessment: “One evidence of professional integrity of the economist is the fact that it is not possible to enlist good economists to defend...the minimum wage laws” (Stigler 1982:60). Of course, Chicago economists and law and economics scholars do not speak for all mainstream labor economists, but certainly the Chicago School’s large influence on post–World War II labor economics, as well as the general shift of the field from an institutional to a neoclassical orientation, noticeably moved the center of intellectual gravity in the 1960s, 1970s, and 1980s toward a more critical-to-skeptical position on the merits of minimum wage legislation (for example, Rottenberg 1981).

The Chicago/neoclassical position was strongly challenged in the 1990s by research on the “new economics” of the minimum wage, led by David Card, Alan Krueger, and Lawrence Katz. They suggested that perhaps the institutionalists had more of the story right than Stigler and allies had allowed. In particular, they found in a variety of contexts and data sets that an increase in the minimum wage did not reduce employment as price theory predicts. Their core findings were presented in Card and Krueger’s (CK’s) book Myth and Measurement (1995). CK dedicated the book to Lester and reached a very Lester-esque conclusion; they stated, for example, that their empirical findings “suggest that the direct test posed by the minimum wage fails to confirm the predictions of the conventional [competitive] model” and “All this evidence suggests to us that the conventional model is incomplete” (p. 397, emphasis in original). CK went beyond Lester, however, by developing a formal model of dynamic monopsony that helped give theoretical grounding to the empirical findings.

The work of Card, Krueger, and Katz stirred up a lively and long-running controversy—largely centered on their empirical findings on the lack of a negative employment effect—the dust of which is only now settling. Three salient points about this follow-up literature deserve brief discussion here.

First, on the surface the amount of scholarly heat generated by CK’s book is surprising, given the well-worn nature of the subject and the fact that the minimum wage affects less than 5% of the American workforce. The reason, Leonard (2000:118) claimed, is that the minimum wage adherents are seen as mounting a direct and potentially highly damaging attack on the competitive core of neoclassical microeconomics, thus raising the intellectual stakes from an argument over a specific labor policy to a strategic battle over the theoretical integrity of the field’s central paradigm. On this matter he states,

The core of modern economics—neoclassical price theory—is seen to be at stake. In particular, minimum-wage research has come to be seen as a test of the applicability of neoclassical price theory to the determination of wages and employment...[It] is not just a technical quarrel over the sign and magnitude of wage-elasticity coefficients; it is the latest chapter in a longstanding methodological dispute over whether and in what domains neoclassical price theory can be said to properly apply.

Second, the popular interpretation is that this paradigm battle is at its core a dispute over whether labor markets are best modeled as competitive or some version of monopsonistic. For example, Neumark and Wascher (2007:1, 123) observed that the debate set off by CK is about “alternative models of the labor market” and described this battle of models as a contest to establish “whether the monopsony model or the competitive model better characterizes the low-wage labor market” (emphasis added).

Third, given that the contest is between the competitive and monopsony models, the key behavioral relationship focused on in empirical research is the employment effect of a minimum wage. If the employment effect is negative and significant (statistically and quantitatively), the conventional interpretation is that the evidence supports the competitive model; if it is approximately zero or even positive, then the evidence is taken to be consistent with a monopsony model. The large preponderance of studies, according to Neumark and Wascher (2007) in a detailed
survey of the post-CK empirical literature, find the minimum wage has a statistically significant negative effect on employment, leading them to conclude (p. 123), “The low-wage labor market can be reasonably approximated by the neoclassical competitive model.” This conclusion dovetails exactly with Stigler’s (1946:359) original claim that “low wage industries are competitive.”

Purpose of the Minimum Wage

Part of my argument in this paper is that the “competitive versus monopsony” way of framing the minimum wage debate is too narrow. In particular, it ignores most of the theory and policy rationale advanced by institutional economists for a minimum wage. One can legitimately disagree with most or all of the institutional side of the argument but, surely, it (like other heterodox perspectives) at least deserves an open hearing and careful examination. Thus, I endeavor to fill in this lacuna by sketching the institutional case for a legal minimum wage, drawing on several articles and books written by other institutionally oriented economists that have so far remained outside the mainstream literature (for example, Linder 1989; Craypo 1997; Prasch 1998; Power 1999; Levin-Waldman 2001, 2009). I start with the stated purpose of the minimum wage, with attention on the American case.

Nearly every mainstream article of the past several decades asserts that the first-order purpose of a minimum wage is to reduce household poverty; other purposes frequently cited are to redistribute income from rich to poor, to protect unions from low-wage competition, and to promote social justice and other normative/ideological goals. This perspective was given its imprimatur in Stigler’s (1946) article on the minimum wage, described by Leonard (2000) as the “locus classicus” of the modern literature. According to Stigler (p. 358), “The popular objective of minimum wage legislation—the elimination of extreme poverty—is not seriously debated.” In a more recent article, Sobel (1999:763) asserted that the minimum wage has two major goals: “lifting families out of poverty” and “alter[ing] the distribution of income in favor of low-income households.”

Given these imputed goals, many assess the policy negatively. Opponents, for example, cite a number of reasons why a minimum wage is a very blunt and sometimes perverse instrument to reach these goals: it reduces jobs for low-wage workers, increases unemployment, does little to reduce poverty (because the majority of minimum wage workers do not live in such households), reduces training opportunities for youth, and reduces wages for low-skilled workers in uncovered jobs (see Neumark and Wascher 2008).

All of these criticisms and negative findings are plausible and in some cases likely true, even if the magnitudes are open to considerable uncertainty and debate. They also, however, are only loosely and often indirectly related to evaluation of the real purposes of the minimum wage, as originally stated in the Fair Labor Standards Act and contained in the voluminous Congressional testimony that preceded its enactment in 1938. Illustratively, Stigler asserted that the primary goal of the FLSA is poverty reduction, but it is revealing that he gave no citation or other evidence to support this claim—perhaps because, as Blum (1947:646) observed in his comment on Stigler, “This writer is not aware of anyone who has advocated minimum wages as a means of eliminating poverty as such.”

Neither recent books (for example, Neumark and Wascher 2008) nor mainstream journal articles of the last three decades examine the stated reasons for enactment of the FLSA. The first section in the FLSA, “Congressional Finding and Declaration of Policy,” outlines the goals of the legislation. One immediately notes that this section says nothing explicitly or implicitly about reducing poverty. Reducing poverty was expected to be a benefit of the FLSA, but an indirect benefit achieved by accomplishment of other direct goals. What were these direct goals? A reasonably close synthesis of the language of the section leads to these four:

• eliminate labor standards that are so low they harm the ongoing efficiency, health, and well-being of workers.

• prevent unrestrained competition in labor markets from further lowering labor
standards in affected industries, or spreading low standards to other industries.

- prevent low labor standards from interfering with attainment of full employment and sustainable economic growth.
- eliminate low labor standards because they lead to labor disputes and divisive relations between employers and employees, thus further harming economic activity.

Confidence in this assessment is further increased by considering the conclusions of other economists who wrote on the matter at the time. First in credibility among them was Paul Douglas, faculty member at Chicago and foremost analytical labor economist of that era in the United States. He wrote two lengthy articles on wage regulation and the FLSA (Douglas 1938; Douglas and Hackman 1938) and gave an in-depth appraisal of the goals of minimum wage legislation. Douglas listed five objectives of the FLSA that exactly parallel the four stated in the act, with the modest difference that he decomposed the third objective in the FLSA into two separate parts: to augment purchasing power, and to improve productivity and growth.

It is significant that Douglas began his analysis with consideration of goal #1—establishing a minimum level of labor standards below which employers and competition cannot go—and immediately referenced Sidney and Beatrice Webb and their book *Industrial Democracy* (1897). The Webbs, proponents of a social and institutional type of economics and co-founders of the field of industrial relations, are generally credited with presenting the first and most influential theoretical rationale for a minimum wage (as well as other institutional mechanisms to establish minimum labor standards, such as unions and collective bargaining). They framed the minimum wage as a device to end “sweating” in the labor market, defined by one person as “the payment of an employer to his work people of a wage which is insufficient to purchase the necessities of life” and by another as “the unfair exploitation by unscrupulous employers of the necessities of the poor and more helpless class of workers” (quoted in Nordlund 1997:2). It may be noted that in some places the Webbs used the term “minimum wage” interchangeably with “living wage,” but in their time the two concepts were regarded as distinct though substantially overlapping. A minimum wage, designed, for example, to end sweating, was a wage to provide a bare subsistence (socially defined); a living wage was at a higher level set to provide a “decent standard of living” (Glickman 1997). The distinction is important.

### Institutional Theory: Core Concepts and Principles

The stated objectives of the FLSA reflect, in part, the economic concerns and problems of the time, most notably the collapse of labor market standards during the Great Depression. They also reflect, however, long-standing theoretical and policy principles put forward by institutional economists that were articulated as part of a larger program of early twentieth-century labor market reform and stabilization—a program, I note, developed to solve the many perceived labor problems of an era when labor markets were closer to “competitive” than at any time before or since (Fishback 1998; Kaufman 1997, 2003). The idea that the objectives of the FLSA reflect institutional theoretical principles will surely strike many readers as a stretch, given the widespread belief among modern economists that the institutional labor economists not only lacked a theoretical framework but were often hostile to theorizing *per se* (Boyer and Smith 2001). As I have tried to show in various publications (for example, Kaufman 1988, 2007a; also see Champlin and Knoedler, 2004), this belief is quite inaccurate. Commons, for example, devoted his last three books to theory and Dunlop and other neo-institutional labor economists likewise wrote numerous books and articles on the theory.

---

3 Stigler was a graduate student at Chicago in the 1930s and surely knew of Douglas’s publications on the minimum wage and the pros and cons of wage regulation, yet he omitted citation to them in his 1946 article. Although conjecture on my part, one possible reason is that Douglas was considered by his Chicago colleagues to be an institutionalist and Stigler was well known as a staunch critic of institutionalism and government interference with the price system (Reder 1982; Freedman 2008).
of wage determination and other labor topics. Given space constraints and the present article’s focus on the institutional theory of labor economics and labor markets. At least a brief summary, however, is necessary to provide the logical foundation for the institutionalist’s advocacy of a legal minimum wage. It may also be helpful as a way to broaden the modern theory and policy debate.

Before proceeding, I should note that the theory I develop comes principally from the original institutional economics (OIE) and, in particular, the writings of John Commons. But at key points I bring in complementary ideas and concepts from Coase and the new institutional economics (NIE). The synthesized institutional economics (IE) presented here also has numerous overlaps and similarities with Austrian, behavioral, post-Keynesian and radical economics (for example, Darity and Goldsmith 1995; Schmid 2004; Bowles 2004) and, indeed, much of the mainstream literature on incomplete contracts, property rights, and strategic bargaining and the burgeoning institutionalist literature in economic sociology.

**Conceptualizing the economic problem.** IE contends that traditional neoclassical economic price theory—certainly its core competitive market model that to this day is the principal tool used in analyzing the minimum wage—assumes away the most important and difficult part of the economic problem. The typical analysis of a minimum wage law explicitly or implicitly starts with “assume a competitive labor market” and then shows how demand and supply via self-interest, market competition, and the Invisible Hand leads to an efficient, welfare-maximizing configuration of prices and quantities. The competitive model, and price theory in general, lead to certain powerful insights but also sidestep the central determinant of whether a country is rich or poor and its people contented or dissatisfied. This is the challenge of devising and strengthening an efficient and just institutional order that promotes material abundance and social advance when individual agents have limited brains and a propensity to selfish, short-sighted, and emotionally driven behavior, and when conflicts, power plays, and non-cooperative behavior define the base-line state of nature. In this spirit, Coase stated that institutional economics studies “the workings of the social institutions which bind together the economic system” in a world where human beings are dominated by “self-love but not without concern for others, able to reason but not necessarily in such a way as to reach the right conclusion, seeing the outcomes of [their] actions but through a veil of self-delusion” (Coase 1994:41,116). Commons (1924:363) asserted that IE is a “return to the true spirit of Adam Smith,” which is to say an exercise in political economy and comparative institutional design with the object of studying how imperfect human beings struggle to collectively construct a prosperous, progressive social order amidst scarcity, conflict, and greed.

**Core concepts and principles.** Listed below are core concepts and principles in IE theory.

- **Ownership and property rights**: ownership and property rights are the foundation of IE theory, for they determine the institutional structure of an economy and how it performs; without prior specification of property rights and ownership, fundamental economic constructs such as commodities, production functions, demand and supply curves, and efficient allocations have no basis.

- **Institutions**: institutions are bodies of rules, both formal and informal and explicit and tacit, that are built out of property rights (broadly defined) and define the rules of the economic game and the resources, constraints, opportunity sets, incentives, and strategic interdependencies faced by economic agents. Institutions determine the structure (and existence) of labor markets, which in turn determine their behavior and performance.

- **Sovereignty**: economics is always “political economy” because the institutions and their derivative rules are in part determined through a political process in which people individually and collectively seek to capture and use the power of sovereignty to shape the institutions and rules to promote their interests and ethical viewpoints.

- **Behavioral/social model of the human agent**: people are modeled as largely pur-
poseful and self-interested, but decision-making is subject to bounded rationality and behavior is influenced by social interdependencies, emotions, and ethical precepts. Workers are an expressly human factor of production.

**Transactions and transaction cost:** a transaction is a legal transfer of ownership; transaction cost is the real resources used to effectuate and enforce this transfer.

**Modes of coordination:** economies have alternative institutional modes for coordinating transactions, including markets and organizations. Markets use prices as the coordinating device; organizations use command and administration.

**Power:** power is the ability to satisfy one’s desires and obtain a greater share of an institution’s scarce goods (material and non-material).

**Reasonable value:** economic agents individually and collectively have a notion of what is fair and reasonable; whenever an outcome or process falls outside the bounds of reasonableness, they undertake action to redress the imbalance and alter the institutional matrix of rules and rights.

**Evolution:** the interaction between outcomes and institutional structure causes economies to evolve over time in a process of cumulative causation along different path-dependent trajectories.

Both Commons and Coase are explicit that IE is not meant to completely replace NE; rather, IE is a partial complement to NE (IE “rounds-out” orthodox theory, says Commons (1934:6)) in that it brings into conventional microeconomic analysis subjects and concepts (for example, institutions) either omitted or treated as exogenous background factors. Nonetheless, at strategic places IE inevitably becomes a substitute paradigm, since taking into account NE’s omitted and taken-for-granted factors leads to negation of core NE theorems and a substantially different view of how economies work (Kaufman 2007a). Indeed, Coase (1992:713), in his Nobel address, went so far as to say that IE will “bring about a complete change...in what is called price theory or microeconomics.” In this spirit, IE argues the centerpiece propositions of neoclassical economics, embodied in the two “fundamental welfare theorems,” to be substantially flawed and inaccurate.

**Analysis of the Minimum Wage**

I now proceed to use these IE principles and concepts to analyze the consequences of a minimum wage law. Four rationales for a minimum wage are discussed, all of which lie partially or completely outside the mainstream debate (e.g., none are substantively examined in Neumark and Wascher’s recent book *Minimum Wages* (2008)). Each is illustrated with the help of Figures 1 and 2. Since the case against a minimum wage is well known, I focus here on the case for it. This tack is taken for purposes of exposition; in practice, IE economists recognize that a minimum wage has both benefits and costs and that a portion of the neoclassical critique is well founded. Hence, many IE economists favor a minimum wage but some oppose it.

**Imperfect Competition and Inequality of Bargaining Power**

The first IE rationale for a minimum wage law is that workers suffer an *inequality of bargaining power* (IBP) because imperfect labor markets and a lopsided distribution of resources and rights put employers in the dominant position in wage bargaining and the individual worker in a weaker and dependent position (Commons and Andrews 1936; Kaufman 1989). With IBP, market competition cannot fully protect the wages and conditions of labor, and thus a countervailing institutional mechanism must be introduced to ensure efficiency and equity. The primary objectives are *protection* of labor and *balance* in social outcomes.

In developing the IBP idea, I proceed in two steps. The first is to consider the role of imperfect competition in labor markets; the second is to consider the role of lopsided resources and rights. The two are distinct and need to be treated separately.

**Imperfect competition.** A situation of equal bargaining power exists in a competitive labor market since both employer and employee are wage takers, meaning neither has power to raise or lower the wage (and other conditions)
above or below the competitive market rate. This yields economic efficiency. At least by one standard the competitive outcome is also ethically just—what Budd (2004) called the standard of “marginal productivity justice” (the fact that under competition workers are paid the value of their marginal product). The labor market also provides full protection to workers since with zero cost they can quit and find jobs elsewhere and, conversely, competition forces firms to provide economically optimal terms and conditions of labor.

But what if labor markets are imperfect? In theory, an imperfect market may give the power advantage to either the employer or employee. Although an imbalance either way can occur, the IE position is that most often it is the workers—particularly those with fewer skills, less education, or from disadvantaged gender and ethnic groups—who suffer IBP. The reasoning is simple: who feels the greater pressure to reach an agreement and fill the job, and who has the greater resources and alternative options to fall back on if an agreement is not reached—the company or the worker? In most circumstances, the answer is the company. A rationale for a minimum wage (or union) is thus to “protect the underdog” and “level the playing field.”

The IE position is that in both theory and practice labor markets are always imperfectly competitive, albeit to varying degrees and in varying ways. We start with the transaction concept advanced by Commons (1934) and integrate it with the transaction cost idea of Coase (1937). The model of perfect competition entails an implicit assumption that property rights to goods and services can be exchanged at zero cost (Dow 1997). A logical implication of such a world of zero transaction cost (TC), Coase argues, is that multi-person firms should vertically disintegrate into single-person entities, such as sole proprietorships and independent contractors. The reason is that with zero TC the market is more efficient at coordinating economic activity than are organizations and management and, hence, the latter disaggregate to their irreducible minimum. Single-person firms, however, have no employees (by definition), so labor factor markets, employment relationships, and the labor demand/supply diagram also disappear by implication (Kaufman 2007b, 2008). In their place, the single-person firms (perhaps some with large capital stocks) obtain labor services through competitive product markets in the form of intermediate goods/services sold by independent contractors, such as John Jones Auto Assembler, Inc. and Nancy Smith Legal Services Ltd. Turning the logic around, if labor markets exist, then they must be imperfectly competitive, since their existence rests on a necessary condition of positive TC—itself a product of imperfect information, fundamental uncertainty, and other such market imperfections. These conditions, in turn, necessarily make labor contracts incomplete, opening the door to a host of contracting problems and market failures, such as principal-agent conflict, moral hazard, and externalities.

IE cites a second reason why labor markets are always imperfect. An essential condition of the competitive model is that labor is a homogeneous (undistinguishable) commodity. But this condition is violated by the very nature of the employment relationship. The reason is that labor services are embodied in the worker (a form of indivisibility) and cannot be separated at the time of sale; thus the worker and employer form a personal relationship at the point of production (Prasch 2004). This fact distinguishes “outsiders” from “insiders,” which, along with search and mobility costs due to imperfect information (an attribute of positive TC), makes incumbent employees preferable to external labor market job candidates as a source of labor services for firms. Hence, workers are not homogeneous but heterogeneous, leading to a situation of monopsony (broadly defined to include structural and dynamic monopsony, oligopsony, monopsonistic competition, and so on) in which the labor supply curve to the firm is upward-sloping (Card and Krueger 1995; Bhaskar and To 1999; Manning 2003; Erickson and Mitchell 2008).

The implication of the two preceding lines of argument is that as a matter of theory and logic, labor markets are always and everywhere imperfectly competitive. Thus, from a theory perspective, imperfect competition should be the base-line for analysis, particularly when
efficiency comparisons are made among alternative labor market outcomes, and the competitive model should be downgraded to a special and somewhat ad hoc case. A convenient but also incomplete representation of imperfect competition is the standard monopsony diagram, shown in panel (1) of Figure 1. In a monopsony labor market the wage is set by the firm, implying, as IE economists (for example, Dunlop 1944; Lester 1964) have long maintained, that wage rates are an administered price—a price set by employers who operate in labor markets with some degree of discretion and wage-making power. If this potential market power is exercised (for qualifications, see Bronfenbrenner 1956), the imperfectly competitive wage $W_1$ will be lower than the competitive wage $W_2$. For inframarginal workers this may take the form of salary compression. The impediments to mobility and limits on competition provide firms an opportunity to practice some degree of compression, discrimination and exploitation in terms and conditions of employment, possibly by providing a wage below the competitive level for new hires, less than competitive pay increases or promotions for tenured employees, or, alternatively, sub-competitive benefits, working conditions, or treatment. In any of these cases, the workers are at an IBP disadvantage; a minimum wage, in these conditions, helps to balance bargaining power and eliminate this less-than-competitive outcome. If well positioned (for example, set at $W_2$), a minimum wage may also lead to an increase—not decrease—in employment, such as from $L_1$ to $L_2$. IE recognizes, of course, that the extent of structural or dynamic monopsony-like power available to firms in low-wage labor markets may be modest; nonetheless, empirical evidence suggests that even here the labor supply curve to firms is often less than perfectly elastic (Manning 2003), particularly for inframarginal workers (Young and Kaufman 1997).

This analysis has important implications for the debate about the employment effect of a minimum wage. For example, if the baseline is a model of imperfect competition then the predicted effect of a minimum wage (or increase thereof) on employment is uncertain, particularly for a small-to-modest boost. In support of this proposition, Doucouliagos and Stanley (2009) find in a recent meta-analysis of dozens of minimum wage studies that after correcting for publication bias the estimated employment effect is not statistically different from zero.

Also, one must note that at a theory level the negative employment criterion is a biased test in favor of the competitive model. As Becker (1962) showed, the law of demand is a fundamental fact of scarcity and is not a unique conclusion of any one theory (for example, it emerges even if people behave irrationally). This implies that a negative employment effect is not discriminating evidence in support of the NE competitive market model. Rather, the real issue separating IE and NE is the tightness and strength of this relationship. NE price theory posits a monotonic well-defined negative relationship between the wage and quantity demanded of labor; IE, on the other hand, posits that the labor demand relationship over a moderate range is “loose” (that is, is discontinuous in places, and forms a “band”) and may have a vertical or positive-sloped section. IE recognizes, on the other hand, that over a large range of wage variation the labor demand curve is surely negatively sloped, particularly in the long-run.

The fundamental divergence between the two theories, therefore, rests on their portrayal of labor and labor markets. With respect to the latter, the competitive model assumes firms are wage takers and labor is divisible and akin to a commodity input (Addison and Hirsch 1997), thus yielding a technologically determined marginal product and a continuous negative wage/labor relationship. IE, however, assumes the opposite and gets a less negative and perhaps zero or positive wage/labor relationship. Also, elementary price theory shows that a monopsonistic firm—broadly defined to include any firm with a less than perfectly

---

4The employment increase in surviving firms may be partially or completely offset, however, by employment declines from firms that go out of business. A monopsony firm need not be profitable and, indeed, may be able to remain in business only with the “subsidy” gained from labor exploitation.
elastic labor supply curve—does not have a well defined conventional labor demand curve, just as a monopoly firm does not have a well defined product supply curve (Fleisher and Kniesner 1980; Manning 2005). The result is a more complex wage/employment relationship and potential irregularities in the law of demand.\footnote{Many economists claim NE is “value-free” and bristle at charges of “bias.” IE claims all theories are “value-laden” (including IE) because a theory depends on choice of assumptions and the factors given little versus much emphasis, both of which turn on underlying human preferences (values). Illustrative of this problem, one can look far and wide in the conventional minimum wage literature and never find mention of the fact that a conventional labor demand curve does not exist in imperfect (labor) competition, despite the centrality of the law of demand to the debate. Related examples are equating the broad field of imperfect competition with the narrow model of monopsony and neglecting other types of non-competitive theory (including transaction cost); ignoring aggregate demand effects, and omitting the labor supply effect of fixed subsistence costs. Are these randomly distributed and harmless omissions, or conscious choices that bias the conclusions of NE research?}

With regard to the labor input, IE insists that workers be modeled as psychological and sociological beings, implying bounded rationality, interdependent preferences, and volitional labor supply are essential parts of the theory of the economic agent \textit{qua} worker. Bounded rationality creates positive TC, and incomplete labor contracts and effort supply can therefore vary with wages, morale, fairness, and other factors, as represented in different strands of efficiency wage theory (Akerlof and Yellen 1986; Ippolito 2003; Bowles 2004). The result is that the workers’ marginal product is variable (perhaps greatly so) and may increase with a minimum wage over a certain range, creating the possibility of a zero or even positive wage-employment effect—just as found by Card and Krueger (1995).

NE theory can also explain CK’s results with various amendments and qualifications to the competitive model that are surely reasonable (for example, costs of adjustment), leading to some unavoidable overlap in theoretical predictions and loss of discriminating power in empirical tests. Setting up the NE competitive model as an overly narrow straw man is thus neither fair nor productive in these debates; by the same token, it seems equally reasonable to insist that meaningful restrictive boundary conditions be placed on extensions of the competitive model lest they become non-falsifiable and clever exercises \textit{in ex post} rationalization.

\textit{Unequal resources and rights.} Labor may also suffer from a bargaining power disadvantage even in a perfectly competitive labor market if the distribution of rights and resources is skewed in favor of employers. I call this a case...
of IBP before the market, in contrast to the case previously described, which might be called IBP within the market. The insight here is that labor market outcomes narrowly viewed may appear competitive but broadly viewed may considerably favor employers due to disguised forms of market power emanating from socially determined pre-market inequalities.

An IBP-before-the-market perspective examines the wage/employment outcome with regard to how alternative distributions and specifications of property rights and ownership influence the bargaining power of workers and employers. The adopted hypothesis is that workers with more resources and rights will enter wage bargaining with (ceteris paribus) a higher reservation wage, mapping into a "higher" supply curve in the market.

We again need a baseline in order to compare situations of equal and unequal bargaining power. For this purpose Commons (1934:683–84) invented the concept of reasonable value. Reasonable value is the community’s conception of the legitimate, morally justifiable upper and lower bounds to the wage bargain, given the economic fundamentals existing at the time and the range of feasible or practical alternatives facing both parties (McIntyre and Ramstad 2002). Reasonable value is inherently subjective and contingent over time and place, yet it is also determinate in that a sovereign governmental body, such as the U.S. Supreme Court (from which Commons derived this concept), has to determine the boundary lines between legitimate and illegitimate contract terms. To do this, the Court devised in the early twentieth century the doctrines of "reasonableness" and "conscionability." The idea is the Court refuses to set aside voluntarily negotiated contracts unless the negotiation process or terms are deemed unreasonable or unconscionable by prevailing community standards.

IBP before the market, therefore, corresponds to wages and other conditions outside the range of reasonable value, outcomes that would ordinarily arise only when bargaining power is sufficiently lopsided to violate community standards of legitimacy. From this perspective, the proposition that there is freedom of contract and that “all sides gain from trade” can be a cruel fiction allowing one side to impose onerous and exploitative terms on the other. As before, in theory IBP before the market can favor either employer or employee; in practice, however, IE argues that it usually favors employers and works against individual workers, particularly in early phases of economic development. With regard to resources, for example, employers have far deeper pockets than workers who live paycheck to paycheck, and can thus survive much longer if no deal is struck. Likewise, employers are less pressured to strike a deal since their revenue stream typically continues even if one job is vacant, while a worker’s revenue stream typically ceases without that job. Also important, the number of alternative job seekers from whom employers can choose typically is higher than the number of alternative job openings for an individual worker (that is, workers are usually on the “long side” of the labor market).

Rights are a second determinant of bargaining power. IE contends that in all capitalist societies legal rights start out heavily skewed in favor of employers (Commons 1924). One reason is that capital is typically scarce while labor is cheap and, therefore, societies give little regard to protecting labor; another is that employers have preponderant access to and influence in the legislative and judicial arenas (Commons 1934:673). Thus, while NE theorizes that the evolution of the common law is driven by pressures of efficiency (for example, Posner 2007), IE posits that in addition to efficiency pressures the common law also evolves from political contestation—that is, the process whereby social “outsiders” struggle to become “insiders” and, to the degree they are successful, judges re-interpret the common law to incorporate their interests.

As seen in IE, many legal rules a century ago, and some today, regarding ownership and property rights create IBP before the market (Adams 1886; Commons and Andrews 1936). One current example is employment-at-will. In any real world labor market the costs of employment-at-will fall disproportionately on workers, undercutting their bargaining (hold-out) power, reducing their minimum supply price, and shifting the
labor supply curve to the right. Also relevant is immigration law. A legal rule that allows large immigration or lax enforcement also shifts the labor supply curve to the right, substantially lowering wages and conditions for domestic workers.

The effect of these considerations is illustrated in panel (2) of Figure 1. Two alternative supply curves are depicted: the “high” supply curve $S_1$ reflects a very favorable regime of labor rights and resources, and the “low” curve $S_2$ reflects the opposite. These different regimes of resources and rights are outcomes of the political process and presumably reflect the differential power in the polity of capital and labor as organized interest groups or classes (e.g., $S_1 =$ Europe; Massachusetts, $S_2 =$ USA; South Carolina). When NE theorists analyze a minimum wage law using a D/S diagram, they insert a supply curve, such as $S_1$ or $S_2$, and find the competitive wage, such as $W_1$ or $W_2$. As long as this wage is competitive, no further questions are asked about its welfare properties. IE insists, however, that economists look deeper and, in particular, examine the relative income shares of employers and workers. If the supply curve is $S_1$, the largest income share goes to workers and the smallest to the owners of capital (labor gets the rectangle $\Delta W_1 \Delta L_1$, capital gets the triangle above it); conversely, if the supply curve is $S_2$ and the demand curve is inelastic (a reasonable assumption), the opposite is the case—capital gets much and labor gets little (compare the rectangle $\Delta W_2 \Delta L_2$ to the triangle above it). The contention of IE is that for the bottom part of the work force the regime of rights and resources most likely resembles the low supply curve $S_2$. The idea of reasonable value, in turn, is that every society accepts some (possibly large) inequality in the group incomes of capital and labor, but not beyond some limit. If supply curve $S_2$ is “very low,” then the resulting (but still competitive!) wage $W_c$ and the labor share of income $\Delta W_c \Delta L_c$ fall outside the range of reasonable value. In this case, labor is on the losing side of IBP before the market; workers also suffer from institutional exploitation—the difference between a minimally reasonable wage and set of employment conditions and the actual level resulting from the skewed rules of the game (Taylor 1977). Enactment of (or an increase in) a minimum wage is one action that can reduce or eliminate this form of social inequality and exploitation.

Before moving on, it is useful to point out that employers are not the only or perhaps even the main party using government to skew the rules of the game against the interests of workers. Consumers also have an incentive to do so, as explicated by the Webbs (1897) in their famous “chain of bargains” argument. Consumers have dual interests to the degree they not only buy goods but also sell labor. Nonetheless, IE surmises that their self-interest on balance tilts toward lower-priced goods, given that lower prices of consumer goods in the economy improve every consumer’s welfare but most forms of higher labor standards improve welfare for only a subgroup. If we look at a minimum wage, for example, most people work at companies that pay considerably above this level, so voting for political candidates who favor a minimum wage increase is likely to reduce the voters’ real income (via higher prices) without any compensating gain in wages. Examined this way, consumers and firms have a shared preference for laws, regulations, and an institutional infrastructure that promote lower labor cost (Freeman 1996). A political economy perspective suggests, therefore, that the erstwhile competitive labor market may well be in fact a site of IBP, absent progressive social norms and labor policies.

Macroeconomic Stability and Full Employment

The second IE rationale for a minimum wage is that it promotes macroeconomic stability and full employment. This argument has been entirely ignored in the modern debate on minimum wages and, until recent months, would have been quickly dismissed by most mainstream economists as irrelevant “history of thought” or anachronistic “depression economics.” However, the specter of economic depression that emerged in 2008–2009 gives the macroeconomic dimension of the minimum wage new relevance and life. IE pinpoints three positive macroeconomic roles for a minimum wage: first, to
boost employment by augmenting household income and aggregate demand; second, to prevent ruinous deflation and “destructive competition” in labor markets; and third, to maintain a better balance between spending and production both by counteracting greater inequality in income and by promoting a more broad-based sharing of the fruits of productivity growth. I discuss each in this order.

Commons (1934:804) asserted that chronic unemployment is the greatest cause of labor problems and capitalism’s greatest vulnerability; the IE case for minimum wage legislation, in turn, rests as much on the evils of excessive competition in labor markets as on the evils of restricted competition. The early institutionalists were “proto-Keynesians”; that is, they rejected Say’s Law, emphasized the link between purchasing power and employment, and rejected wage reductions as a method to eliminate unemployment (Kaufman 1997). Indeed, recent research shows that part of Keynes’ theoretical inspiration for the General Theory came from Commons and Wisconsin institutionalism (Whalen 2008; Kates 2008). Thus, progressive economists promoted minimum wage laws to offset the downward drag on labor standards caused by unemployment and to augment aggregate demand and job creation.

Neoclassical economists draw the D/S diagram and start the analysis of a minimum wage law at the equilibrium market wage without noting one hugely important assumption—namely, that the labor market is thereby presumed to be at a point of full employment (number of jobs offered equals number of job seekers), indicated by an equality between labor demand and labor supply (with, in reality, some positive frictional unemployment). This presumption rests, in turn, on the most celebrated idea of Adam Smith and the core proposition of Marshallian/Walrasian neoclassical economics: the idea that a competitive market economy is self-regulating and flexible prices rise or fall to bring the market back to a demand/supply equilibrium located on the production possibility frontier (Kniesner and Goldsmith 1987). Yes, neoclassical economics recognizes that many frictions and imperfections may impede and interfere with this process, but nonetheless its adherents hold that as a central tendency the demand/supply model and Invisible Hand theory capture the reality of a market economy (Reder 1982; Lazear 2000).

IE, in contrast, explicitly denies both the Invisible Hand theory and Say’s Law. The first part of the IE argument is that involuntary unemployment is the normal or “default” condition in the aggregate labor market. The only necessary change in assumptions from the NE model is that economic agents, to better reflect reality, are modeled as human beings. The rest flows as a matter of logic. Why is involuntary unemployment the default option in a capitalist labor market? According to competitive theory, persistent involuntary unemployment is a logical impossibility, illustrated by Reynolds’s (1991:176) assertion that “all unemployment is by choice…it all boils down to a question of price.” But involuntary unemployment is however, a partial equilibrium model of an industry or sector is used, or a model with a covered and uncovered sector. The former hides by construction the effect of a minimum wage on aggregate demand, and the latter (for example, Mincer 1976) typically demonstrates how a minimum wage may generate greater unemployment from higher labor cost but then ignores potential employment gains from greater consumer spending. The Cat is out of the Bag. There is no invisible hand. There never was. If the depression has not taught us that, we are incapable of education.” More recently, Nobel laureate Joseph Stiglitz (2002:1) reiterated this same proposition, stating, “Adam Smith’s invisible hand—the idea that free markets lead to efficiency as if guided by unseen forces—is invisible, at least in part, because it is not there.” These views are diametrically opposed to self-adjusting macro models of both Walrasian general equilibrium and Chicago-inspired “new classical” macroeconomics, as well as the NE first fundamental (“Invisible Hand”) welfare theorem.

---

6 On this matter, Keynes’ biographer Robert Skidelsky (1992: 229) writes: “Commons, an institutional economist who taught at Wisconsin University, is an important, if unacknowledged, influence on Keynes. Indeed, Keynes wrote to him in 1927 that ‘there seems to me to be no other economist with whose general way of thinking I feel myself in such general accord.’”

7 Since the U.S. federal minimum wage covers at least 90% of the work force, economists should presumably use a model of the aggregate labor market. Most often,
a logical impossibility in NE competitive theory only because workers are modeled as quasi-lifeless human commodities rather than thinking, feeling people. By substituting a social/behavioral model of the human agent for *homo economicus*, economists have demonstrated that firms may actually maximize profit by paying above-market wages as a deliberate strategy to attract and retain the best workers and motivate them to contribute maximum cooperation and work effort (Akerlof and Yellen 1986). They do so for at least two separate reasons (Slichter 1931; Bowles 2004): because of asymmetric information and price/quality interdependence (for example, as in signaling models of labor selection), and because of the positive link between high wages and work effort/cooperation (for example, as in efficiency wage models). Above-market wages, in turn, lead to a semi-chronic condition of excess labor supply in this part of the aggregate labor market—absent, of course, other forms of intervention, such as activist government fiscal and monetary stabilization programs.

In proceeding, it is next useful to introduce the institutional concept of segmented labor markets (Kerr 1977), such as in a dual labor market model (Dickens and Lang 1988). If the aggregate labor market is prone to chronic unemployment and jobs are in short supply in primary sector firms, it is likely that a portion of the job seekers will spill into the low-wage secondary sector as they become more desperate for work (Bulow and Summers 1986). Similar to a NE minimum wage model with a covered and uncovered sector (for example, Mincer 1976), the result is a rightward shift of the labor supply curve in the secondary sector and downward pressure on this already low wage. When NE economists draw a D/S diagram and start the analysis of a minimum wage law at the "competitive" wage, they are thus engaging in a partial equilibrium exercise that hides the reality that involuntary unemployment in the overall market has earlier forced this wage down to a level that would not exist in a truly competitive (full employment) world.

If panel (2) of Figure 1 is taken to represent the low-wage secondary labor market, its supply curve shifts rightward from $S_1$ to $S_2$ due to this spill-over effect from unemployment, leading to a form of "crowding" that lowers wages for people already at the bottom end of the pay scale.

But the situation in low-wage labor markets can be much worse, which brings us to the second macroeconomic role for a minimum wage: preventing destructive competition. According to NE theory, the solution to an excess supply of labor is a fall in the wage until labor demand and supply are equal. But IE argues this logic is false and, in fact, claims that the process of wage reduction (and deflation in general) is likely to make unemployment worse, not better. There are two steps in this argument.

The first is to show that wage reductions, even in a perfectly competitive economy, cannot cure unemployment. This proposition was embraced by Keynes (1936, Chap. 2) and is the revolutionary part of his message. Keynes advanced two reasons, neither of which depends on wage/price rigidities. The first is that the capitalist macroeconomy suffers a coordination failure due to a "missing institution"; that is, workers cannot "buy" a job because the property right they have to trade—a lower money wage—is not commensurable with the property right employers want for their job opportunity—a lower real wage. In practical terms, a money wage cut leads to a price cut, thus preventing a fall in the real wage. The second reason is that a reduction in money wages not only reduces the cost of labor and moves firms down their labor demand curves (toward a new full employment equilibrium) but also reduces aggregate income and expenditure and thus shifts labor demand curves leftward. The result is to perpetuate and worsen macroeconomic disequilibrium and unemployment (with debt deflation offsetting other presumed corrective forces outside labor markets, such as a fall in interest rates).

IE goes a step farther and demonstrates that wage reductions can lead to more harm by unleashing a dynamically destabilizing downward spiral in wages and prices, potentially culminating in an economic catastrophe. IE theorists call this process *destructive competition* (an oxymoron term in NE); today it is also referred to as a "race to the bottom"
The key conditions leading to destructive competition in the macroeconomic labor market are an excess supply of labor, large fixed costs for workers (ongoing costs of food, shelter, health care, and so on), limited mobility to other more buoyant labor markets (for example, immigration to Europe or Australia), and lack of a social safety net (for example, no unemployment insurance). Workers, squeezed financially by fixed survival and family costs as their spell of unemployment lengthens, bid down wages and working conditions in an increasingly desperate effort to get jobs. In NE theory, this process of wage reduction leads to a demand/supply equilibrium and full employment (Say’s Law); in IE and Keynesian theory, it shifts the aggregate labor demand curve leftward from $D_1$ to $D_2$ to $D_3$ in panel (1) of Figure 2, and wages and employment spiral downward. A legal minimum wage (along with other protective laws, such as those placing a ceiling on allowable hours of work and banning child labor) prevents destructive competition from proceeding by establishing a wage floor in the labor market, such as at $W_1$.

The third IE macro role for a minimum wage is to help ensure that the gains from productivity growth are distributed in a balanced way between labor and capital and rich and poor, thus ensuring that consumer spending keeps pace with expansion in production capacity. Absent a minimum wage law, a welfare state with a progressive income tax, or some form of collective bargaining, it is likely that a growing share of the national income will be paid to capital owners and the top tier of the professional/managerial class. The reason is that unskilled labor, having the most elastic supply curve, receives proportionately the fewest economic rents from growth, while capital and skilled labor receive larger (sometimes very large) rents as their demand curves shift rightward along inelastic supply curves (Bok 1993; Frank 1995). The modern-day exemplar of this process is the meteoric rise in CEO compensation relative to the average hourly earnings of production workers. The result from the early 1990s to 2007—parallel to the 1920s—is a growing maldistribution of income, a booming stock market, and a growing imbalance as aggregate demand growth is able to keep up with aggregate supply growth only through massive increases in debt among the middle and lower classes and luxury spending among the affluent. This imbalance is ultimately unsupportable and the economy goes into recession or depression for lack of broad-based purchasing power. Key to preventing this underconsumption scenario is an institutional mechanism to ensure that some of the income gains and economic rents from productivity growth are channeled back to...
the middle and working classes to maintain strong demand growth. A minimum wage is one such device.

Although these IE macroeconomic arguments for a minimum wage are conspicuously omitted from today’s mainstream research, they were key considerations that led not only to the passage of the FLSA in the 1930s but also to much of the rest of the New Deal labor program (Mitchell 1986; Linder 1989; Kaufman 1996). These macroeconomic rationales, along with other institutional ideas such as IBP, lost salience, however, after World War II with the revival of neoclassical economics, the spread of the modern welfare state and associated safety net programs, and the success of Keynesian countercyclical fiscal and monetary policies. Indeed, growing postwar concerns about inflation caused the minimum wage and collective bargaining to become increasingly seen as a macro “minus” on the supply side of the economy (a source of cost-push pressure) rather than a macro “plus” on the demand side. Recent events suggest, however, that the IE demand side rationale for a minimum wage retains relevance as depression fears re-emerge and policy makers desperately strive to prevent the deflation in wages and prices that NE textbooks teach is the market’s mechanism for regaining full employment.

Efficiency and Growth

The third rationale advanced by IE economists for a minimum wage law is that it promotes greater long-term economic efficiency and growth. The hallmark of the NE criticism of a minimum wage is that it distorts the price system and leads to resource misallocation and static inefficiency. This is one side of the story, and certainly one that IE admits has some truth. But IE points out a different and more positive side that is widely neglected—factors that lead to gains in both static and dynamic efficiency. The entry point is the concept of ownership and property rights, claimed by Commons (1934:5) to be the foundation of institutional economics. Secure property rights are crucial to the success of a market economy, and one of the core functions of government in the neoclassical/neoliberal paradigm is to protect and enforce these rights. Who will invest in productive enterprise, after all, if one’s property can be easily confiscated without compensation? An insight of IE is that a competitive labor market puts the worker exactly in this situation, leading to under-investment in work effort and human capital (Commons 1921).

The neoclassical conception of property in a market exchange context is a “commodity,” typically some physical or measurable good or service. Perfect competition assumes, in turn, that all aspects of a commodity’s property rights are well defined, priced, and protected. In Legal Foundations of Capitalism (1924), Commons described in considerable detail the evolution of the legal conception of property. A great transformation, he wrote, occurred in the late nineteenth century, when the U.S. Supreme Court ruled that property is not only the physical item itself (use value) but also the exchange value of the item. This distinction arose, among other places, in labor disputes where the Court granted firms an injunction against striking employees, not because they were damaging the employer’s physical property but because they damaged the market value of the property by keeping away customers and preventing the shipment of goods.

This expanded notion of property opens up a Pandora’s Box for NE theory, however. Perfectly secure property rights are a cornerstone of the competitive model, but if secure property rights are taken to mean legally guaranteed exchange values, then prices/wages can never deviate from some original equilibrium level (at least without offsetting compensation). That is, a change in D/S leads to a change in the value of property, and part of the property belonging to the parties against whom the price works is “stolen” by the market just as surely as if a thief trespassed and carried it away. Absolute security of property, therefore, can only be attained by rigidly fixed prices, no doubt much to the harm of allocative efficiency.

On the other hand, if property rights are taken to mean only secure use values, then large changes in D/S and market price can dramatically alter the exchange value of property. NE labels this a “pecuniary exter-
nality,” but concludes it does not interfere with attainment of efficiency. IE suggests, however, that pecuniary externalities, and more generally the much-touted “flexibility” of competitive markets, may actually reduce efficiency in an economy of real people. This proposition rests on the observation that competitive markets create large amounts of insecurity for economic agents and that greater insecurity, beyond some point, makes workers less productive (an inverse U relationship). Here are two reasons why.

Employment insecurity—the opposite of entitlement—up to some point is a positive, constructive force that motivates workers to perform well and do what is best for the long-run interests of the firm; beyond some point, however, greater insecurity reduces work motivation and performance by creating dysfunctional levels of stress, impaired decision-making and attention (bounded rationality), constant job search, and unduly short time horizons (Polanyi 1944; Kaufman 1999b).

Second, it is widely recognized that workers develop a stronger perceived property right in their jobs as they accumulate additional years of tenure with the firm (“sweat equity”), even though, of course, the courts and legislatures have generally refused to legally recognize such a right (Fogel 1982). When firms cut wages or benefits (or both) or order layoffs, as would routinely happen in a competitive market, workers instinctively regard these actions as tantamount to theft of part of their property right in the job, leading them to react in a variety of ways that erode efficiency (Perlman 1928; Polanyi 1944). Examples include reduced work effort, greater absenteeism, and calling in the union organizer. Thus, measures that reduce excessive wage/employment insecurity—without going to the other extreme of a completely rigid wage/employment system—promote efficiency. This idea is one of the foundation stones for the modern welfare state, and labor law and social insurance plans are vehicles for implementing it (Moss 1996).

Much the same idea applies to investment in human capital. An NE indictment of any form of labor market regulation that puts a floor on wages is that it reduces firms’ willingness to provide general on-the-job training (OJT), since workers can no longer offer to work for a lower wage to compensate firms for the cost of the training. As in all these matters, however, there is another side to the story. Rarely asked is this question: will firms invest in specific OJT, and will such training be attractive to workers, if the value of this asset is at great risk from layoffs due to large shifts in D/S in laissez-faire labor markets? Probably not. According to Galbraith (1967), firms are unwilling to invest the billions needed to finance new products and plants without stable market conditions; similarly, in the case of labor markets, without stabilization firms cannot risk the fixed costs of internal labor markets (ILMs), implying that competitive labor markets are destructive to the extent they undermine ILMs and the many efficiency advantages they bring (Doeringer and Piore 1971; Williamson 1985). The heretical conclusion, therefore, is that the optimal (most efficient) level of labor market competition is not the maximum amount associated with perfect competition and as envisioned in NE’s first fundamental welfare theorem but, rather, some intermediate and balanced level, as with some degree of imperfect competition and market regulation. A minimum wage is one essential element in achieving this needed market stabilization.

Labor market regulation and mandates may promote greater efficiency and growth through several other channels. All involve an aspect of human agency and incomplete contracts. One example is the IE “shock effect” argument. Institutionalist Sumner Slichter (1931) argued that collective bargaining and minimum wages might not produce a negative employment effect because the increase in labor cost shocks management into tightening up on other elements of cost in order to maintain profitability. NE economists (for example, Stigler 1946) typically dismiss this argument as ad hoc or lacking empirical evidence. IE, however, provides a logical account for the shock effect. Assuming economic

---

9To better appreciate this point, the reader may wish to consider the research productivity of professors starting from 0 job security (a daily spot market) to 100% lifetime tenure protection.
agents approximate human beings, managers and workers have bounded rationality, volitional effort supply, and face positive TC. The first casualty is the NE hypothesis of cost minimization and profit maximization. Positive TC and incomplete contracts open the door for the principal-agent problem in firms. The interest of the owners is maximum profit, but the interest of the salaried managers is their own utility maximization—a function of many things not only in addition to profit (required to keep their jobs) but at the expense of profit (for example, big expense accounts, shorter work hours). Institutionalists such as Berle and Means (1932) have labeled this the “separation of ownership and control,” and they note that it can lead to satisficing behavior with regard to cost and profits (Simon 1982; Kaufman 1999a; Altman 2001). Satisficing behavior, in conjunction with a minimum required profit level, leads management to tolerate organizational slack and above-minimum costs. A rise in labor cost from a labor law or mandate does not, therefore, necessarily translate into a decrease in employment or efficiency, since the managers may well be able to find equivalent cost savings in other areas, including their own effort supply and (in some cases) multi-million-dollar compensation.

Minimum wage laws may enhance efficiency in another way as well, by protecting not only workers but also “high road” employers who make long-term investments in human capital, physical capital, and R&D. Research shows that productivity is higher at firms using a high performance work system (HPWS) with self-managed work teams, job security provisions, extensive training, employee involvement methods, and formal dispute resolution programs (Appelbaum, Berg, Kalleberg, Bailey 2000). These kinds of organizational investments are crucial for long-run growth but may be seriously impeded by the instability and hyper short-term competition found in competitive markets. A minimum wage law can protect and encourage new forms of work organization, such as HPWS, by putting a floor under competition so “low road” firms are not able to undercut and drive out high road firms.

Last but not least, we come to the link between efficiency and fairness. This link was already touched on with regard to efficiency wages, but it has a much broader and more compelling role to play. An implication of the second NE fundamental welfare theorem is that the attainment of Pareto efficiency in a competitive economy is independent of fairness in endowments and outcomes (Stiglitz 2000). IE denies this proposition at both a micro and macro level. The micro-level insight is that whereas commodities do not care if they receive a high or low price and have no conception of fair treatment, people do. Research in behavioral and experimental economics systematically shows that when procedural and distributive norms of fairness in the workplace are violated, workers retaliate by reducing work effort, cooperation, and organizational citizenship behavior, thus exacting a reciprocal “price” in the form of reduced profit and efficiency (Falk, Fehr, and Fischbacher 2003; Schmid 2004). Fairness also promotes efficiency and growth at the macro level (Kitson, Martin, and Wilkinson 2000). Societies that have a more balanced income distribution (at least up to a point) show higher growth rates (Gobbin, Rayd, and Van de Gaer 2007). One reason for this is that when people at all levels of a society feel they are sharing equitably in the fruits of productive enterprise, they also feel more committed to and respectful of the enterprise. This sense of shared gain and social solidarity helps maintain and expand both a firm’s and a nation’s single most productive asset—a cohesive, cooperative, and lawful institutional order. Without such an institutional order, and the sense of inclusion and fair treatment it rests on, organizations and societies fall into the Pareto suboptimal trap of the Prisoner’s Dilemma game (Miller 1991). Labor market regulations, such as a minimum wage, are one solution to maintaining cooperation and social justice—and avoiding industrial conflict and adversarial employment relations.

Externalities and Social Costs of Labor

The fourth rationale advanced by IE economists for a minimum wage law concerns labor market externalities and social costs of labor (Blum
NE typically brings externalities into the analysis as “exceptions” to the operation of competitive markets, but IE maintains that externalities are logically present in all labor market situations because, as earlier noted, the employment relationship always and everywhere involves an incomplete contract. An externality arises any time one or more dimensions of a good or service are not fully priced and covered in a complete contract, thus causing part of the benefits or costs to be omitted and shifted onto third parties. When this happens, there develops a divergence between the private benefit/cost realized by the buyer and seller and the social (or total) benefit/cost realized by the buyer, seller, and all affected third parties. Since buyers and sellers make decisions based on private benefits/costs, this divergence leads to incorrect decisions, false price signals in the market, and economic inefficiency from misallocated resources and inequity from misplaced or unanticipated gains/losses in exchange.

The externality and social cost ideas have several applications to the minimum wage debate. A serious gap in NE labor supply theory (see Killingsworth 1983), for example, is the focus on marginal cost (the trade-off between income and leisure) and neglect of most elements of fixed living cost. In contrast, in nineteenth-century classical economics, labor’s subsistence wage (where “subsistence” is defined relative to prevailing socioeconomic conditions) was a central focus of attention. Adam Smith (1937:67) noted, for example, that “a man must always live by his work, and his wages must be at least sufficient to maintain him.” This insight was later developed by Sidney and Beatrice Webb (1897) and John M. Clark (1923) into a social rationale for a “national minimum” in terms of wages, earnings, and benefits. This national minimum is sometimes called the “social wage.”

Sidney Webb explains the idea behind the social cost rationale for a minimum wage in his article in the *Journal of Political Economy* (1912:986–87):

The continued efficiency of a nation’s industry obviously depends on the continuance of its citizens in health and strength. For an industry to be self-supporting, it must, therefore, maintain its full establishment of workers unimpaired in numbers and vigor, [and] with a sufficient number of children to fill all vacancies caused by death or superannuation. If the employers in a particular trade are able to take such advantage of the necessities of their workpeople as to hire them for wages actually insufficient to provide enough food, clothing, and shelter to maintain them permanently in average health; if they are able to work them for hours so long as to deprive them of adequate rest and recreation; or if they can subject them to conditions so dangerous or insanitary as positively to shorten their lives, that trade is clearly obtaining a supply of labor force which it does not pay for…he [the employer] is clearly receiving a subsidy or bounty…[and is] economically parasitic.

The idea is that the wage paid workers must cover not only the opportunity cost of leisure but also the maintenance and depreciation of their human capital, or otherwise private production cost understates social production cost. This means that the wage must cover all items that define the long-run subsistence cost of labor, such as minimal necessary health expenditure, minimal retirement income, minimal income support during periods of unemployment, and minimal income for dependent children (so the nation has a future work force). Firms, however, may be able because of market imperfections and incomplete contracts to partially or completely avoid paying these costs, which in effect also shields consumers from these costs in the form of higher product prices. Instead, the costs are shifted to the workers themselves, their families, local communities, or the nation at large. For example, a firm may be able to obtain employees at a low wage and not pay health insurance; or it may opportunistically renege on pension payments by firing workers when they get closer to retirement age; or it may routinely continue to pay the maintenance cost of capital during slow periods but shift the maintenance cost of labor to third parties.

---

10 The social cost rationale for a minimum wage is developed in considerably greater detail in Kaufman (2009b). Also see Power (1999).
through layoffs. Consumers and firms are, in Webb’s term, “parasitic” in that they enjoy lower prices and more material abundance at workers’ expense—particularly low-wage workers who often work in unsafe jobs, have the least financial ability to withstand ill health, and have the least income for and access to alternative suppliers of health care.

A shifting of social labor cost is also facilitated by a second factor—a large supply of labor in the market. Even in a perfectly competitive labor market, the wage may not cover the subsistence cost of labor. A huge supply of firms, for example, floods the market with product, causing the price to fall until enough firms “die,” which constrains supply and restores profit to a normal (“subsistence”) level. The same process works in labor markets. Assume that the labor market is opened to unrestricted legal and illegal immigration. To balance supply and demand, the wage may have to fall so far that it does not cover the minimum subsistence costs of labor. The parallel market solution is for some workers to “go out of business” so that the labor supply shrinks until wages again cover minimum fixed and variable labor cost. This process may happen in a variety of ways: for example, workers may withdraw from the labor market and maintain themselves through crime or in the underground economy; they may become homeless and beg for food and live in community shelters; or they may die through sickness and starvation. If, however, workers had a recognized legal and human right to the social wage, the social cost problem would disappear. Commons (1898) proposed this solution (a right to work, or the minimum income therefrom) but was denounced as a radical (Kaufman 2003); it is, however, a logical step in making the competitive labor market even more “perfect” by filling in a missing human property right. Absent this step, a better-than-nothing solution is a minimum wage.

These arguments are demonstrated in panel (2) of Figure 2. When NE economists analyze the economic effects of a minimum wage law, they draw a diagram with demand and supply lines $D$ and $S$ and a “competitive” wage labeled $W$. But in doing so they elide several considerations about the location of the supply curve, including not only all the factors that create IBP and cause the supply curve to lie further rightward but also the social wage that covers the full cost of the labor input. For purposes of illustration, assume the social wage is $W$. In fact, the social wage may be above, equal to, or below the prevailing market wage, determination of which is ultimately an empirical issue.\footnote{The social wage differs across demographic groups; for example, it is lower for teenagers and higher for married adults with children. A uniform national minimum wage is therefore a relatively blunt instrument for solving the social cost problem, and is becoming more so as the work force becomes more diverse. To solve this “one size fits all” problem, the early institutional economists advocated creation of tripartite wage boards that could tailor different minimum wage rates by state and industry (and also introduce greater economic democracy). This idea was incorporated in the FLSA but eliminated shortly after World War II in favor of a uniform national standard.}

We may call the market wage $W$ a \textit{pseudo competitive wage}—pseudo because, first, it is an illusion obtained only by omitting consideration of the full social cost of labor and the extra-market sources of labor’s IBP; second, because it purports to yield an efficient resource allocation when in fact it yields an inefficient allocation (as with any externality). A market-determined wage at or above $W$ may be called a \textit{true} (or “full”) \textit{competitive wage} since it covers all labor costs and yields economic efficiency. In Figure 2, the difference $W - W$ represents the per-unit \textit{social tax} on labor and “social subsidy” to capital and consumers.\footnote{To the degree that long-run competition erodes profits to a normal break-even level, the entire social subsidy eventually passes to consumers. But not all consumers benefit equally, since consumer spending varies with income. Krugman noted in 2002 that the 13,000 richest American families had more income than the poorest 20 million (Krugman 2002)—thus, the rich and affluent most likely disproportionately gain when there is no minimum wage to end the social cost subsidy (embedded in lower prices), while the working poor bear the largest share of the “tax.”}

Given this market failure, Coase (1960:18) argued, “The problem is one of choosing the appropriate social arrangements for dealing with harmful effects.” Clearly, a legal minimum wage is one such social arrangement.
The idea is to set the minimum wage at the level of the social wage $W_2$, thus creating a wage floor that covers labor’s minimum fixed and variable cost (pro-rated on an hourly basis). This wage floor is depicted by the solid horizontal line at $W_2$. A yet higher income floor, such as is envisioned with a “living wage,” would lie above $W_2$.

Now consider the effect of the minimum wage on employment and unemployment. It is possible, as Card and Krueger’s (1995) study found and Doucouliagos and Stanley (2009) confirmed in their meta-analysis, that a moderate minimum wage hike has on average a close-to-zero employment effect. Assume for purposes of debate, however, that the neoclassical theory is correct and the imposition of a minimum wage causes an employment decline, say from $L_1$ to $L_2$. This loss of jobs is at the heart of the neoclassical critique of a minimum wage, but by the welfare standards of orthodox economics it should be welcomed rather than deplored.

The reason is that the minimum wage reduces or eliminates the externality-like gap between the private and social cost of labor and thus improves economic efficiency. The effect is analogous to placing a tax on a paper mill that dumps pollutants into a river. The higher cost causes the firm to reduce production and cut employment, but economic welfare is improved—not hurt—because the tax corrects a market failure (a missing property right) that allows the firm to use a valuable social resource (the river) without paying the cost. A minimum wage is also, in effect, a tax on firms, but these firms—like the paper mill—are using a resource to make profit without paying the full social cost. The minimum wage, therefore, has exactly the desired effect: it ends (or reduces) the subsidy on low-wage labor and causes firms to cut back on production and employment to the efficient level that would prevail if the labor market were truly at a competitive equilibrium (for example, $W_2$, $L_2$). Society gains from this loss of jobs because the human capital can be transferred to alternative uses that yield a higher return.

Other indirect benefits also arise. The loss of jobs from a minimum wage (if such occurs) forces society to confront and solve a problem it otherwise prefers to ignore: that is, why do $L_1 - L_2$ workers have such low productivity that they cannot earn at least a subsistence wage? Moreover, not only do some workers potentially lose their jobs, some firms also go out of business. But again this result has to be regarded as in the social interest, since it weeds out the least efficient and most backward firms and concentrates capital and managerial talent in the most efficient and advanced firms. In this regard, the Webbs observed, “The Common Rule has another, and even more important result on the efficiency of industry, in that it is always tending to drive business into those establishments which are most favorably situated, best equipped, and managed with the greatest ability, and to eliminate the incompetent and old-fashioned employer” (1897:727–28). The Webbs also noted that the wage floor usefully serves to shift firms’ search for competitive advantage from additional cheapening of already low-priced labor to other methods, such as technological advance, higher product quality, capital investment, and improved business methods, that collectively promote higher dynamic efficiency.

Critics may nonetheless assert that it is socially misguided to destroy these $L_1 - L_2$ jobs when they are held by low-wage workers who presumably need them and, further, voluntarily accepted them. Dwelling on this last point another moment, critics will often ask: how can the government in good conscience prevent workers and employers from freely negotiating employment contracts when it is self-evident that both sides gain from the exchange? The answer is given by taking the question in reverse direction. That is, if providing jobs to the poor is the overriding consideration, then why stop at abolishing the minimum wage? Why not also abolish numerous other protective labor laws, such as occupational safety and health or even the ban on child labor? Doing so would get rid of additional “burdensome,” “inflexible,” and “inflationary” labor regulations, thus reducing the price of labor and inducing firms to hire more people. All sides gain from trade, so is not welfare increased? The answer is no: on efficiency grounds these outcomes are harmful to both workers and society if
the market wage is less than the social wage, whereas on normative grounds they represent a retrogression to inhumane labor conditions that advanced societies long ago repudiated.

The unemployment effect of a minimum wage also deserves attention. Critics note that a minimum wage not only reduces employment but most likely increases unemployment in the market. Is this not also a harmful and perverse outcome? Perhaps, but it may also have a beneficial and constructive side. If the minimum wage is $W_2$, unemployment is $L_3 - L_2$. The portion represented by layoffs and reduced hiring is $L_1 - L_2$. For reasons just cited, the “new unemployed” in this group are being inefficiently utilized, and the disappearance of their jobs allows them to shift to more productive employment (or requires society to adopt other policy measures to create such work). The $L_3 - L_1$ portion represents new labor force entrants, induced to search for work by the higher wage. This may be judged a social virtue on two counts. Some of these new entrants are presumably substituting toward market work and away from underground or black market work; also, encouraging people to seek gainful employment is often espoused as a desirable social value, which is exactly what a minimum wage does.

One observation and two caveats are required. First the observation. The social wage, expressed on an annual basis, is approximated by the federal government’s “poverty line.” As noted earlier, however, the direct purpose of the minimum wage is not poverty reduction; rather, the objective is (in part) to cover the full social cost of labor. Thus, if the market wage is less than the social wage, a valid case exists on economic efficiency grounds to enact a minimum wage even if the labor demand curve is elastic and the higher wage reduces total labor earnings (the wage bill), or many minimum wage workers are not in poverty households, or both. But these problems—admittedly a large concern for policy—are muted in significance since empirical studies find that in most cases labor demand curves are inelastic (Hamermesh 1993) and, further, that on balance a higher minimum wage may well reduce—albeit perhaps quite modestly, particularly given the increasingly diverse nature of the work force—the number of poverty households (Card and Krueger 1995; see Neumark and Wascher 2002 for conflicting evidence). Thus, I conclude that a minimum wage law is potentially a double win—it can not only directly contribute to increased efficiency but also indirectly contribute to poverty reduction. It then holds the potential for a triple win if we also count a more balanced and just society.

Now the caveats. A minimum wage may solve the social cost problem, but as Coase (1960) observed, there are also numerous alternative ways to accomplish the same end, one or more of which may be superior. For example, one approach is to eliminate the gap between the market wage and social wage through government programs that cover the overhead costs of labor, such as universal health insurance and old-age pensions. Alternatively, the same could conceivably be accomplished through universal collective bargaining. In effect, government or union provision lowers the wage from work that is necessary to cover labor’s social overhead costs, which in Figure 2 is equivalent to lowering the social wage from $W_2$ to $W_1$ (thus eliminating the social cost gap). A second approach is to shift the labor demand curve to the right through some type of wage or job subsidy to employers until it intersects the wage floor on the supply curve $S_1$. The virtue of this approach is that it closes the social cost gap and increases employment (Macpherson 2004). Yet a third approach is an income supplement for low-wage workers, such as the Earned Income Tax Credit (EITC) in the United States (Neumark and Wascher 2008).

A complete analysis of the minimum wage, therefore, requires a comparative institutional analysis of the pros and cons of alternative policy instruments. I do not undertake that project here. It is worth noting, however, that a minimum wage is likely to become less useful and attractive as the degree of union organization increases and as a country’s social welfare program expands in breadth and depth (as in many European countries), in part because these other instruments fill the social cost gap. Going further, even in lightly regulated neo-liberal labor
markets a legislated minimum wage could be a second-best solution relative to one or more of the other approaches cited above. Institutionalisists, therefore, take a pragmatic position on the minimum wage and are certainly willing to consider other options, consistent with Commons’s observation that “the problem is one, not of ideals, but alternatives” (1919:185). What they do not countenance, on the other hand, is doing nothing when a first-best solution is unattainable but an admittedly imperfect but helpful second-best option is available.

Conclusion

Anyone who has completed elementary economics and can draw a demand/supply diagram knows the essence of the neoclassical critique of a legal minimum wage. The institutional position is that this diagram, and the critique that flows from it, provide useful insight on aspects of the minimum wage issue and labor markets in general. For example, institutional economists would not claim that the minimum wage can be substantially raised all at once without engendering a significant negative employment effect, nor would they dismiss the many virtues of free, flexible labor markets. They do not view the competitive model of labor markets as completely wrong or useless; rather, their position is that it provides certain useful insights and ways of thinking about markets but is also prone to yield systematically misleading answers and conclusions because it unduly neglects human and institutional considerations (Dunlop, 1984). In this spirit, Commons (1919:17) remarked, “The commodity theory of labor… is not false, it is incomplete”; in the same vein, Coase (1992:714) called NE theory “a great intellectual achievement” but also stated, “What we have is a very incomplete theory…. [it] lives in the minds of economists but not on earth.” The incomplete nature of conventional theory was, as earlier noted, also the critique made by Card and Krueger (1995).

Following in this line of thought, my contention is that reliance on the NE paradigm tends to yield a narrow and one-sided portrait of the minimum wage. It puts most of the emphasis on negative features of a minimum wage, places excessive weight on the employment effect at the expense of other economic and social considerations, and downgrades or omits numerous positive features. As evidence, compare the truncated list of positive rationales for a minimum wage given in Klein and Domple (2007) (reproducing a survey sent to hundreds of economists about the pros and cons of a minimum wage) with the broader list given here. Also instructive are these anomalies. According to a 2006 national poll 83 percent of the American public said they favored a $2.00 increase in the minimum wage, even though economists have been criticizing the minimum wage for decades.13 Similarly, despite the highly negative picture of the minimum wage painted in neoclassical economics, Britain recently adopted a national minimum wage with little subsequent measurable negative effect (Arrowsmith, Gilman, Edwards and Ram 2003; Metcalf 2004). Lastly, since publication of CK’s Myth and Measurement (1995), the economists who have most actively and influentially led the countercharge against the minimum wage are Neumark and Wascher. In their book Minimum Wages (2008: 289) they conclude, “we find it very difficult to see a good economic rationale for continuing to seek a higher minimum wage.” The evidence they advance to support this conclusion is detailed and impressive; it is also adduced, however, with little-to-no substantive attention to the four positive rationales presented here or in earlier IE-related studies. Possibly, therefore, a more balanced and inclusive theoretical analysis would have led to a more balanced and inclusive empirical investigation and policy conclusion.

In fairness, one must also admit that institutional economists have not done well in putting forward their case for the minimum wage, thus allowing neoclassical economists to more easily and legitimately pass it by. To remedy this situation, I have endeavored in this paper to re-broaden the theory and policy debate on the minimum wage by outlining the core arguments advanced

for it by earlier generations of institutional economists. Emphasized here are efficiency arguments, although normative arguments are also important. The institutional case for a minimum wage law is partially distinct but substantially complementary to the theory presented by other supporters, such as Card and Krueger. In addition, the analysis presented in this paper also helps integrate OIE and NIE.

I do not deny that some or many of the IE arguments cited in favor of a minimum wage can be captured in some expanded/amended NE-based model; I do claim, however, that if carried too far, this risks making NE a “theory for all seasons” and therefore fundamentally ad hoc and non-falsifiable. More important than critiquing NE is demonstrating that IE has an analytical core and provides labor economists with an alternative paradigm and perspective on labor issues. In addition, the NE/IE dichotomy usefully puts into the open fundamental issues that are often sidestepped, assumed away, or ignored; for example, that labor markets either are or are not self-correcting, the Invisible Hand in labor markets either achieves efficient outcomes or does not, and relatively unrestricted free trade in labor either promotes or harms the social interest. The answers to these questions mark the ultimate divide between orthodox and heterodox in contemporary labor economics, an area in which NE and IE occupy overlapping but nonetheless distinct spaces.

In keeping with the evolutionary character of IE, the benefits and costs of a minimum wage vary by country, stage of economic development, extent of unemployment, and the breadth, depth, and structure of the labor market regulatory regime. Certainly some of the rationales for a minimum wage advanced in the 1930s are less important today (because of successful institutional innovation and regulation!), whereas other rationales may be of equal or greater importance. A pragmatic weighing of the evidence is required, with a balanced consideration of the costs of market failure versus the costs of government (and union) failure. Furthermore, IE is not committed to a minimum wage per se, and if other means can accomplish the same purpose (for example, an Earned Income Tax Credit, universal health insurance) but in a more efficient way, then they should be adopted. Thus, on one hand IE is open about means and welcomes market solutions; on the other hand, it is also convinced that relatively unfettered laissez-faire in labor markets is detrimental to social welfare, and that some degree of social protection and regulation of labor is therefore in the public interest. Balance, pragmatism, progress, and human values are the leitmotifs of institutional economics (and industrial relations) in this and all other policy issues.

REFERENCES


UNION INFLUENCE ON VOTER TURNOUT: 
RESULTS FROM THREE LOS ANGELES COUNTY ELECTIONS

J. RYAN LAMARE *

Using voting records of several thousand people in South Los Angeles over three local elections in 2003 and 2004, the author examines the effects of political mobilization contacts by the Los Angeles County Federation of Labor on voter turnout. Many view Los Angeles as a key example of U.S. labor movement revitalization and regard the County Federation's political acumen paramount to the local labor movement's success. Using logistic regressions, the author measures changes in voter propensity based on union contact for each election. He finds that all types of union contacts (including personal visits and live phone calls) significantly affect the turnout levels of voters, particularly Latinos.

*J. Ryan Lamare is a research scholar in the Kemmy Business School at the University of Limerick, Ireland. The author would like to thank the members of his dissertation committee—Harry Katz, Ron Seeber, Lowell Turner, and especially David Lipsky—for all their tremendous support and assistance throughout the writing of this paper. He benefited also by the helpful comments provided on earlier versions of the manuscript by John Budd and Peter Gahan. The author also thanks Larry Frank, Anne Kamsvaag, and Anthony Thigpen for providing the data. He especially appreciates the advice of and help given by James W. Lamare and Jessye Wojtusik, in this and all his endeavors. Requests for data and more detailed discussion of the mobilization process may be obtained by emailing the author at ryan.lamare@ul.ie.

Political scientists suggest that the act of voting relies on interactions between a multitude of short- and long-term influences that change, to various degrees, the turnout prospects of any potential voter. Mobilization efforts enhance turnout odds by acting as a significant short-term factor in particular, and by overcoming some of the obstacles associated with political participation. Labor groups have played a role in U.S. politics for more than one hundred years, employing tactics ranging from the pluralism of Samuel Gompers to the current multi-million dollar endorsements in support of top-level (that is, presidential) political candidates and innumerable Committee on Political Education (COPE) efforts; yet only a limited literature assessing labor’s influences on turnout exists. Some recent studies have concluded that in certain circumstances, union efforts might increase turnout.

Researchers and journalists (see Rodriguez 1998; Cooper 2003; Frank and Wong 2004; Meyerson 2003, 2005a, 2005b, 2006; Milkman 2002, 2006a; Milkman and Wong 2002; Zahniser 2006) have argued that the Los Angeles County Federation of Labor (effectively the local political arm of the AFL-CIO) should be considered an immensely powerful political entity, capable of exerting enormous influence on the formation of local government. These scholars highlight the County Federation’s endorsements of pro-union candidates and those candidates’ subsequent victories, buoyed by increased voter turnout often from the Latino community, as evidence of the labor group’s political strength. Yet this proffering of success lacks a necessary empirical foundation; no researcher has tested whether the
Federation’s efforts alone—independent of typical voting determinants—have played the critical role in increasing turnout. In this paper I empirically study the Federation’s mobilization efforts in three elections, testing a number of hypotheses developed from the political science and industrial relations literatures while recognizing the strong ties between labor and Latinos in Los Angeles.

The Los Angeles Context

The three Los Angeles County elections I analyze comprised generally part of a larger scheme developed by the County Federation and more particularly Miguel Contreras, the Federation chief who took over in 1996 and was credited with transforming the organization into a political machine before his untimely death in 2005. Following the ascension of Latinos into prominent citywide union positions in the 1980s and 1990s, concurrent with the changing demographic trends of the region itself (wherein Latinos increasingly began to occupy blue-collar union jobs), the Federation shifted its political position from one of insider politics to focus on voter registration drives and targeted mobilization efforts.1 Contreras took advantage of the Latino backlash against Proposition 187 in the mid-1990s, forging an alliance with the Latino community by engaging in voter registration efforts and supporting pro-labor, pro-immigrant politicians in local and state-wide elections through targeted mobilization efforts (Milkman and Wong 2002; Milkman 2006b). 2

Given the success of its political efforts, the County Federation established a reputation as a major player in local elections, with Contreras viewed as a political power broker (Zahniser 2006). The labor group began to garner praise during the mid-2000s for its ability to assimilate immigrant communities into politics (Meyerson 2003). David Reynolds (2007: 82) noted, for example, “Union organizing and political action [in Los Angeles] has become a path for Latino empowerment, and in turn Latino political development strengthens labor.” The Federation’s vote-getting efforts have been lauded as worthy of national-level attention (Meyerson 2005a, 2005b). Further, a number of authors have tied the political gains in Los Angeles to implications for labor movement revitalization, maintaining that a virtuous cycle connects increases in union membership with coalition building and political action (see Turner 2005; Milkman 2006b; Reynolds 2007). Hauptmeier and Turner (2007: 129) asserted that the labor-Latino political and social alliance built by the Federation allowed Los Angeles to develop into “a poster child for labor movement revitalization.” Rusty Hicks, political director of the Federation, maintained that the labor group’s political mobilization efforts “were designed to help our bargaining and our organizing” and noted that there were several national-level organizing campaigns occurring during the summer of 2008 wherein L.A. was at the center of the effort because of “the political clout that had been built up in Los Angeles over time” (telephone interview, July 3, 2008). Given the weight placed on L.A. (and particularly the County Federation) by those studying union revitalization, the implications of this study, which empirically tests the labor group’s political success, may be considerable. 3

1The changing union leadership and worker demographics have been highlighted in a number of sources. See Ruth Milkman (2006a); Milkman and Wong (2002); and Frank and Wong (2004).

2Proposition 187 was a 1994 ballot initiative that would have denied a number of benefits to illegal immigrants, including public education, health care, and social services.

3For more evidence of the critical role given to the city by those studying the topic of revitalization, see also Waldinger et al. (1998); Erickson et al. (2002); and Turner (2005), in addition to the references found in the paper.
Runoff, and the March 2004 California State Assembly District 47 (or SA 47) Democratic Primary were suitable for study, moreover, not only because the same general geographic area (that is, South L.A.) was studied in each campaign and the elections were all local races featuring local candidates but also because the time frame spanned only one year, with ostensibly little change in the demographic makeup in the population. At the same time, the elections were unique enough to be measured independently. For example, in the March and May CD 10 races, the major difference was the push made by the labor movement on behalf of one candidate. In the SA 47 Primary the following year, the election occurred for an entirely different seat, the candidate was not the same as in 2003, only Democratic voters were included in the analysis, and the geographic area varied slightly from the other races.

In the March 2003 City Council elections, two candidates friendly to organized labor ran for office—Martin Ludlow in District 10 and Antonio Villaraigosa in District 14. When both Villaraigosa and Ludlow chose to run, the Federation faced a dilemma in terms of how to allocate its resources. Ludlow had been the political director of the Federation, yet Villaraigosa had long been associated with labor, having come out of the United Teachers Los Angeles (UTLA) ranks. Ultimately, the Federation decided to split its resources between the two candidates. Acting under the assumption that Villaraigosa could win his seat outright but that Ludlow would need a runoff, the Federation heavily backed Villaraigosa in the March race whereas Ludlow obtained full Federation support in the May runoff (Meyerson 2006). There was, however, still a concerted mobilization effort in support of Ludlow’s Primary in South L.A.

In the SA 47 Primary campaign, labor supported Karen Bass, a South L.A. community activist who was less associated with the labor movement than Ludlow. In fact, Bass’s work had largely involved community-based support programs rather than union campaigns. Despite this, labor was a key contributor to the coalition pushing to increase turnout in her SA 47 race. The Federation saw Bass, an African American working in South L.A., as a candidate capable of garnering support from (and possibly uniting) both the African American and the Latino populations in the area (Frank and Wong 2004). Notably, Bass’s and Ludlow’s paths diverged dramatically after being elected. Bass was elected to the California State Assembly in 2005 and elected Speaker in 2008; her victory in the SA 47 election was a significant milestone in her rise to political prominence beyond the local level. By contrast, Ludlow left his City Council post in 2005 to lead the Federation after Contreras’ death and was seen as a rising star within the national labor movement, but resigned in 2006 following allegations involving campaign violations during his 2003 election bid.

In each of these elections, it is critical to return to the point discussed above: the County Federation was particularly emphatic in mobilizing Latinos. South L.A., traditionally comprised mainly of African Americans, has seen rapid growth in its Latino population in recent years. Many of the labor leaders orchestrating the mobilization drives were Latinos, as were a substantial number of the potential voters. A large part of the labor-Latino alliance in Los Angeles involves vote-getting campaigns in Spanish-speaking neighborhoods; though Latinos are typically unlikely voters, labor has made a significant effort to bring this group into the political fold. The Latino ethnic connection between the union leaders, their members, and potential voters should not be understated. Regarding the Federation’s efforts to harness the Latino vote, Hicks maintains:

---

4 The contextual information surrounding these elections comes from interviews with mobilizing strategists and County Federation staff, and from Frank and Wong (2004) and Meyerson (2006).

5 Bass stepped down as Speaker early in 2010, having spent two years as the first black woman to head a state legislative body.

6 Both Ludlow and Bass are African Americans. The Federation supported these candidates in part because it hoped it could integrate Latino organizers and voters with African-American candidates and issues.
The Latino vote has been a large part of the success that the Federation has had in Los Angeles, primarily because Miguel Contreras had the idea to basically take janitors and hotel workers, put them on leave from their jobs, and then pay them to become the army to go out and talk to voters, specifically to Latino voters. Because of the high numbers of Latinos in the Federation’s membership, there’s no other group out there that has [a comparable] organized way of communicating with Latino voters on their issues (telephone interview, July 3, 2008).

Theory and Literature on Voter Turnout and Mobilization

The Federation’s mobilization efforts in 2003 and 2004 were concentrated within a particular setting; however, analysis of these efforts requires an understanding of the broader political science literature on voter turnout. In answering the question of why some individuals vote and others do not, Downs (1957) applied a cost/benefit analysis to voting behavior. The approach, adopted by others (see, for example, Riker and Ordeshook 1968; Fiorina 1981; Rosenstone and Hansen 1993; Hill 2006), concluded that voting is a paradox. The benefits of voting, if measured as the opportunity to influence an election outcome, always equal zero given that a single vote has never determined a major election—and yet people still vote in massive numbers. Downs and others presented a question to the political science community, which was termed the paradox of participation: if utility always equals zero, why do individuals vote?

To resolve this paradox, researchers have considered the necessity of mobilization efforts, which are designed to enhance the likelihood of participating by overcoming any obstacle that devalues casting a ballot. A rich literature on this topic has developed over time, beginning with Gosnell (1927), who was the first to explicitly examine the outcomes of mobilization campaigns. Subsequent studies tested vote-getting efforts using various techniques (see Cutright 1963; Kramer 1970; Blydenburgh 1971; Nagel 1987), all confirming the political importance of mobilization (Goldstein and Ridout 2002). Rosenstone and Hansen (1993) determined that political entities use direct mobilization (such as mailings, personal contact, and phone calls) and indirect efforts (including activating a key person in a social network that encompasses a larger web of potential voters) to overcome the paradox of participation in any single election. They maintained that the lack of political mobilization in the 1970s and 1980s was the key cause for declining turnout during this period.

Recent studies, led by Alan Gerber and Donald Green and their colleagues, have established what one might call a new paradigm for testing mobilization efforts (see Gerber, Green, and Green 2003; Green and Gerber 2005; Bergan et al. 2005). These authors have argued that the methodologies used in previous studies were fundamentally biased (Gerber and Green 2000; Green and Gerber 2005). First, the previous studies relied on self-reporting of turnout levels through surveys (without public record verification), which led to an overstatement of turnout levels by the surveyed individuals. Second, the sample sizes of these studies were often small, which led to unreliable results. Finally, the prior studies did not reveal whether the mobilizing groups had focused on likely voters, which also biased the results. The Gerber and Green cohort have asserted that the best way to analyze the role of mobilization efforts is to use an experimental approach, wherein a large sample of potential voters is randomly assigned into control and treatment groups.

Personal mobilization techniques influenced turnout, but less personal approaches such as direct mailings and phone calls were essentially ineffective in increasing turnout levels (Gerber and Green 2000). More recent studies by the same authors and their associates have supported these conclusions; although personal visits remain strong indicators of turnout in the newest studies, live phone calls have been found in some instances to raise turnout slightly (Ramirez 2005). Ironically, however, Gerber and Green (2000 and 2001) have called their own techniques into question. Relying on large out-of-state telemarketing firms to conduct phone calls, they used graduate students to perform personal contacts. The mobilizing groups had no vested interest in the out-
come of the election—the phone operators delivered the message “in a routinized and at times rushed manner,” sounding more like “a professional firm rather than local volunteers or neighbors” (Gerber and Green 2001: 77). The contacting agents’ lack of commitment may have negatively biased their results.

Turning to unions in politics, some researchers (including Uhlaner 1989; Verba, Schlozman and Brady 1995; Radcliff and Davis 2000) have maintained that union membership may influence a member’s propensity to vote, especially if the individual contacted is active in the labor organization. Freeman (2003) concluded that union members were more likely to vote than those not in unions, but that the union effect was mitigated by socioeconomic status. Delaney, Masters, and Schowchau (1988) highlighted the methodological considerations in measuring union status and voting. The results depended on the data source (either survey-based answers or responses garnered through public records). Their research, however, omits a substantial facet of union political activism: the concept of unions as mobilizing agents.

Few empirical studies have investigated the effects of union mobilization, the majority having either ignored or marginalized the role of union efforts (Radcliff and Davis 2000). Recently, however, a small body of work has aimed to fill this void. Radcliff and Davis (2000) compared state-by-state union density levels with turnout, finding that states with higher density had stronger turnout levels, though the scope of their analysis ignored the localized nature of many mobilization efforts.

Zullo (2004) measured the effectiveness of a union mobilization drive on grocery workers in Wisconsin in 2000 using a combination of public records and union contact lists. He found that contact in the days immediately preceding the election increased voter turnout. Although my approach is similar to Zullo’s, important differences exist. First, Zullo’s study controlled for age, marital status, and workplace variables, but it failed to account for vote propensity of the contacted individuals, which creates a significant limitation and proves problematic when determining the independence of the variables. Second, his study used less precise measurements of factors influencing voter turnout, relying on the employment setting and work-related variables as proxies for all the demographic and socioeconomic influences on turnout. Third, Zullo depended on broad assumptions regarding the homogeneity of the sample and measured only the effects of individual phone calls and workplace-wide political education efforts. Finally, he noted that the study could not be generalized, given the highly local nature of the case and the limited sample (Zullo 2004: 328).

**Hypotheses**

I have drawn three hypotheses from the literature and theories related to voting behavior and mobilization efforts, the Los Angeles context, and the relationship between the labor movement and the local Latino community.

Hypothesis 1: Any type of County Federation contact increased the turnout odds of potential voters.

Hypothesis 2: Personal contacts of potential voters were more successful than phone calls.

Hypothesis 3: The County Federation was particularly successful in raising the turnout levels of contacted Latinos.

Hypothesis 1 directly tests the initial question of this paper: does empirical evidence support the praise accorded to the Federation regarding its political influence? Hypothesis 2 tests Gerber and Green’s assertion regarding mobilization types, a claim echoed by others in the political science community. This study is distinct from previous ones in that the Federation’s mobilization agents were highly committed to the outcome of their efforts, unlike the “mechanical” contacting.

---

7 The concern with omitting vote propensity is that it becomes unclear whether contacted individuals voted because they were approached by the union, or because they were already intending to vote given their previous turnout patterns. It is necessary to separate contact from vote propensity to satisfy Gerber and Green’s (2000: 653) argument that “the apparent link between contact and turnout may be spurious” if likely voters are also the most heavily contacted.
groups of the studies conducted by Gerber and Green (2000 and 2001). Hypothesis 3 tests the connection between labor and Latinos in Los Angeles in terms of turnout; the labor-Latino alliance has been identified as critical to the success of citywide union revitalization.

The Data

The data set used in this study includes 188,551 individuals from the broad geographic area known as South Los Angeles.8 Labor and community groups interested in mobilizing the population of South L.A. asked a third-party organization (called Political Data) to provide a list of the potential voters in the region, which included the individuals’ names, addresses, and phone numbers, along with several known pieces of demographic information.9 A substantial number of listed people were contacted by union groups acting either alone or in conjunction with community organizations. The contacting agents were paid or voluntary union members and staff, and, when labor and community groups both mobilized individuals, some agents were active members of the various community organizations. In the three elections studied, only contacts by the Federation alone (not in conjunction with community groups) were included. The mobilizing parties were highly committed to the outcome of their efforts, given that the Federation had a vested interest in influencing voter turnout in each particular election.

Tables 1a and 1b list all the variables used in the analysis along with their coding schemes and frequency information. Turnout constitutes the dependent variable in each model; the data set provides every individual’s voting record for all elections from the 1990s through the March 2004 race. From these vote records, I have also established each individual’s vote propensity (an independent variable). Using a zero-to-five scale, I tallied vote histories for the five elections prior to the studied race. An individual who did not vote in any of the past five elections received a 0; someone who voted all five times was given a 5, and so on. Each additional increase in vote propensity would likely yield stronger odds that the individual would vote in the election of interest. The data set also provides several demographic variables for each person in the population, such as measurements of age, gender, party affiliation, ethnicity, birthplace, and affluence, which are generally predicted by the literature as relevant to voting behavior (see Campbell, et al. 1960; Lipset 1960; Verba and Nie 1972; Wolfinger and Rosenstone 1980; Teixeira 1992). These variables have been incorporated as controls in the data, so that the effects of contact can be measured independently from these influences.

Age (a continuous variable) and gender (male or female) are clearly defined; the literature suggests that older individuals generally vote more, but the influence of gender on voting is not clear (Lipset 1960; Flanagan and Zingale 1975; Milbrath and Goel 1977). Party affiliation, cited as a major contributing factor to turnout (Campbell et al. 1960; Abramson and Aldrich 1982; Hill 2006), has been divided into three groups: Democratic Party affiliation, Republican Party affiliation, and Minor Party affiliation.

Studies on this subject suggest that ethnicity has a substantial effect on turnout likelihood, with minorities voting less than non-Hispanic whites (Cassel 1979; Presser, Traugott, and Traugott 1990; Leighley and Negler 1991; Teixeira 1992). Ethnicities have been categorized as Latino, Asian, Jewish and generic. Ethnic distinctions were determined based on the last names of each individual, which is a technique commonly used by third-party sources tasked with gathering information on large populations. This approach results in some inaccuracies, given that ethnic identity cannot always be tied to an individual’s last name.10 Further, the

---

8 See the Appendix for a discussion of the filtering process that led to the N values in the regression analysis.
9 Political Data is a group that provides voter universes to campaigns, consultants and pollsters in California using voter registration rolls.
10 For instance, I use last names identified as “Spanish” as a proxy for the Latino ethnicity in this paper. Although a last name of apparent Spanish origin may not always equate to Latino ethnic identity, my methods mirror the Federation’s own mobilization techniques.
technique does not allow for discernment of generic names: a name like Smith or Johnson does not relay any type of ethnic identity. This is unfortunate in that there is no discernible African American group, given that it is difficult to identify African Americans when looking at last name alone. 11

Birthplace constitutes another variable in the data; individuals were either U.S.-born or foreign-born. It could be argued that individuals born in the U.S. would be more inclined to vote than those born in foreign countries. Finally, the data provide a measurement of each individual’s socioeconomic status. In fact, studies highlight affluence as a significant contributor to turnout (Bennett and Klecka 1970; Milbrath and Goel 1977; Verba, Nie, and Kim 1978). The data provide information detailing whether each individual lived in an apartment, which can be considered a reasonable proxy for affluence given that individuals of voting age living in apartments are generally less likely to be affluent than those who own homes.

When each potential voter was contacted, the mobilizing agents noted the contact type and the organization performing the contact. In terms of contact types, the groups used either personal visits (door-to-door walks) or live phone calls to relay their message. If a potential voter was targeted via a personal visit in any given election, he or she was not contacted over the phone, and vice versa, which allows for separation by contact type across the campaigns.

Models and Suitability Tests

To test each of the three hypotheses, it is necessary to create statistical models

---

11 One way to compensate for this is to consider the overall known demographics of the region. Roughly 40 percent of South L.A. residents identify themselves as African American; as such, I can safely assume that the majority of those individuals in the generic category are African Americans, though a definitive number cannot be determined. Techniques such as comparing census data with my figures cannot be performed due to the vagueness of the geographic area. It is critical to remember that the only known way to determine ethnicity without directly contacting all the studied individuals is to use their last names as a proxy. Given the size of the data set, it is unlikely that any errors in coding ethnicity are so large as to invalidate the use of the variable.
wherein the determinants of voter turnout are measured by a set of independent variables. In this work, the key independent variable is whether an eligible potential voter was contacted by the County Federation in each studied election. The four key variations on contact include contact of non-Latinos, contact of Latinos, contact by personal visits, and contact by phone calls. The variables regarding non-Latino and Latino contact test Hypothesis 1 and form the first model. The Latino contact variable within this model also tests Hypothesis 3. The variables regarding contact type form the second model and test Hypothesis 2. I have also included a third model that looks only at the Latino population in the data (removing all other ethnicities) and measures the relative turnout rates of contacted Latinos versus non-contacted Latinos alone, to provide an additional study of this key subset. Across all the models, I also include the remaining factors suggested by the studies mentioned above to influence turnout. Thus, the models demonstrate that voter turnout in any particular race depends upon all of the demographic variables, upon prior vote history, and, critically, upon the County Federation contact measurements, which test the paper’s three hypotheses.

It is important to establish whether the models will provide correct estimates of the coefficients to be attributed to each independent variable. One concern within a study of this type lies in the nature of the data collection itself. Potential voters in South L.A. were not intentionally divided into groups at random before being contacted. This could, theoretically, lead to a problem of cause and effect between the dependent and independent variables. For instance, if the Federation chose to contact only the likely voters, then the effect on turnout attributed to contact alone would be overstated. However, both my interviews with the Federation’s political director and my observations from the literature (see Alexander 2005) suggest that the labor group did not intentionally contact likely voters more often than any other type of voter. Further, because I included a vote propensity variable, I was able to capture and control for the vote likelihoods of individuals.

Additionally, I used a number of collinearity diagnostics (condition indexes, correlation matrices, and tolerance/VIF tests) to assess whether the variables provided truly independent measurements within each model. None of the tests indicated collinearity problems, and the outcomes showed no substantial connection between union mobilization and the other independent variables included in the models. It remains possible that there existed some influence on contact that the data could not measure (which might bias the selection of individuals into the two groups and perhaps overstate the results attributed to contact), but this work has mitigated the problem in every conceivable way by accounting for a substantial number of potential influences on

---

### Table 1b. Coding Schemes and Frequency Information for Vote Propensity Scale and Dependent Variables

<table>
<thead>
<tr>
<th>Vote Propensity Scale</th>
<th>Coding Schemes</th>
<th>Frequency Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Did not vote in past five elections</td>
<td>23.6%</td>
</tr>
<tr>
<td>1</td>
<td>Voted once in past five elections</td>
<td>20.5%</td>
</tr>
<tr>
<td>2</td>
<td>Voted twice in past five elections</td>
<td>16.2%</td>
</tr>
<tr>
<td>3</td>
<td>Voted three times in past five elections</td>
<td>14.1%</td>
</tr>
<tr>
<td>4</td>
<td>Voted four times in past five elections</td>
<td>13.7%</td>
</tr>
<tr>
<td>5</td>
<td>Voted five times in past five elections</td>
<td>11.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Coding Schemes</th>
<th>Frequency Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnout (CD 10 Primary)</td>
<td>0 = no turnout; 1 = turnout (regist. voters only)</td>
<td>32.9%</td>
</tr>
<tr>
<td>Turnout (CD 10 Runoff)</td>
<td>0 = no turnout; 1 = turnout (regist. voters only)</td>
<td>34.5%</td>
</tr>
<tr>
<td>Turnout (SA 47 Primary)</td>
<td>0 = no turnout; 1 = turnout (regist. voters only)</td>
<td>46.4%</td>
</tr>
</tbody>
</table>
turnout and establishing their independence from Federation contact.\textsuperscript{12}

\textbf{Percentage Point Differences}

The effectiveness of mobilization drives in each election can be first tested by running simple crosstabs comparing the percentage point difference in turnout by those contacted versus those not contacted. Table 2 displays these percentage point differences across the three elections for contact of non-Latinos, contact of Latinos, personal visits and live phone calls alone, as well as contacts within the Latino subset. In the CD 10 Primary, for example, a 13.5-percentage point difference characterized the variance in turnout among non-Latinos contacted by the County Federation relative to those not contacted. Similarly, in the CD 10 Runoff, turnout was 19.5 percentage points higher among those contacted compared to those not contacted. The numbers were similar in the State Assembly 47 race: non-Latino contacts were associated with increased turnout of 17.8 percentage points. Federation contacts of Latinos raised their turnout rates by 2.4, 26.1, and 1.4 percentage points in each respective election.

The results indicating when personal visits took place illustrate a different picture. In the CD 10 Primary, these were associated with a 10.6-point increase whereas personal visits in the 2004 SA 47 race resulted in a turnout increase of 4.2 points. When individuals were contacted via phone, turnout increased by 14.1 percentage points in the CD 10 Primary, and by 6.2 percentage points in the SA 47 election. Surprisingly, those who were contacted by phone voted at larger rates in the CD 10 Primary and March 2004 elections than those contacted via visits. Finally, within the Latino subset, the results also indicate that contacted Latinos voted at a higher percent than non-contacted Latinos.

It would be spurious to suggest at this point that the increases in turnout were directly caused by organized labor’s mobilization efforts. Other factors may have been more, or even exclusively, influential in affecting turnout. In order to separate these other factors from other investigate only the effects of contact on voting behavior, logistic regressions must be used.

\textbf{Logistic Regressions}

I use logistic regressions to isolate the effects of contact, the critical independent variable, from other factors that could contribute to turnout. Given that the dependent variable in each model is dichotomous, logistic regressions are suitable. Ethnicity and party affiliation are classified as categorical variables; each was assigned a respective reference point of generic ethnicity and Democratic Party affiliation.

\textbf{The City Council District 10 Primary (March 2003)}

The March 2003 CD 10 Primary was the election given the lowest priority by the

\begin{table}[h]
\centering
\begin{tabular}{lccc}
\hline
Type of Contact & CD 10 Primary (March 2003) & CD 10 Runoff (May 2003) & SA 47 Primary (March 2004) \\
\hline
Contact (Non-Latino) & +13.5 & +19.5 & +17.8 \\
Contact (Latino) & +2.4 & +26.1 & +1.4 \\
Contact (Visit) & +10.6 & N/A & +4.2 \\
Contact (Phone) & +14.1 & N/A & +6.2 \\
Contact (Latino subset) & +10.2 & +29.9 & +18.5 \\
\hline
\end{tabular}
\caption{Percentage Point Difference in Vote Turnout between Individuals Contacted and Not Contacted by the County Federation of Labor}
\end{table}

Note: No information was given for the contact types in the CD 10 Runoff election.

\textsuperscript{12} The correlation matrices, condition indexes, tolerance/VIF tests, and a discussion of the mobilization process, which supports the notion that contact efforts were independent of any influences, are all available on request.
County Federation of Labor. Potential voters were nevertheless approached by mobilizing agents, and the results of the regressions prove interesting (see Table 3). Non-Latino individuals contacted by the County Federation (Model 1) were 1.335 times (33.5 percent) more likely to turn out than those not contacted (p < .01), whereas contact of Latinos in Model 1 was not significant. Those contacted by personal visits (Model 2) had a 1.236 likelihood of turning out (a 23.6 percent increase), but individuals contacted over the phone were 1.461 (or 46.1 percent) more likely to vote; surprisingly, phone calls were more successful than in-house visits (p < .01 for both). Each yearly advance in

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstd. β</th>
<th>S.E.</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact (Non-Latino)</td>
<td>.289***</td>
<td>.060</td>
<td>1.335</td>
</tr>
<tr>
<td>Contact (Latino)</td>
<td>.332</td>
<td>.214</td>
<td>1.393</td>
</tr>
<tr>
<td>Contact (Visit)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Contact (Phone)</td>
<td>.212***</td>
<td>.078</td>
<td>1.236</td>
</tr>
<tr>
<td>Contact (Latino subset)</td>
<td>.379***</td>
<td>.081</td>
<td>1.461</td>
</tr>
<tr>
<td>Age</td>
<td>.020***</td>
<td>.001</td>
<td>1.021</td>
</tr>
<tr>
<td>Republican Party</td>
<td>-.328***</td>
<td>.055</td>
<td>.721</td>
</tr>
<tr>
<td>Minor Party</td>
<td>-.176**</td>
<td>.076</td>
<td>.838</td>
</tr>
<tr>
<td>Male gender</td>
<td>.015</td>
<td>.030</td>
<td>1.015</td>
</tr>
<tr>
<td>Latino ethnicity</td>
<td>-.338***</td>
<td>.047</td>
<td>.713</td>
</tr>
<tr>
<td>Asian ethnicity</td>
<td>-.140**</td>
<td>.065</td>
<td>.869</td>
</tr>
<tr>
<td>Jewish ethnicity</td>
<td>-.447***</td>
<td>.091</td>
<td>.639</td>
</tr>
<tr>
<td>Foreign birthplace</td>
<td>.180***</td>
<td>.045</td>
<td>1.197</td>
</tr>
<tr>
<td>Live in apartment</td>
<td>-.119***</td>
<td>.033</td>
<td>.888</td>
</tr>
<tr>
<td>Vote propensity</td>
<td>.508***</td>
<td>.009</td>
<td>1.662</td>
</tr>
<tr>
<td>Constant</td>
<td>-.296***</td>
<td>.073</td>
<td>.701</td>
</tr>
</tbody>
</table>

Note: The coefficient differences in Model 1 (between Non-Latino and Latino contact) and Model 2 (between visit and phone contact) were statistically significant, with the 95% CI of the coefficient variances equaling (1.407, 1.493) and (.936, 1.270) respectively. The omitted reference categories were Democratic Party affiliation and generic ethnicity. Missing values were excluded listwise.

*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level (two-tailed tests).

13 The coefficient differences between phone calls and personal visits, and between contact of Latinos and non-Latinos, are statistically significant in all three elections. I employed Hosmer and Lemeshow’s 95 percent confidence interval test of the variance of the differences, finding that the intervals did not include zero. The methods I used are found in Hosmer and Lemeshow’s Applied Logistic Regression (2000: 268–269). The outcomes for these 95 percent CI tests are shown in Tables 3–5, and the equations are available on request.

14 Given the sample size, significance tests should be considered generally less important in these regressions than the magnitude of the effect (particularly the odds ratio) attributed to each variable.

15 There is no R^2 value entirely analogous to that found in an OLS regression. Arguably the closest approximation is the Nagelkerke R^2, which is commonly reported in logistic regressions and is a modification of the Cox and Snell pseudo-R^2. However, R^2 values are given less weight in logistic regression than in OLS regression and their implications must be considered with caution (for more on this, see Nagelkerke 1991).
age increased by 1.021 (2.1 percent) (1.013 for the Latino-only model) the likelihood of turnout (p < .01 for all models). Individuals affiliated with Republican and Minor parties were less likely to vote than Democrats in the first two models, though Minor Party members were not significantly different from Democrats in Model 3. Ethnicity variables factored into voter turnout in the first two models, resulting in a decreased likelihood of turnout relative to the generic reference point. Those born in foreign countries were more likely to turn out than those born in the United States (p < .01) in all the models. Affluence influenced voting in the first two models, with those living in apartments less likely to vote than those living in homes (p < .01). Finally, an individual’s vote propensity appeared to be the most substantial indicator of turnout relative to the generic reference point. Those born in foreign countries were more likely to turn out than those born in the United States (p < .01) in all the models. Affluence influenced voting in the first two models, with those living in apartments less likely to vote than those living in homes (p < .01). Finally, an individual’s vote propensity appeared to be the most substantial indicator of turnout across all three models.

Federation efforts appeared to increase turnout odds substantially, particularly insofar as the role of contact relative to the control variables is concerned. The contacts provided for turnout increases ranging between 23.6 and 46.1 percent, which trailed only vote propensity in overall impact on voting likelihood. This increase in voting likelihood in an election with low priority raises an important question: How would the results differ for an election in which labor engaged in a concerted mobilization effort when essentially all of the external political factors influencing the voting population (that is, the candidates, the geographic setting, the year, and so on) remained the same?

### The City Council District 10 Runoff (May 2003)

The May 2003 CD 10 Runoff election was heavily targeted by the County Federation. The results in Table 4 suggest that contact of non-Latinos by the Federation (Model 1) almost doubled a potential voter’s likelihood of turning out, with an odds ratio of 1.830 (p < .01) whereas contact of Latinos raised turnout by more than three times (p < .01). The results therefore indicate that contacted Latinos were more likely to vote than contacted non-Latinos. No information was available regarding the relative effectiveness of personal visits and phone calls. The Latino subset (Model 2) illustrates that Federation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstd. β</th>
<th>S.E.</th>
<th>Odds</th>
<th>Unstd. β</th>
<th>S.E.</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact (Non-Latino)</td>
<td>.604***</td>
<td>.044</td>
<td>1.830</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact (Latino)</td>
<td>1.179***</td>
<td>.358</td>
<td>3.251</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact (Latino subset)</td>
<td></td>
<td></td>
<td></td>
<td>1.118***</td>
<td>.358</td>
<td>3.059</td>
</tr>
<tr>
<td>Age</td>
<td>.013***</td>
<td>.001</td>
<td>1.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Republican Party</td>
<td>–.237***</td>
<td>.056</td>
<td>.789</td>
<td></td>
<td>.006***</td>
<td>.002</td>
</tr>
<tr>
<td>Minor Party</td>
<td>–.115</td>
<td>.078</td>
<td>.891</td>
<td></td>
<td>.109</td>
<td>.090</td>
</tr>
<tr>
<td>Male gender</td>
<td>–.022</td>
<td>.031</td>
<td>.979</td>
<td></td>
<td>.140**</td>
<td>.055</td>
</tr>
<tr>
<td>Latino ethnicity</td>
<td>.126***</td>
<td>.049</td>
<td>1.134</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian ethnicity</td>
<td>.035</td>
<td>.068</td>
<td>1.034</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewish ethnicity</td>
<td>–.181</td>
<td>.095</td>
<td>.834</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign birthplace</td>
<td>.263***</td>
<td>.046</td>
<td>1.300</td>
<td></td>
<td>.433***</td>
<td>.067</td>
</tr>
<tr>
<td>Live in apartment</td>
<td>–.198***</td>
<td>.034</td>
<td>.820</td>
<td></td>
<td>.117*</td>
<td>.061</td>
</tr>
<tr>
<td>Vote propensity</td>
<td>.673***</td>
<td>.010</td>
<td>1.960</td>
<td></td>
<td>.671***</td>
<td>.018</td>
</tr>
<tr>
<td>Constant</td>
<td>–2.930***</td>
<td>.075</td>
<td>.053</td>
<td></td>
<td>–2.497***</td>
<td>.119</td>
</tr>
</tbody>
</table>

Note: The coefficient difference in Model 1 (between Non-Latino and Latino contact) was statistically significant, with the 95% CI of the coefficient variance equaling (1.183, 2.333). The omitted reference categories were Democratic Party affiliation and generic ethnicity. Missing values were excluded listwise.

*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level (two-tailed tests).
contacts also increased vote odds by more than three times relative to non-contacted Latinos (p < .01).

The other variables influencing voter turnout were age, party affiliation, vote propensity, birthplace, affluence, and ethnicity (for Model 1); gender was meaningful in Model 2 but not in Model 1. A one-year advance in age led to a 1.3 percent increase in turnout in the first model and a .6 percent increase for Latinos only. Asian ethnicity mattered in March but not May (for the non-Latino model). Similar to the CD 10 Primary, individuals born outside the United States were more likely to vote than those born in the United States, individuals living in apartments were less likely to vote than those living in homes, and vote propensity positively affected turnout.

A comparison of the March and May contact results ostensibly reveals that labor’s concerted effort to increase the vote worked. The March contacts resulted in an increase in turnout odds of about 1.4 for non-Latinos and was insignificant for Latinos whereas the May effort nearly doubled the likelihood of voting for non-Latinos and tripled the vote odds of Latinos. Though these results are helpful in answering the question of whether or not union mobilization efforts increase turnout, there are some limitations. For example, both sets of results focused on a single campaign and both involved the same endorsed candidate, which may have contributed to the similar voting patterns across the control variables. Further, the voters were contacted in close temporal proximity. It would be interesting to see whether the same patterns of labor influence on turnout existed when looking at an entirely different election in a different year.

The State Assembly 47 Primary (March 2004)

The Federation attached a considerably high priority to the 2004 SA 47 race. Though SA 47 and CD 10 overlap geographically to a degree, the populations are not identical (see the Appendix for details). Table 5 shows that Federation contact (Model 1) led to an increase in the likelihood of voting for both non-Latinos and Latinos, with odds ratios at 1.419 and 1.721, respectively (p < .01 for each). As with the May 2003 CD 10 Runoff, contacted Latinos were more likely to vote than contacted non-Latinos. Contact by personal visits (Model 2) resulted in a 1.620 likelihood of turnout whereas live phone calls led to a 1.526 odds of voting (p < .01 for both). Contacts of individuals within the Latino subset also resulted in increased odds of turnout, with a ratio of 1.714 (p < .01) compared to non-contacted Latinos (Model 3).

The estimates of the effects of the models’ other measured variables follow some but not all of the 2003 results. Age and vote propensity played a significant and positive role in influencing turnout, though gender was insignificant. The ethnicity variables were significant relative to the generic base (for Models 1 and 2) though those born in the United States were more likely to vote than individuals born in foreign countries. Affluence appeared to be insignificant in the 2004 race.

The results across the three elections demonstrate that among those contacted in 2004, non-Latinos were slightly less likely to vote than they were in May 2003 but more so than in March 2003. Latino turnout among contacted individuals in 2004 also fell between the results for the March and May 2003 elections. Though personal visits and live phone calls were not measured in May 2003, the March 2003 and March 2004 elections are useful comparators. Both in-house visits and phone calls yielded more turnout success among those contacted in 2004 than they did in 2003. Phone calls appeared to have increased turnout more strongly than personal visits in March 2003, whereas the converse was true in March 2004.

Implications and Discussion

These data analyses offer considerable implications for the study of labor-run mobilization efforts and turnout and strongly confirm opinions promulgated by labor experts studying Los Angeles. It appears that any type of contact by the County Federation increased voter turnout among the sample in the three elections, confirming
Hypothesis 1. Although many researchers have asserted this to be the case, none has empirically supported the proposition. The effects of County Federation contact fluctuated across the elections, yet this is likely to have occurred because of the weight placed on each mobilization effort by the Federation itself. Given that the effects of contact hold up across three elections and remain significant even when controlling for known influences such as age, party affiliation, ethnicity, affluence and vote propensity, the results should be given significant weight.

Two results must be considered regarding type of contact: one is that phone calls influenced turnout meaningfully across all three elections, and the other is that phone calls worked better than personal visits in one of the elections (supported by statistical tests of the coefficient differences). The first point challenges the studies of Gerber and Green (2000 and 2001), insofar as the authors determined that phone calls offered little to no contribution to turnout. The divergence in outcomes is presumably explained by different data collection techniques. Gerber and Green relied on out-of-state telemarketing firms with no clear interest in the election result to act as their phone-based contacting agents. The Los Angeles individuals who contacted voters were local community and union members with close ties to the candidates for whom they sought to increase turnout, and they often shared key similarities (particularly in terms of ethnicity and language) with the neighborhoods in which they contacted voters. My research analyzes what is probably a far more persuasive mobilizing group than those studied previously; if this assumption is correct, then it helps explain the magnitude of the results as a whole, which is larger than that reported in previous research.

The second point, that phone calls were statistically more influential than personal visits in the 2003 Primary, also needs consideration. It is possible that unmeasured factors (or their relative persuasiveness) may explain the success of phone calls relative to personal visits. Assuming my results are valid, the success of phone calls relative to personal visits may be due to the relative ease of contacting potential voters and given the time and resource-intensive nature of door-to-door campaigns and the relative ease of contacting potential voters.
via phone banks, my findings suggest that union mobilizing groups strongly consider using phone calls in some instances. The individual mobilizing group needs to decide whether the possible bump in turnout that occurs because of personal visits is worth the extra resources when phone calls appear to be highly successful.

Regarding the contact of Latinos, the County Federation was generally successful in raising turnout of this ethnic group. The Model 1 results demonstrate this notable success in the May 2003 CD 10 Runoff and March 2004 SA 47 Primary, in which turnout odds were higher (and supported by statistical tests) among contacted Latinos than among contacted non-Latinos. Similarly, Model 3 illustrates that the contact of Latinos was also statistically significant and positive when compared with non-contacted Latinos. However, the hypothesis is not confirmed across all elections, given that the March 2003 results (for Model 1) showed no statistical significance for contact of Latinos.

It has been theorized that the labor-Latino alliance in Los Angeles allows unions to affect Latino individuals markedly in terms of organizing campaigns and political mobilization drives. My results generally support this claim. Considering that many of the mobilizing agents were Latinos, any assertion regarding the success of the Federation in harnessing the Latino vote could be connected to the notion that (ethnic) similarity breeds affinity when it comes to mobilizing voters. The explanation for the differences across the elections in labor’s influence on Latino voting may be the same as the explanation offered for the sample as a whole, namely that higher voting patterns reflect the priority the Federation gave to the election.

The outcomes as a whole suggest some implications beyond the local level although the results may not be considered applicable in a blanket manner across all contexts. On the one hand, it can be argued that certain characteristics of Los Angeles highlight the uniqueness of the case studied. For instance, without specific anti-immigrant legislation, the Latino community might not have been as receptive to labor’s political overtures. More broadly, the ethnic makeup of the city (and my sample in particular) does not mirror the demographics of the United States as a whole. The bonds formed between the Federation and Latino voters in South L.A. might not be as easily developed in other contexts, largely because Latinos are not a particularly significant voting population in some locations.

On the other hand, there are growing suggestions that circumstances similar to those found in the Los Angeles context may arise in other situations. Consider the immigration battles across the United States in the mid-2000s. Massive uproar occurred in response to national anti-immigrant legislation like the Sensenbrenner Bill (HR 4437). Milkman (2006b) asserts that the reaction to the bill has mirrored the response to Prop. 187 in California, contending that, although Los Angeles may be a somewhat peculiar case, the conditions that led to Latino immigrant organizing and political action in L.A. are likely to unfold on a larger scale in the near future. Assuming Milkman’s assertions are valid, it may be disingenuous to argue that the Los Angeles case holds little relevance for labor beyond the local setting, particularly given both the continued influx of Latinos into large urban areas (such as Houston, San Antonio, Phoenix, San Diego, and so on) and the possibility of this ethnic group becoming politically active in response to anti-immigrant legislation.

Critically, the Federation’s mobilization success relied on both an invigorated local immigrant population as well as labor leaders willing both to embrace this group and to offer substantial resources in support of targeted political efforts. In settings where Latinos provide a large blue-collar and union base, the labor movement might consider building political inroads with Latino communities (following the pattern set by Miguel Contreras), possibly developed around legislative reforms. This may prove a significant roadblock in some instances, considering the apparent divide within the labor movement over immigration policy, indicated in part

---

16 A near-identical sentiment can be found in Milkman’s response to critics of L.A. Story (see Milkman 2007).
by labor’s response to the proposed guest worker programs (see Raine 2006).

The recommendations based on this analysis are not necessarily limited to Latinos. There is no particular ethnic characteristic unique to the Latinos in Los Angeles that makes them more susceptible to political mobilization than any other group. For instance, the Federation reports that it has run successful mobilization drives among African American and Asian voters in recent years, using similar methods to those used with Latinos (Hicks 2008). It could be argued that two critical components that allowed labor to make inroads within the Latino community were the perceived political threat to the group and labor’s ability to communicate with the threatened entity. The idea that a population might be awakened politically by a threat and thus be more receptive to mobilization efforts is certainly not novel, nor is it limited to the Los Angeles case. However, the ability of the local labor movement to integrate itself into this threatened population built around shared political aspirations has not yet been demonstrated in other settings, at least with the same intensity as in Los Angeles. In appropriate circumstances, the labor movement might consider developing or strengthening ties with any threatened groups who would be receptive to union-led political action, so that the individuals within these groups might be mobilized in a manner similar to that discussed in this paper. In so doing, labor could build on the example set by the County Federation in Los Angeles.

Appendix

The data set provided registered voters in “South Los Angeles,” which is not tied to one exact political boundary, but encompasses several locations, including CD 10, SA 47, and other political areas. CD 10 and SA 47 overlap to some degree but are not identical. Some potential voters lived in South L.A. but not in CD 10 or SA 47, and thus were excluded from the final sample since they were geographically ineligible to vote. A large number of individuals lived in both CD 10 and SA 47, whereas others lived in only one of these political boundaries. Further, the March 2004 election was a Democratic Primary campaign, in which those not identified as Democrats were ineligible to vote and were accordingly excluded from the final analysis. Additionally, only individuals registered to vote at the time of the election were included in the analysis. Finally, the variables for age, ethnicity, party affiliation, gender, and birthplace contained a small number of missing values, none of which biased the included data (that is, some names could not be classified into any ethnic group and a portion of the individuals declined to state their political affiliations). Any cases with missing values were excluded listwise from the analysis. When the non-eligible (by way of location, party affiliation, registration status or missing data) cases were removed, the final samples consisted of 26,736 potential voters in the CD 10 Primary, 26,456 in the CD 10 Runoff, and 31,476 in the SA 47 Primary (totaling 84,668 potential voters summed across all three elections). For the Latino subset, the samples consisted of 8,797, 8,840, and 7,652 potential voters, respectively.
REFERENCES


