Title
Essays on Behavioral Tax Policy and Political Violence

Permalink
https://escholarship.org/uc/item/82s628mg

Author
Huet-Vaughn, Emiliano Raphael

Publication Date
2013

Peer reviewed|Thesis/dissertation
Essays on Behavioral Tax Policy and Political Violence

by

Emiliano Raphael Huet-Vaughn

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Emmanuel Saez, Chair
Professor Edward Miguel
Professor Shachar Kariv
Professor Patrick Kline
Professor Ernesto Dal Bó

Spring 2014
Essays on Behavioral Tax Policy and Political Violence

Copyright 2014
by
Emiliano Raphael Huet-Vaughn
Abstract

Essays on Behavioral Tax Policy and Political Violence

by

Emiliano Raphael Huet-Vaughn

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Emmanuel Saez, Chair

This dissertation includes original research in the fields of behavioral public economics and political economy.

The first chapter provides evidence from a field experiment testing whether exposure to relative earnings information impacts worker effective labor supply. I exogenously manipulate access to information about relative position in the distribution of worker earnings as well as the shape of the distribution among workers engaged in piece rate, web-based clerical work. There are four main findings. First, those exposed to information about their placement in the earnings distribution provide significantly more labor effort on average than those with no information about peer earnings. Second, labor supply elasticity with respect to net of tax wages, a key sufficient statistic for optimal income tax policy, is unchanged between the two groups. Third, the higher productivity observed among workers exposed to relative earnings information is driven by those workers who experienced an exogenously assigned high relative earnings rank and low average comparison group earnings. Fourth, this later finding is gendered in the sense that women supply more labor regardless of whether they learn they occupy high or low relative standing while men supply significantly more labor only upon learning they occupy high relative standing. A model of worker preferences that incorporates status concerns is shown to reconcile these seemingly disparate findings in contrast to several alternatives considered. These findings suggest that governments can potentially use relative earnings information to grow the tax base - but not to affect optimal labor income tax rates - and that firms can generate significant productivity boosts simply by providing workers with information about the earnings of their peers.

The second chapter addresses an entirely different question, namely, the efficacy of violent forms of protest. It takes as its point of departure the acknowledgment that estimating the effect of violent forms of political protest on protest success is complicated by endogeneity and omitted variable bias. To address this problem, I utilize instrumental variables methods to estimate a causal effect of violent protest on the likelihood that protesters win policy concessions. Using daily French protest data and a set of weather and school holiday instruments, I find a significant and negative relationship between property destruction associated
with protests and the chance of near-term success in changing policy. The IV estimates are larger than OLS estimates and are robust to a variety of alternative specifications. Such findings are predicted by several posited endogeneity channels, and, they suggest that political violence does not, in fact, pay off.
To my parents and family, for making me someone who could do this. 
And, to Luma, for sustaining me, in uncountable ways, through the process.
Contents

List of Figures

List of Tables

1 Using Status to Motivate Workers: A Field Experiment on Relative Earnings and Labor Supply
   1.1 Introduction 1
   1.2 Related Literature 5
   1.3 Labor Market and Experimental Design 7
   1.4 Empirical Results 11
   1.5 Explaining the Results 17
   1.6 Policy Implications 20
   1.7 Conclusion 24
   1.8 References 26

2 Quiet Riot: A Causal Effect of Protest Violence
   2.1 Introduction 50
   2.2 Existing Literature 53
   2.3 Data and Measurement 55
   2.4 Estimation Framework 61
   2.5 Main Empirical Results 62
   2.6 Implications for Future Modeling 67
   2.7 Conclusion 69
   2.8 References 71
   2.9 Appendix 76
List of Figures

2.1 Plot of unique protest locations (in blue) and weather stations . . . . . . . . . . 59
2.2 Disorderly Conduct Calls as a Function of Temperature (reprinted from Rotton and Cohn, 2000a) 60

List of Tables

1.1 Descriptive Statistics .......................................................... 30
1.2 Test for Balanced Treatment and Control Groups ............................ 31
1.3 Test of Differential Attrition .................................................. 32
1.4 Relative Earnings Information and Worker Output ........................... 33
1.5 Deflated and Inflated Rank Treatments Exogenously Affect Rank .......... 34
1.6 High Rank Revelation and Worker Output .................................... 35
1.7 No Differential Elasticity Among Treated and Control ....................... 36
1.8 Estimates of the Effects of the Placebo Treatment ........................... 37
1.9 No Gender Differences Among Deflated Rank Workers ....................... 38
1.10 Average Output By Gender and Treatment Status ........................... 39
1.11 The Labor Effect of High Rank for Men ..................................... 40

2.1 Descriptive Statistics .......................................................... 58
2.2 Protest Violence and Concession to Protesters (OLS Results) ................. 62
2.3 Weather, School Holidays and Protest Violence (First-Stage) ............... 63
2.4 Effect of Protest Violence on Concession (IV-GMM Results) ................ 64
2.5 Weather-only Instruments and Overidentification Tests ....................... 66
2.6 Expanded OLS, IV, and First-Stage Results with Additional Variables ..... 76
Acknowledgments

This dissertation owes a great deal to the input and influence of my colleagues and mentors at UC Berkeley. Enumerating the many ways they have improved my scholarship is no easy task. But, as is the custom, I will try to do so here, even if done incompletely.

Many know Emmanuel Saez as the brilliant scholar whose work has transformed not only the economics profession but also the policy debates of our time. This is indeed true. His work, without fail, engages the first order questions that make me care about economics, and, he has led by example in expanding the space for research at the intersection of what is academically rigorous and what is socially useful. But, Emmanuel is also the most dedicated of advisers, seemingly driven by a real sense of duty to the students he advises. His door was always open. His feedback was always timely. His advice truly discerning and invaluable. I thank him for his constant guidance and support throughout my graduate career, and, for being an inspiration to me as to what an economist can be and can study.

Ted Miguel in many ways acted as a second adviser to me during my time at Berkeley. Just as Emmanuel’s work motivated my substantive interests in public economics, and the first chapter of this dissertation, Ted’s work helped to expand my interests to the area of political economy, including the study of conflict and political violence, as reflected in the second chapter of this dissertation. He encouraged my interdisciplinary interests and introduced me to a new academic community. Additionally, Ted’s work served as a model of how scientific research truly is an act of great creativity, the novelty in his research designs reflecting the hard-to-teach inspirational origins of research. In a personal capacity, he has been a friend and a source of strong encouragement for my work. I thank him profusely for this.

I also offer special thanks to Shachar Kariv and Gabriel Lenz. Few people are as likable as Shachar. He can eviscerate your work while still making you smile. The high standard he holds for experimental work gives a young researcher a high bar to aspire to (a welcome measuring stick), but, it also makes his encouraging words that much more encouraging when offered. His confidence in my work gave me added confidence in it and his career and job market advice was among the most illuminating one can find. His dedication to economics as a rigorous empirical science that interacts symbiotically with theory will continue to motivate my work.

From the very start, Gabriel Lenz was the most gracious, welcoming professor (and person) I could have hoped to find when crossing campus and departments in search of advice. He was immediately warm and open, sharing knowledge casually and expansively, like others do gossip. My work has benefited greatly from his impressive and vast expertise, and, I have benefited from his friendship. He always went the extra mile, making numerous calls from overseas while on sabbatical, and, offering deep and detailed feedback on drafts of my work. If I can approach even half the level of scholarship or mentorship that he demonstrates, my future employer will be very well served.

Patrick Kline, thank you for being so damn easy to talk to and for making every word in those conversations replete with encyclopedic knowledge and a relatable brusqueness. Ernesto
Dal Bó thank you for sharing your extensive intellectual range and for the seriousness with which you think about your students’ research when they sit down with you. Stefano DellaVigna, thank you for making time when there wasn’t really time, and thank you for what you are doing to expand the sphere of the profession. Matthew Rabin thank you for making that expansion possible to begin with, thanks for reading ahead of time, and, thank you for simply being you. You are beautiful. Many other professors also contributed valuable advice and wisdom, but, I want to particularly note the valuable input, guidance, and support I received from Alan J. Auerbach, Frederico Finan, Hilary Hoynes, Robert Powell, David Card, Dan Silverman, Leonardo Arriola, Jonah Levy, Peter Lorentzen, Kevin O’Brien, Alexander Gelber, and Leonard Green.

As definitive as the Berkeley faculty are to a Berkeley education, the graduate school experience is shaped probably more so by one’s fellow graduate students. A special kind of economist (or economist-to-be) comes to Berkeley to study (bearing quite a resemblance to the faculty that selected them, naturally). Creative, collaborative, interdisciplinary in orientation, generous in the sharing of knowledge, largely lacking in ego. You go to another school if you want to be ensconced in orthodoxies. You come to Berkeley to draw on the best of what is in the discipline while simultaneously expanding the frontier of what could be. You go to another school if you want to be surrounded by people who discuss economic ideas as a form of intellectual combat. You come to Berkeley if you want peers more concerned with mutual intellectual growth. I’ve been very fortunate to learn together with people such as these.

Liang Bai, you have been my friend and my consistent intellectual springboard. You are also a gem of a human being. Fred K. Ghansah you were always there for late night econometric discussions and I found in you a kindred spirit with a heart planted in the world of the humanities and the arts while our heads were buried in math and statistics texts. Youssef Benzarti, by a happy accident we taught together, leading us to become office mates, where I learned over time how surprisingly similar we are in research approaches and personalities. It’s been a pleasure discussing research and everything else with you. Vladimir Asriyan you’ve been there when it really counted. Thank you for that and for the way you still get intoxicated with the act of thinking. Ana Rocca, I’ve always appreciated how you seek to get to the depths of knowledge. I feel like I learn better when we actively learn together. Michel Serafinelli you inspire me with your drive and with your truly comradely spirit. Tristan Gagnon-Barstch, I still refer back to things you taught me first year. Much respect for your sustained nonchalance. Matt Botsch, I am always secretly impressed with how much you know about econometrics whenever we discuss it, and, by how decent and good-hearted you are as a person. Antonio Rosato, you went first and it was sad to see you go. I missed our discussions in and out of the office, but, I benefited so much from your tireless willingness to share what you’d learned. Thank you for being there for me. Miguel Almunia I think not knowing anyone in a new department, you’d be the colleague I’d want to meet first. Carl Nadler, I get you. Let’s talk more about the paths for scholarship and its use, and, hold each other accountable in the future. Xiaoyu, it’s been a pleasure maturing together. Alisa Tazhitdinova, you have been a friend and advocate, and, a most forgiving
office mate. Vico Vanasco, I think you make everyone around you better and more alive. Many thanks also to others in the program who helped me along the way including Matt Leister, Takeshi Murooka, Juan Carlos Suarez Serrato, Francois Gerard, Zach Liscow, Willa Friedman, and Josh Tasoff. And a very sincere thank you to the remarkable staff at UC Berkeley, especially Patrick Allen, Rowilma Balza del Castillo, Camille Fernandez, Vicky Lee, John Ridener, and Joe Sibol. In very tangible ways you made this work possible. I am so grateful to receive your consistent, committed support. Patrick, in particular, I am indebted to you.

I also want to acknowledge generous financial support from the Center for Equitable Growth, the Robert D. Burch Center for Tax Policy and Public Finance, and the Experimental Social Science Laboratory at UC Berkeley. And, to the people of Oakland: your struggle for justice inspired me to think seriously about how the public can most effectively get there.

Lastly, in ways that are indirect and distant, but, also foundational, David and Yolanda, I owe my education and its fruits, and more fundamentally my love of an intellectual life, to you. Degrees aren’t the thing in life that give me pride, but, if there is any sense of accomplishment here, the accomplishment is yours. Luma, finding you during this graduate career is without a doubt the greatest accomplishment of my life. Thank you for sharing your strength with me as I wrote this dissertation. I will aspire to always share mine with you.
Chapter 1

Using Status to Motivate Workers: A Field Experiment on Relative Earnings and Labor Supply

1.1 Introduction

Throughout the economy we observe firms that publicly disclose information about the earnings of their employees. This includes firms in a wide variety of industries encompassing many types of workers, from sales people at department stores to software developers at tech firms to upper management at Fortune 500 companies.\(^1\) Many governments similarly disclose the relative earnings information of workers. State governments in the United States, for instance, make the salaries of all state employees public, and, in certain Scandinavian countries, the federal government allows the reported taxable income of every taxpayer to be public and immediately accessible to anyone with access to the internet. The rationales for such disclosure vary.\(^2\) This paper considers one possible rationale, namely, that peer earnings disclosure may be justified as a matter of firm or government policy on the grounds that it positively affects worker labor supply or alters labor supply elasticity with respect to net of tax wages.

Dating back to Veblen (1899), Duesenberry (1949), and, more recently, Frank (1985), economists have been generally interested in how concern over relative standing may enter into individual preferences and thereby affect behavior. Much of the literature has been theoretical, focusing on how optimal income tax rates change when introducing preferences over

\(^1\)See Grote (2005) for several cases studies. For an example of disclosure of peer earnings information among salespeople see Barankay, 2011a. For discussion of a mandated earnings disclosure of top executive pay for Fortune 500 firms see Gartenberg and Wulf, 2013. For further discussion of pay disclosure practices among Wall Street bankers, lawyers, non-profit employees, and tech employees, among others, see Belkin, 2008; Williams, 2008; Indiviglio, 2011.

\(^2\)Often these policies are promoted by transparency advocates or on the grounds that disclosure will limit tax evasion (see Hasegawa et al., 2010, for an investigation of this second rationale).
CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY

not only absolute consumption or earnings but also relative position (Boskin and Sheshinski, 1978; Oswald, 1983; Frank, 1985; Ireland, 1998, 2001; Allgood, 2006; Beath and Fitzroy, 2007). In recent years, there has been growth in empirical work demonstrating that social comparisons do in fact matter in a variety of contexts. For instance, social comparisons have been shown to have effects on job satisfaction and job search intention (Card et al. 2012), electricity consumption (Alcott, 2011), choices over lotteries and costless charitable giving (Kuziemko et al., 2011), and self-reports of happiness (Luttmer, 2005).

While such outcomes are important for many considerations, to firms and governments looking to maximize profit and tax revenue, respectively, the decision to adopt a policy disclosing peer earnings hinges in large part on its effect on worker output. Additionally, economists, given the priority our discipline places on revealed preference, may be particularly interested in knowing whether the documented effect exposure to peer earnings has on self reports of job satisfaction and happiness translates into changes in actual worker productivity. This paper investigates this outcome using a natural field experiment (in the taxonomy of Harrison and List) where worker access to relative earnings information is exogenously controlled, allowing for a treatment group exposed to peer earnings information and a comparison group that is not. Specifically, workers in an online labor market perform bibliographic data entry work for which they are paid a piece rate (framed as a net of tax wage) for each journal article whose bibliographic information is correctly entered. As they have a fixed work period, labor supply in this context amounts to labor effort as measured by effective output, or, the number of correctly inputed journal articles. After one period of work, all workers are told of their earnings from this previous work period, while a randomly selected group of workers is also informed of how their earnings in that period compare to the earnings of a group of fellow workers. All workers then engage in a second period of work with information about own and (for some workers) peer earnings in that period of work also revealed (in the same fashion and to the same workers as before) following their work, with employment then terminating. In this setting there is no direct financial reward for improved status and a perception of indirect financial reward (from increased chance of continued employment for higher performing workers) is unlikely given the institutional setting and that workers are told the work relationship will finish upon the end of the second period of work. An additional unique feature of the experimental design is that in this setting we are able to control the makeup of the presented peer distributions and randomly assign some of the workers treated with relative earnings information a low performing peer group and some a high performing peer group. This exogenous assignment of relative position allows for an assessment of heterogenous effects of the peer earnings disclosure treatment that avoids confounding ability bias or mean reversion present in other designs.

We find that on average worker effective labor supply increases among workers who are presented information about their placement in the distribution of earnings among fellow workers, a novel experimental finding in the field. This productivity boost in the second round of work amounts to roughly 10% of average output, a significant economic effect suggesting there may be substantial gains in some settings from disclosing peer earnings information.
CHAPTER 1. USING STATUS TO MOTIVATE WORKERS: A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY

Additionally, in a first for the literature, we are able to use varying piece rates between workers to test whether relative earnings information can be used to not only increase total tax revenue, but, also, as an instrument to alter optimal marginal tax rates on labor income. In the standard optimal income tax problem, there is an equity-efficiency tradeoff whereby governments’ desire to spend on various ends is limited by the disincentivizing effect of taxes. This disincentivizing effect results from a diminished marginal reward for costly effort when a tax is increased. When workers care about absolute levels of earnings and not their earnings relative to others, this marginal reward is the consumption value of earnings, since the amount one can consume is lowered as the absolute net-of-tax earnings fall. However, if workers care about relative earnings as well, deriving additional utility from having higher income relative to a peer group, the marginal reward for the “status” value of earnings may not be diminished by an increased tax if one’s peer group also experiences the same increase in tax rates.\(^3\) One possible implication is that stimulating worker concern for their relative earnings position via disclosure of peer earnings may be a way to minimize the efficiency loss of taxation. We test this hypothesis by measuring worker labor supply elasticity with respect to the net of tax wage, a sufficient statistic for optimal income tax rates (Saez, 2001). We find virtually no difference in the elasticity for workers in the control and treatment groups, suggesting that while governments may use the disclosure of earnings information to grow the tax base it is not an effective tool to alter the efficiency costs of taxation and optimal income tax rates.

A third finding presented in this paper addresses possible heterogeneity in the response to relative earnings information. With identification assisted by the abovementioned random assignment of low and high performing comparison groups to workers in the treatment group, we conclude that the increase in worker output among treated workers is primarily driven by increased effort from those who learned they ranked higher on average in the earnings distribution (who provide significantly more labor effort than workers in the control group who received no relative earnings comparisons whatsoever). While the performance of those treated workers who learned that they ranked lower on average in the earnings distribution is less definitive for reasons discussed in the text, this heterogenous treatment effect suggests the design of the program revealing relative earnings information may matter for the optimal incentivization of workers, and that, when feasible, firms and governments may want to selectively manipulate the exact distribution of earnings offered as the comparison group earnings.

Fourthly, we also present heterogenous treatment effects among men and women, illuminating gender differences in the response to positive and negative information about one’s own relative standing. While men seem to have an advantage at the data entry work in the absence of non-pecuniary incentives (as measured by higher productivity in the control group), the gender gap in output is maintained among those workers incentivized with relative earnings information from low performing comparison groups but is muted among those workers exposed to relative earnings information from high performing comparison groups.

\(^3\)Falling tides may lower all boats, but individuals will still compete to have their sail fly highest.
While the relative earnings information treatment does not lower the average performance of men relative to the control group even when the information suggests low relative standing, the small and insignificant improvement in performance for men (but not for women) in this “bad news” case suggests that men are more sensitive to learning they occupy low relative standing.

Finally, this paper also contributes to two other strands of literature. First, a large peer effects literature now exists showing that individual performance is greatly affected by the presence of peers even in contexts where there are no team incentives and where an individual’s payoff is independent of the peer (Sacerdote, 2001; Falk and Ichino, 2006; Mas and Moretti, 2009). To the author’s knowledge, the results presented here mark the first experimental evidence of peer effects with absent peers, where a positive effect on worker productivity is observed to result simply from the introduction of information about co-workers’ performance without the co-workers being physically present at the site of work or engaging in any interaction at all. This effect seems to come about because of competitive preferences that motivate workers to improve their relative standing, and we think such a mechanism may offer one explanation for the broad class of results that have been termed peer effects: workers may be learning something about the productivity or future earnings of their co-workers in peer workplace settings and this may inspire them to compete to perform better or earn more.

A second touchpoint for this work is the nascent literature documenting labor supply elasticity using randomization of wages in the field. It is well known that observational data relying on non-random wage assignment is subject to endogeneity concerns (for instance, high skilled workers earning higher wages may work harder naturally for reasons unrelated to the wage). Despite neo-classical theory, there is a lack of robust lab evidence of an upward sloping labor supply function (see Camerer and Hogarth, 1999; Bonner et al. 2000; Charness and Kuhn, 2011), and, to the author’s knowledge, in real labor markets the theory has only been tested using exogenously manipulated wages in three existing studies (Fehr and Goette, 2007; Goldberg, 2010; Dal Bó et al., 2012). Our work contributes to this literature with experimental labor supply estimates in a new labor market, and, we develop an online platform which researchers may utilize to study questions that require conformance with basic labor supply theory.

The remainder of the paper is organized as follows. In Section II we review related literature. Section III describes the labor market and the experimental design. Section IV presents the main empirical results. Section V discusses models of worker preferences consistent with these results. Section VI discusses the implications of the findings for firm and government policy. Section VII concludes.

4This platform is currently being used for other experiments and the programming code will be made available for interested researchers upon request.
1.2 Related Literature

This work builds on the existing economic literature on social comparisons. Much of the work from the 1970s onward focuses on the negative externality other people’s consumption or income implies given individual preferences over relative position, and, seeks to derive the consequences for optimal taxation rates (Boskin and Sheshinski, 1978; Oswald, 1983; Frank, 1985; Ireland, 1998, 2001; Allgood, 2006; Beath and Fitzroy, 2007). More recently, there has been a shift toward empirical study of the behavioral consequences resulting from provision of information about relative position. This work fits most closely into this literature, which we will discuss in further detail here, focusing on the most relevant papers.

In the present study we are interested in the way information about relative standing affects performance in contexts in which there is no direct financial benefit from higher rank. Recent papers falling under this rubric use both observational and experimental data, in both the lab and the field. In the lab, Ericksson et al. (2009) find that giving feedback on the number of correctly solved addition problems done by oneself and a peer does not significantly influence the average number of correct submissions relative to a comparison group, (while, if the feedback is continuously provided, it does increase the number of errors submitted by those who were low performers in the first half of the task prior to the rank information revelation). The comparison group in this experiment, however, is not comparable to the setting here, since all participants know they will be shown their relative performance at the end of the session (allowing for an ex-ante effect of rank information to motivate even in the control group). In another laboratory experiment, Kuhnen and Tymula (2010) inform students of their relative performance at a multiplication task among 6-9 co-participants. They find that individuals work harder and expect to rank better when told they may learn of their rank (an ex ante effect of rank information), and, that individuals who rank better (worse) than expected decrease (increase) output in the subsequent round of work. As with Ericksson et al. the comparison group here is not one deprived of exposure to relative earnings information at all points in the experiment, but rather, all participants are in some rounds exposed to relative earnings information and in others not (with the possibility that in the next round they could again fall into either group), so that relative earnings information from previous or future periods contaminates the control group in any given round of work, making for a counterfactual analysis that is hard to analogize to any real-world policy alternatives. This is not an issue with Charness et al. (2010), who conduct a lab experiment showing that performance in a real effort task (the number of successes in decoding letters to numbers) is significantly higher over time (or more precisely, fails to decline in later periods of work, compared to the baseline control group) when participants are given information about how they and two other participants performed relative to one another in the current round of work (the experiment uses a random-subject matching protocol for the peer group each round, so the peer group changes). Freeman and Gelber (2010) find a similar improvement in performance, also under a fixed rate compensation scheme like Charness et al., following exposure to information about relative position in a previous round of piece rate work solving mazes. Clark et al. (2010) also conduct a
lab experiment manipulating exposure to peer comparisons to study effort choices, but, they utilize a non-real effort task and an experimental set-up that lends itself to different interpretations of the causal mechanism at work than that tested here.\footnote{Clark et al. (2010) perform a lab gift exchange game in which the “employee” is told about the income offers that four other “employees” received from their respective “firms” in the same period, allowing them to compare their income sans effort selection to that of others in the labor market, before making a decision about how much “effort” to offer. They find that conditional on own income offer, an individual’s rank in the income offer distribution more strongly determines effort (positively) than does others’ average income, and that individuals who receive higher income offers or enjoyed higher income offer rank in the past exerted lower levels of effort for a given current income offer and rank. However, the effort in this design is not real effort, but gift-exchange game effort, so the effort they are modeling should not be understood to involve the usual tradeoff we are after - that between leisure and remunerative work. Instead, effort here is a decision to split money that likely involves preferences such as reciprocity or fairness that are not present in many performance pay scenarios and go beyond the income rank concerns we want to isolate.}

In the field, there are papers by Blanes i Vidal and Nossol and by Barankay which explore similar territory. Blanes i Vidal and Nossol (2011) utilize a quasi-experimental introduction of information about relative position in the average hourly earnings distribution among piece-rate warehouse workers (including information on the max, min, and average hourly earnings among co-workers). They find that workers exhibit a significant increase in productivity (measured by daily rates of goods dispatched per hour holding order size constant) in the period anticipating the future ranking information (an ex ante effect of the rank information) and an even larger ex post increase in productivity after receiving the information. This improved performance did not occur at the expense of either a long-term decrease in the quality of the work produced or an increase in the quit rate among the workforce, and seems to have been evenly distributed across workers of high, medium, and low skill level (with some differences when breaking down ex ante and ex post effects). As the work is an event study using observational data, one cannot rule out the possibility that unobserved factors coincident with the rank information onset may have generated the change in worker performance, though the authors do convincingly dispel some possible confounding theories. In two complimentary papers, Barankay finds an opposite negative effect of peer earnings information on worker effective output. In one (Barankay, 2011a), which employs a group of online workers as we do here, the peer rank information provided is not of earnings but of accuracy at a work task for which workers are not paid on the basis of accuracy but only for completion, making this finding somewhat hard to interpret and compare to this and other existing work. In the other (Barankay, 2011b), salespersons at a furniture store earning piece rates are shown their national relative rank in sales to date (and thus imputed income rank to date given a proprietary formula for assigning commissions) and exhibit a large reduction in sales as a consequence relative to a control group of workers shown nothing about their rank. Relatedly, there is also a strand of work looking at how relative performance revelation affects test scores (Tran and Zeckhauser (2009), Azmat and Iriberri (2010), and Ashraf et al. (2013)). The first two find that presentation of academic rankings increased performance in two very different settings (Vietnamese English-as-a-second-language students and Basque high school students studying a variety of subjects), and the third finds a negative result (in
a Zambian health worker training program).

Design-wise, of the above mentioned works studying work or workplace-like settings, only the papers by Freeman and Gelber, Charness et al., and Barankay have a real-effort work task with a well-defined control group that is like the treated workers in all respects other than that they are not asked to consider peer earnings information at any point during the experiment (what we consider to be the proper test for assessing the introduction of policies of peer earnings disclosure from a status quo without them). In the discussion of the results we compare our findings to these works in greater detail, and discuss additional tests which may be used to distinguish the reason for the conflicting findings about the level effect peer comparisons have on productivity. As noted, our finding of a positive effect on worker productivity from exposure to relative earnings information is the first such finding using field experiment evidence. Additionally, regarding our second reported finding, none of the existing studies reviewed, and no others to our knowledge, have designs that allow for testing for a difference in labor supply elasticity (with respect to net of tax wages) among those treated with relative earnings information and those who are not (which our design allows). This is for a variety of reasons. In some cases, the experimental set-up has involved fixed wages only (Charness et al., 2010; Kuhnen and Tymula, 2010) while in others piece rates or wages have not varied during the period of study (Freeman and Gelber, 2010; Barankay 2011a, Barankay 2011b), and in still other cases, there have been varying wages but either because the variation is so small as not to incentive more work or because the tasks induce intrinsic motivation (Gneezy et al., 2011), or for other reasons, the wage changes yield no changes in worker output or effort (Ericksson et al., 2009; Blanes i Vidal and Nossol, 2011). We, thus, are the first to be able to consider the implications for optimal income tax policy of government policies that publicly reveal worker earnings. Finally, we are also the first to experimentally assign the distribution of peer earnings information that workers see, allowing for a clean test of heterogenous treatment effects for those who learn they rank high and those who learn they rank low compared to their peers.

1.3 Labor Market and Experimental Design

Labor Market Under Study

Our field experiment takes place in the online labor market hosted by Amazon Mechanical Turk. Amazon’s Mechanical Turk website serves as an online marketplace where employers and employees virtually meet to contract for web-based, usually short-term work, with Amazon charging employers a percentage of all wages paid for its services as the middleman. There are more than 500,000 workers registered with Amazon Mechanical Turk (henceforth, MTurk) and thousands of employers who use the site to recruit and contract with these workers. Common tasks contracted on the site include clerical work, image tagging, data entry, transcription, guided web searches, and other low-skill tasks.

Many researchers in other social sciences have utilized MTurk as an experimental site
in recent years. For studies of the incipient MTurk marketplace see Berinsky, Huber, and Lenz (2012) in political science and Buhrmester, Kwang, and Gosling (2011) in psychology. These evaluations find general consistency of results using MTurk populations with classic work in their respective fields that use more traditional subject populations. Horton, Rand, and Zeckhauser (2011) replicate three classic experiments in economics (a dictator game, the Asian disease problem, a priming effects experiment) and a simple test of labor supply with MTurk and find comparable estimates to existing work. Other recent economics papers situated in MTurk or similar online labor markets include papers by Kuziemko et al. (2013) and Pallais (2013).

The lack of face to face interaction with the employees who work online invites questions about the motivation and demographics of MTurk workers, which the above works and others have investigated. Surveys of MTurk workers by Paolacci, Chandler, and Ipeirotos (2010) find that 61% of MTurk respondents report that earning money is an important driver of their participation on MTurk, while Horton et al. (2011) find approximately 80% of MTurk respondents report that their participation is primarily motivated by money (though most of these workers use MTurk as a supplementary source of income). Other motivations include learning new skills and intrinsic interest in the work, not unlike other workplace settings. In terms of demographics, the MTurk worker population is found to be more representative of the general American population than traditional experimental subject pools, with gender, race, age, and education matching the population more closely than commonly used college undergraduates samples (Paolacci et al., 2010; Berinsky et al. 2012). Still, compared to the general population, MTurk workers analyzed in these other works are on average somewhat younger, more educated, more female and with less income, while having a similar fraction of white and non-white workers and broadly similar geographic dispersion (Paolacci, Chandler, and Ipeirotis, 2010; Berinsky et al. 2012). In our sample, gender is more closely balanced and similar to the population average, as is the share of white and non-white workers, while both mean age and income are lower than the general population (Table 1 of the Appendix).

The MTurk platform has several distinct advantages for this study when compared to both laboratory studies and many other alternative labor markets. For one, in comparison to lab studies, MTurk increases internal validity by allowing workers to complete the experiment without interacting with the experimenter, and in many cases not knowing they are in an experiment at all - thus, removing experimenter bias, or demand characteristics (Orne, 1962). In our case, the average worker spends less than ten seconds on the screen for informed consent, hardly enough time to do more than perfunctorily check the box that they have read the long consent form and wish to continue, making it very likely that workers do not perceive the work they are undertaking to be part of an experiment.6 Furthermore, since we registered with MTurk as an employer with our own firm name and external work website we appeared to MTurk workers as just another employer. Secondly, compared to other workplace settings in the field, the geographic dispersion and anonymity of MTurk workers makes it

6Moreover, consent forms themselves are not anomalous on MTurk, with MTurk employers using release of liability forms to avoid future possible legal actions, as is common in other online applications.
very unlikely that there is contamination of the control group, a particular concern in settings of information treatments such as ours, or, spillover effects whereby untreated individuals are affected by the treatment indirectly (see Duflo and Saez, 2003 and Miguel and Kremer, 2004). Thirdly, MTurk also allows for the creation of faux enterprises (see Kube et al., 2012; Falk and Ichino, 2006) with relative ease, allowing us to have complete control over worker pay, incentives, and exogenous variation in the comparison groups presented, without mediating (dare we say meddling) partnered firms imposing constraints on the optimal research design. This permits clean identification of the causal effect of the relative earnings information. Lastly, the online nature of the work provides ample opportunity for on-the-job leisure and substituting away immediately from effort to another tab in the browser containing a favorite website when the compensation for work is sufficiently low. This is an attractive feature when studying elasticity of labor supply with respect to the net of tax wage, and approximates the assumptions of the neoclassical labor supply models over continuous consumption and leisure space.

Experimental Work Task

The disincentivizing effect of taxation is premised on an upward sloping labor supply function. To conduct a useful test for whether relative earnings information can be manipulated to minimize the efficiency loss of taxation, and to test for differential labor supply elasticity when relative income considerations are made salient, we must first have a task that ordinarily responds positively to monetary incentives. However, finding real effort tasks that demonstrate the expected upward sloping labor supply function with increasing pay is not a trivial matter. Camerer and Hogarth (1999) and Bonner et al. (2000) survey the economics and psychology literatures testing the effects of financial incentives on performance in a host of lab experiment settings. They find performance in a wide variety of tasks is not positively affected by increased compensation, and that only a few classes of laboratory tasks exhibit a positive relationship between piece rate and the level of performance. One class common to both surveys is clerical tasks, which inspire little intrinsic motivation, require little skill, and where effort increases actually improve performance. As such, we choose a simple clerical task to use in this experiment.

The task is a bibliographic entry task, modified form Tonin and Vlassopoulos (2012). The work involved filling in the correct bibliographic information for academic articles (author name, journal, article title, etc. for a series of published articles) and workers were paid a piece rate for each correctly entered article. See Appendix Figure 1 for an example. This task is derivative of earlier typing tasks, such as that by Swenson (1988) and others.\footnote{Charness and Kuhn (2011), in a review of laboratory labor experiments, describe the Swenson work as the “first laboratory experiment to examine labor supply response to wage changes among humans that is couched in economic theory.” In it, subjects are paid a variable piece rate for the number of exclamation points (followed by hitting the enter key) they type out on a computer that requires sequential hits of the return key (to prevent active leisure during the experiment if continuously holding down the keys was allowed). Sillamaa (1999) reports a replication of Swenson’s experiment and finds an upward sloping labor}
CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:  
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY

an ideal work task since the only special skills needed to perform the work are the ability to type and read - prerequisites that any internet user, and thus MTurk worker, will possess. Thus, incentives, whether pecuniary or non-pecuniary, will not be rendered irrelevant due to ability limits when in fact they could potentially motivate in other settings. The data entry work is also incredibly simple and requires little attention to understand, thus rendering one possible drawback of an online platform - that subjects can not be sufficiently instructed by an experimenter or manager in the proper rules of their work - largely moot. Lastly, differences in ability and effort will generate a wide dispersion of performance.

Experimental Design

After registering with MTurk as an employer we posted an advertisement (see Appendix Figure 2) on the MTurk employer advertising bulletin where potential workers can scroll through alternative job postings and read short descriptions of the work opportunities. In our advertisement, workers were informed of the opportunity for work at a bibliographic data entry task for which they would be paid a piece rate plus a flat one dollar fee for answering a few worker survey questions. The nature of this work is very similar to other work posted on MTurk, and by all appearances comparable to work offered by many other MTurk employers. Our advertisement indicated the work would take about 45 minutes to an hour and interested workers were told to access the work through an external website. Once on the website workers were presented with longer instructions (Appendix Figure 3a-c) about the work task and the actual piece rate per correctly inputed article for the first twenty minutes of work. Following the instructions, workers were asked about their earnings expectations (see Appendix Figure 4) and then a 20 minute work period began (with the work screen looking like Appendix Figure 1 with the addition of a clock counting down the 20 minutes). Following this, a series of demographic questions were asked and workers were informed of their earnings in the previous period of work. Then a second 20 minute round of work began, with workers earnings expectations for this round of work again elicited immediately beforehand. In the second round of work some workers received higher or lower piece rates than in the previous work period and some workers’ piece rates remained constant (workers were told in the instructions that their wage in the second round of work may or may not be the same as that in the first and would be determined independently of their performance in the first period of work). The piece rates were framed as net-of-tax wages with workers told ”any withholdings required by Amazon consistent with state law will have already been applied, so the entire X cent bonus per correct article is yours, like an after-tax wage.” Following the second round of work, workers were informed of their earnings in this round of work and then the work relationship was terminated.

supply function over the entire wage range (whereas, Swenson’s was upward sloping but backwards bending at the highest piece rate). Similarly, Ariely et al. (2011) report an upward sloping labor supply function for an alternative typing task involving repeated typing of certain letters. The bibliographic data entry task used here is a far more plausible task for an employer to pay works to undertake.
CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:  
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY

Piece rates were randomly assigned to workers as they successively entered the external website, as was the relative earnings information treatment status. Following each round of work, those in the treatment group were exposed to information about co-workers’ earnings in the previous period of work in addition to information about their own earnings in the previous period (information they are told they would receive in the instructions), while those in the control group learned only about their own earnings in the previous work period. For an example of what these screens looked like see Appendix figures 5 and 6. The sole difference between the control and treatment groups was this additional exposure to earnings information about the distribution of earnings for a comparison group of 4 co-workers, and, the solicitation of two additional earnings expectations questions prior to each round of work (one eliciting the worker’s expected earnings quintile relative to co-workers working for the same piece rate performing the same work, and, the other asking what earnings were expected on average of this comparison group; see the second two questions in Appendix Figure 4). Those receiving the peer earnings information treatment were randomly assigned into two different types of peer earnings information treatments (without any knowledge of the distinction): one in which the peer group was drawn from low-performing peers and one in which the peer group was drawn from high performing peers, leading to a greater chance that workers assigned to the former group learn they rank highly relative to their co-workers (and the reverse for those assigned to the later group). We call these two different types of treatments inflated rank and deflated rank treatments, respectively. Pay was independent of earnings rank for all workers, and the workers were clearly told as much at the outset. Piece rates ranged from 2 cents to 16 cents per article, though workers only knew of their own piece rate in a given period of work and not about this range.

1.4 Empirical Results

We conduct a balancing test (Wilcoxon-Mann-Whitney rank sum test) to test if randomization of treatment status leads to balance of observables, as would be expected with successful randomization. The findings, presented in Table 2 of the Appendix, suggest that based on observable characteristics like gender, age, race, and income, randomization was successful. In the first 20 minute round of work (when treated workers expect relative earnings information but have yet to receive any) there is some evidence of higher performance by the treated workers, indicating a positive ex ante effect of relative earnings information. The significance of this ex ante effect is less robust than the ex post effect, and in all that follows we concen-

---

8A test for whether these additional questions may themselves have had an impact on labor supply independent of the relative earnings information is performed in Section IV.

9In fact, what is referred to as an ex post effect here and in the literature review actually encompasses both the effect of the revelation of relative position information from the previous period (a pure ex post effect) and the effect of anticipating information revelation about the current period of work, a period 2 ex ante effect. In a policy context it is this aggregate effect that is relevant, since for any ongoing policy of peer earnings disclosure there will be both an effect of the previous period’s disclosure today and an anticipation of the current period’s disclosure tomorrow.
trate on the round 2 performance of workers (the period after relative earnings information exposure for the treated workers). We note though that there is no evidence that the robust positive second period effect documented below is offset by a dip in the productivity of the treated workers ex ante (an important consideration for policy implementation), and that, if anything, workers may be motivated by competitive preferences in anticipation of their rank revelation at the announcement of the onset of such a policy.

Necessarily, we restrict the analysis to those workers who took up the treatment/control assignment and did not leave the website, accounting for approximately 75% of all initial visitors to the introductory page of the work website and 95% of those visitors who began work, a high participation rate. There is no evidence of differential attrition between those assigned to the treatment group and those assigned to the control group, as can be seen in Column 1 of Table 3 of the Appendix. While Amazon takes efforts to assign only one unique MTurk worker ID to each person, in the analysis that follows we drop observations whose work comes from the same IP address, as is standard with MTurk studies, to prevent the possibility that these may be repeat workers using different MTurk worker IDs. We also focus on workers who supplied positive levels of output in each period of work to allow for intensive elasticity estimates and a meaningful comparison of actual earnings in the treatment (though the results are not significantly affected by this).

Results in Table 4 of the Appendix indicate the positive effect relative earnings information has on worker output. In column 1, revelation of last period’s relative earnings information is found to increase the number of articles correctly entered in the second period by 2.4 articles (controlling for second period wage level). Results are significant at the 1% level and remain so upon the addition of demographic control variables (column 2) and experimental session controls (column 3). In column 2, with additional age, gender, race, and income controls, relative earnings information exposure increases worker output by 2.85 correct article entries, or approximately 10% of average worker output. In column 4, the sample is restricted to those workers who provided MTurk IDs that could be matched by Amazon. This leads to the exclusion of only 7 observations from workers who did work but provided missing or erroneous MTurk IDs that could not be matched to those in Amazon Mechanical Turk’s database of workers who accepted our work. The results do not change in this sample. Finally, in column 5 the regression is run with the log of worker output in period 2 and the log of period 2 wages in order to look at percentage changes and calculate a labor supply elasticity with respect to wage rate. With full compliance, given that all treated workers must go through the relative earnings revelation screen, and with an improbable

---

10 Table 3 reports results on the attrition of control and treatment groups in the second round of work for those workers who went through round 1 of work and were exposed to information about own (and, for treated workers, peer) earnings, thus assessing whether the relative earnings information revelation led to differential attrition. There is similarly no evidence of differential attrition if instead looking at all entrants to the website, including those who never took up work but only received initial instructions, though coefficients and t-stats do rise in magnitude.

11 Restricting analysis to these workers does not significantly alter the differential attrition results or the balancing test.
chance of contamination, given that workers are geographically dispersed, anonymous and not working together, the estimates on the relative earnings information treatment constitute a local average treatment effect (Duflo et al., 2006).

While the above results are perhaps the most policy-relevant results, demonstrating that earnings transparency leads to significant productivity boosts overall, there remains the question of whether the treatment effect is driven primarily by those who learn they were low ranked or those who learn they were highly ranked. To test this, the experiment was designed, as previously mentioned, so that treated workers were randomly assigned either a low-performing or high-performing comparison group (unbeknown to them), thus, on average generating artificially inflated or deflated ranks, respectively, among workers who should be similar otherwise due to randomization. Table 5 in the Appendix shows that such an assignment did indeed generate better revealed relative performance for the inflated rank treatment vis a vis the deflated rank treatment, with about 65 percent of workers in the former group learning they were in first place compared to their peer group in the previous period of work, and about 60 percent of workers in the later group learning they were in last place. Column 2 in Table 3 shows there is no differential attrition among the two different treatments.

Table 6 of the Appendix reports the heterogenous effect of relative earnings information on labor supply, showing how the productivity boost among treated workers observed in Table 4 appears to come primarily from those treated with a low-performing comparison group and who, thus, on average learned they were ranked highly. In columns 1 and 2 of Table 6, we document how those workers given inflated ranks increase output by approximately 3.5 articles relative to the control group, with results that are highly significant both with and without controls. In the case where we control for demographic variables (column 2) the increased output for the inflated rank workers amounts to an increase of 14% of the average worker performance. There is a smaller but still positive coefficient on the indicator for assignment to the treatment with artificially deflated ranks. The performance for this deflated rank treatment is not quite significantly different at the 10% level than that of the control group unexposed to relative earnings information in the specification without demographic controls (Column 1 of Table 6) while it is significantly different at around the 6-7% level in Columns 2 and 3 where demographic controls are added to the base sample and the sample excluding workers without valid MTurk IDs. The p-value for the hypothesis that the inflated and deflated rank coefficients are equal is about 0.20 in these specifications and the inflated rank workers correctly input about 1.5 more articles than the deflated workers (about 5% of the average output). If more observations are added (not reported) by allowing for a less restrictive exclusion rule for observations coming from the same IP address (excluding the second observation from an IP address but not the first, as long as the first has completed all work periods before the second’s entry - an exclusion rule not without problems\textsuperscript{12}) then the p-value for this hypothesis approaches 0.08. Taken in total,

\textsuperscript{12}The decision to exclude all observations from the same IP address is undertaken to avoid repeat workers who might enter as observations twice in the data by using multiple MTurk IDs. The alternative exclusion
CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY

the evidence inclines us to believe the deflated workers supply more labor than the control
group and less than the inflated rank workers, though, we can not reject at the standard 5%
level the null that they supply labor at the same level as the control workers nor the null
that they supply the same level of labor as the inflated rank workers (though we know both
nulls can not be true given that the inflated rank workers do supply significantly more labor
than the control workers).

Causal inference from these results hinges on the successful randomization of the low-
performing and high-performing comparison groups to workers who on average would be
equally productive otherwise. We can perform a placebo test to see if first round work
was somehow significantly different between these two groups of treated workers. Given
that the assignment of low/high performing comparison groups takes place following the
completion of work in round 1, a significant difference between those with low and high
performing comparison groups in their round 1 performance would raise questions about
whether we can actually attribute the significant and higher output relative to the control
group among workers with inflated rank to information about their inflated rank. However,
we find no significant difference in round 1 output for treated workers later assigned to low-
performing and those later assigned to high-performing comparison groups (the coefficient
on the difference in round 1 performance between the inflated and deflated rank workers is
about 0.25 and the t-stat is close to zero, results unreported).

Taken together, the results in Table 4 and 6 provide robust evidence that exposure to rel-
ative earnings information increases worker productivity (in a statistically and economically
significant way) and inclines us to interpret the average effect as likely driven by workers
who experienced a favorable ranking following the previous period of work. In Table 7 of
the Appendix, we show that the elasticity of labor supply with respect to net of tax wage
changes are almost identical for the treatment and control groups. This is confirmed by a
Chow test, with the null that the elasticities are the same, yielding a p-value of over .9.
While relative earnings information affects level of output, as shown in Table 1, it does not
seem to lead to differential sensitivity to wage changes. Table 7 also provides estimates of
the labor supply elasticity in the field using exogenously assigned wages. The result (an
elasticity of 0.16) conforms with moderately sized elasticities found in existing observational
studies (e.g. Blundell et al., 1998).

To test whether the higher productivity in the treatment group results from asking treated
workers to consider their peer earnings (see the second two questions in Figure 4 of the
Appendix) rather than just from the policy of relative earnings information disclosure itself,
rule avoids a repeat user but does not prevent systematically different behavior by users who plan to work
on the same task again after finishing it the first time. Workers who somehow have access to another MTurk
ID and intend to use it to work at the same task a second time may behave systematically differently than
other workers. For instance, if receiving a high piece rate in round 1 and a low piece rate in round 2 the
worker, planning to work again, may decide to lower her effort in the second period of work anticipating the
ability to repeat the task at the high round 1 piece rate again in the future, thus biasing down the actual
effect of the round 2 wage in comparison to all the other workers who follow the MTurk protocol of one ID
per person and only one assignment of the work task per person.
we conduct a placebo treatment to test whether there is a priming effect from these additional earnings expectation questions. In the placebo treatment workers are never shown any peer earnings information or told to expect it, however, in advance of each work period they are asked to consider what a peer group doing work under the same instructions might earn (as the actual treated workers are). Results (presented in Table 8 of the Appendix) suggest that the placebo treatment has little effect on labour supply.

Revisiting Table 7, one sees an interesting gender difference in performance: men are more productive in the control group but in the treated group there is no significant difference in productivity between the sexes. At first glance, this seems to be at odds with recent literature (for a review see Niederle and Vesterlund, 2011) that finds women performing worse than men in some competitive settings when there is no gender difference in comparable non-competitive environments (Gneezy et al., 2003; Günther et al., 2010; Shurchkov, 2011). If workers indeed care about relative standing then the treatment group is in a competition of sorts for status and one might expect, given the existing literature, men and women to perform comparably in the control group and men to perform better in the treatment group - the opposite of the finding. To make sense of this discrepancy, in Table 9 of the Appendix we decompose the treated workers into those treated with co-worker earnings information drawn from low-earning peers and those treated with co-worker earnings information drawn from high-performing peers. The table shows that among the former group, the inflated rank workers, men perform better than women by about as much as they do in the control group. It is among the deflated rank workers - the majority of whom learned that they were in the bottom of the earnings distribution (see Table 5) - that the significant male advantage in productivity disappears (the point estimate on the male indicator variable is small and the estimate has a very small t statistic). This suggests that men may have a natural advantage at the particular data entry work task at hand (as evidenced by the gender difference in the control group), which they can maintain in an environment that elicits competitive preferences and increased performance from both genders (the inflated rank treatment), but, that this advantage is offset when the relative earnings information revealed in this competitive environment indicates the worker occupies a low relative standing. Table 10 of the Appendix shows that this closing of the gender gap comes about from men in the deflated rank treatment increasing their average effort (relative to the control) by less than the men in the inflated rank treatment do, while women increase their effort by the same amount on average in both deflated and inflated rank treatments. Among men, the difference in output between the inflated and deflated rank treatments is significant and the difference between deflated rank and control worker output is not. Among women, both deflated and inflated rank workers perform significantly better than control workers but not significantly different than one another. We conclude that men are more sensitive than women to revelation of

For some workplace settings the distinction is immaterial. For instance, in sales jobs where every time workers log in to their work stations they view daily performance and relative performance information, thus, naturally invoking comparisons with co-workers. It is also not hard to imagine a policy of relative earnings revelation by governments that invites people to consider their relative placement as well (thus, potentially priming them).
their low relative standing, with a bad rank lowering their performance relative to a good rank - but note that whatever demotivating effect the low relative standing information has on men, it is limited (perhaps by a desire to compare favorably the next time the earnings distribution is revealed) so that the male deflated rank workers do not ever perform worse than the control group on average.

As a final illustration of the importance of high relative standing among men, the type of treatment (deflated or inflated rank) is used to explicitly instrument for the effect of informing a worker that she/he is in first place. Such an instrumental variables analysis can reveal the actual effect of being informed that one is in first place as distinguished from confounding factors that result since one’s place in the earnings distribution is generally a function of not only peer earnings but also one’s own earnings in the previous round of work, which may be correlated with this period’s earnings through an omitted ability variable or mean reversion. As the inflated and deflated rank treatments are exogenously assigned and uncorrelated with the outcome or other potential influences on it, and as there is a clear first stage relationship between own rank in the revealed earnings distribution and deflated vs. inflated rank treatment assignment, IV estimation is a natural extension of the previous results. Table 11 presents the IV estimates of the effect of learning that one is in first place in the earnings distribution rather than another place in the distribution. This leads to an increase in subsequent output of approximately 6.5 articles for men, an estimate that is statistically significant at the 1% level. This represents the causal effect for those who would not have been in first place if compared to the high performing peers in the deflated rank treatment, but who receive a first place assignment when compared to the low performing peers in the inflated rank treatment. This may include people who would have moved from 5th, 4th, 3rd, or 2nd place to 1st as a result of being assigned to the inflated rank treatment rather than the deflated rank treatment. Clearly, for men, learning that one has high relative standing can motivate greater effort relative to learning that one has low relative standing. For women there is no such significant effect.

Our initial result showing that individuals work more when exposed to information about relative standing is consistent with the lab experimental findings of Charness et al. (2010) and Freeman and Gelber (2010), but at odds with the field experimental results of Barankay (2011a, 2011b), who finds a negative effect from exposure to information about relative position. A word about the possible reason for this difference is in order. One possibility is that the work settings where a negative finding is observed are ones in which there is an ability ceiling reached in the performance of the work task absent the relative earnings information. If this is the case and there are heterogenous treatment effects, then those who might be positively incentivized by the relative earnings information have no room to improve their performance, while those (even if a few) who might be negatively incentivized by the relative earnings information will lower their effort, bringing down average output. This seems plausible in the multi-tasking setting studied by Barankay (2011b) where workers are furniture salespeople, since the job is one in which the mapping from effort to actual effective output (furniture sold) is not one to one and one in which salespeople may be at a loss for ways to effectively improve their sales in the face of considerable chance in their sales
efficacy. One way to test this theory would be to use the design in this experiment but with a work task that does not respond positively to monetary incentives, suggesting that perhaps an ability limit has been reached that monetary incentives can not overcome. If in this setting the introduction of a non-monetary incentive like relative earnings information leads to a negative or absent effect on output then this would be consistent with the explanation offered here. In a related ongoing experiment performed with a memorization task in a lab setting, we find only weak evidence of monetary incentives improving performance and no evidence of improvement of performance among workers exposed to relative earnings information. However, the sample size is small and the signs and significance of the coefficients on wage and relative earnings treatment are such that they are tending toward those reported in this paper, and with the collection of more data it is unclear if this ability-ceiling story will continue to be supported.

1.5 Explaining the Results

The previous section revealed three findings that must be explained by any model of worker preferences and labor supply in this setting.

Finding 1: Those exposed to relative earnings information supply more labor effort on average than those without access to this information.

Finding 2: The increase in labor supply for those exposed to relative earnings information seems to be driven primarily by those shown lower performing reference groups (resulting in inflated rank for themselves) in the previous period of work. These workers supply significantly more labor than control group workers, while for those workers with deflated ranks there is weaker evidence that they exert effort somewhere between the control and inflated rank group on average.

Finding 3: There is no significant difference in labor supply elasticity with respect to wage changes for those treated and untreated with relative earnings information, and both are positive.

In this section we provide a simple stylized, descriptive model of worker utility that can explain the results observed in this field experiment. It is by no means the only specification that may be consistent with the results, but it does explain the observed findings better than certain other plausible candidate models, which we discuss briefly as well.

To derive worker labor supply and labor supply elasticity with respect to net of tax wages, we use a standard model of effort provision that imposes a separability assumption between consumption, or, post-tax earnings, and cost of effort. Specifically, worker utility is taken to be quasilinear in post-tax earnings \( c \) with a constant wage elasticity of labor supply and an additive term indicating relative earnings considerations, \( s(c, \bar{c}) \)

---

14These findings summarize an “average” worker’s behavior ignoring the gender differences and attempting to summarize the findings presented in the discussion of Tables 4, 6, and 7 in the previous section, though, as we note below, the model presented can nest the different observed behavior for men and women.
where $\bar{c}$ is the average of peer post-tax earnings and $s(c, \bar{c})$ is weakly increasing in $c$, meaning one’s status utility is never negatively affected by one’s own increased earnings. The parameter $\lambda$ represents the weight worker preferences place on relative earnings concerns, with a larger $\lambda$ indicating greater concern with the status-value of earnings relative to the consumption-value of earnings. For a worker in the control group who is not exposed to relative earnings information, $\lambda$ is presumed to be zero since no co-worker peer group was ever mentioned to these workers, and, a decision about the allocation of consumption and labor seems unlikely to include concerns about an out of sight and out of mind comparison group. The information about co-worker earnings then serves to “turn on” concerns for relative earnings $s(c, \bar{c})$.\(^{15}\) In the case of $\lambda = 0$, the preferences are standard neoclassical preferences, and, optimal labor supply\(^{16}\) from maximizing (1) subject to the budget constraint $c = (1 - \tau)w^g$ is

$$e(w) = \left(\frac{w}{\theta(1+1/\epsilon)}\right)^\epsilon$$

where $w^g$ is the gross wage, $1 - \tau$ is the net of tax rate, and $w = (1 - \tau)w^g$. The corresponding labor supply elasticity with respect to $w$ is $\epsilon$.

Workers in the treatment group, who are exposed to relative earnings information, maximize the expected value of (1) subject to the budget constraint, now with a positive value of $\lambda$. The sub-utiliy function $s(c, \bar{c})$ now matters, as does the distribution of worker beliefs about peer earnings, which we summarize by beliefs about the average peer earnings $\bar{c}$. We simplify by assuming there are two states of the world that a treated worker believes exist, one where $\bar{c} = \bar{c}_L$ (that is, the belief that he faces a low-earning peer group) that occurs with probability $P_L$, and, one where $\bar{c} = \bar{c}_H$ (the belief that he faces a high-earning peer group) that occurs with probability $P_H = 1 - P_L$. Status concerns have been modeled before in the literature using the simple functional form $s(c, \bar{c}) = c - \bar{c}$ (Charness et al., 2010; Allgood, 2006; Clark and Oswald, 1998) and we modify this functional form so that

$$U(c, e, \bar{c}; \theta, \lambda) = c - \theta e^{1+1/\epsilon} + \lambda s(c, \bar{c})$$

which allows for the

\(^{15}\)This assumption is used to explain Finding 1, but it need not hold. It is possible that workers may be constantly comparing themselves to hypothetical peer groups even in the absence of any mention of them. We think this unlikely, but, even in this case, Finding 1 can be shown to hold under less restrictive conditions that depend on the specification of $s(c, \bar{c})$ used. In the specification used in the text, the results will hold for some positive values of $\lambda$ for control workers as a function of the values of $P_L$ for control and deflated rank workers; however, going as far as giving control group workers the same priority on status as treated workers (the same $\lambda$) yields the prediction that control workers work more than deflated workers, a result at odds with the data.

\(^{16}\)Given the quasilinear preferences assumed here uncompensated labor supply equals compensated labor supply.
possibility that the marginal status utility of one’s own additional earnings is non-linear. Optimal labor supply for the treated workers then becomes

\[ e(w) = \left( \frac{w(1 + P_L \lambda + P_H \lambda^1)}{\theta(1 + 1/\epsilon)} \right)^\epsilon \]  

Thus, comparing (2) and (3), we see that Finding 1 implies that at least \( \lambda \) or \( \lambda^1 \) is greater than zero given that \( P_L \lambda + P_H \lambda^1 \) must be greater than zero for the finding to hold and given \( P_L, P_H > 0 \) and the assumption \( s_c \geq 0 \). Workers exposed to relative earnings information provide more effort (all other things being equal) since when compared to some group of peers their concern with relative standing provides additional incentive to marginal effort in the form of not only increased consumption returns on earnings, but, now, also, increased relative standing and status returns. Finding 3 can be shown to follow from this specification as well, with worker labor supply elasticity with respect to net of tax wage equal to \( \epsilon \) for both workers with labor supply defined by (2) and workers with labor supply defined by (3).

Finding 2 can be seen to follow from the expression in (3) if those workers who were exposed to low-performing peers following round 1 performance (the inflated rank workers) have a larger \( P_L \), and thus smaller \( P_H \), than the deflated rank workers exposed to high-performing peers. This seems intuitive and if it is true that the different information they received indeed changed their expectations, as we would expect, then the inflated rank workers should expect their average peer earnings to be lower than the deflated rank workers, an assumption that is confirmed by the data. A rank sum test testing the null that inflated and deflated rank workers have the same reported expectation of peer earnings in the second round of work (following the relative earnings revelation for these two types of treated workers) has a p-value of 0.01, indicating that the larger reported expected average peer earnings for the deflated rank workers is significantly larger. This fact together with Finding 2 implies \( \lambda > \lambda^1 \).

It is worth noting that the successful prediction of all three findings is sensitive to the choice of \( s(c, \bar{c}) \). An alternative definition of \( s(c, \bar{c}) = \frac{\bar{c}}{\theta} \) that has been presented elsewhere in the literature (e.g. Boskin and Sheshinski, 1978) can explain Findings 1 and 2 easily. Finding 3 can also be explained, but, only if limiting the peer groups being compared or assuming utility maximizing behavior that is less than perfectly rational.\(^\text{17}\) Other candidate models of

\(^{17}\)For this functional form \( U(c, e, \bar{c}; \theta, \lambda) = c - \theta e^{1+1/\epsilon} + \lambda s(c, \bar{c}) = c - \theta e^{1+1/\epsilon} + \lambda \bar{c} \) the labor supply function for treated workers is \( e(w) = \left( \frac{w(1 + \lambda \bar{c})}{\theta(1 + 1/\epsilon)} \right)^\epsilon \). This can be shown to lead to an elasticity of \( \epsilon \) as well, but only if we assume workers take into account wage changes that affect them without thinking about the global effect of a wage change on other’s earnings. If otherwise, in this specification we would see an elasticity for treated workers that is less than that of control workers - contrary to the findings. Outside of worker myopia, such an assumption is intuitively more plausible in a workplace setting where an individual worker may receive a raise while the peer group does not or in a tax setting where the comparison group is in a tax bracket not affected by a tax change that affects the worker under question, as with a group of former classmates, for instance.
worker utility that do not consistently explain all three findings include a reference dependent model that takes last period’s average peer earnings as the reference point, predicting a labor supply function for deflated workers that is weakly greater than the labor supplied by inflated workers for all wages, and, an alternative reference dependent model that treats rank rather than earnings as the gain-loss commodity and takes last period’s rank as the reference point, yielding predictions that are consistent with each of the findings at some wage level but never simultaneously consistent with all three of them using the same draws from the wage distribution.

The gender differences observed in the data can also be explained through the treated worker labor supply in (3). Recall, for men the pattern of observed worker productivity is inflated worker labor supply > deflated worker labor supply = control worker labor supply. This suggest that $\lambda > \lambda_1 = 0$ for men. That is, while men get extra utility from improving their relative standing when facing low earners, who make them likely to go to the top of the earnings distribution with their extra effort, they experience no real marginal status gain from extra effort when compared to a group of high earners, presumably because this peer group makes achieving high relative position unlikely and men value improved social status only when it allows them to approach the position of top dog. Given the shift in expectations implied by the different treatments, a $P_L$ approaching zero for deflated workers combined with $\lambda_1 = 0$ predicts the near convergence of output for deflated treatment and control workers\(^\text{18}\) while inflated treatment workers outperform both. For women, the pattern of observed worker productivity is inflated worker labor supply = deflated worker labor supply > control worker labor supply. This implies women do not have a non-linearity in $s(c, \bar{c})$ and that $\lambda = \lambda_1 > 0$, i.e., women are not as singularly concerned as men with achieving status at the top.

### 1.6 Policy Implications

Caution should be exercised in applying this work’s findings to settings far afield from the one studied here as there are many dimensions of variation between the present labor market and others, and, it is not clear how important some of these are for the extension of the results. Such uncertainty invites further work. However, in some ways at least, the current field experimental setting makes for a weak test of the importance of relative earnings considerations. For instance, existing literature suggests that competitive preferences can be more pronounced when workers share identifiable characteristics with their co-workers (e.g. when women compete against same-gendered peers, Gneezy et al., 2003), while in the present scenario the peer groups are made up of anonymous co-workers with no known identifiers in common with the worker. Similarly, it seems reasonable to expect that concern with relative earnings may be more pronounced when workers personally know the members

\(^{18}\)For $P_L = 0$ the two converge and for $P_L$ approaching zero we get a slightly larger level of effort supplied by the deflated workers, which is consistent with the data that finds a larger, but statistically insignificant, level of output for deflated workers than that for control workers.
of their comparison group or work in close physical proximity to them - neither of which are the case in our setting with anonymous, geographically dispersed peer groups.\textsuperscript{19} Thus, there is some reason to believe that in other workplace settings the reported effects may be more pronounced. In this section we discuss the possible implications of our findings if indeed future work confirms their generality.

**Implications for Tax Policy**

The first and most obvious implication of this work for governments is that tax revenue can potentially be increased significantly without a corresponding increase in tax rates via policies that publicize taxpayer earnings. Governments do already disclose such information and have historically done so. In post-war Japan, for instance, laws mandated the public release of any taxpayer’s tax returns, and this regime was kept in place for high-income taxpayers until the mid-2000s (Hasegawa et al., 2010). Public disclosure of income tax information existed in the United States well before this, dating back to the first income tax imposed during the Civil War, and, then again, in the 1920s and 1930s with the introduction of the modern income tax (Lenter et al., 2003). Currently, Norway, Sweden, and Finland all require public disclosure of taxable incomes (Rege and Solli, 2012), and, many state governments in the United States require a more limited form of public disclosure for the subset of the population employed by the state, as studied by Card et al., 2012. The advent of the internet and online taxable earnings databases hosted by newspapers truly makes these earnings disclosures public and available to almost anyone. These policies have usually been advocated on the grounds of transparency or to deter tax evasion (see the study by Hasegawa et al. for a challenge to this rationale). Here, we find there may indeed be an additional rationale for such disclosure policies, namely, that disclosure of peer earnings information can stimulate increased worker productivity at a given net of tax wage. Thus, without raising marginal rates, tax revenue on labor income could potentially be increased by sharing peer earnings information with workers.

On the other hand, this work also suggests the possible limitations for tax policy purposes of a peer earnings disclosure policy. Many optimal tax formulas, from the formula for the revenue maximizing linear tax rate to the formulas for both linear and non-linear optimal labor income tax rates, take some form of average elasticity of earnings supply with respect to the net of tax rate as a key sufficient statistic. For instance, in the simple case of a government attempting to maximize revenue from labor income with a linear tax rate, the optimal tax rate $\tau^* = \frac{1}{1+\epsilon}$ is decreasing in the uncompensated elasticity of aggregate pre-tax earnings with respect to the net of tax rate ($\epsilon$). For a given gross wage, this can be shown to be reducible to the elasticity of labor supply with respect to the net of tax wage, which is measured in this study. Given our finding that this elasticity is unchanged with the introduction of peer earnings information, the possibility that workers may engage in a tournament for relative

\textsuperscript{19}See Bandiera et al. (2010) for evidence of how matching workers to co-workers with whom they have social ties can increase total productivity in a different piece rate employment context with no explicit relative earnings information revelation.
earnings one-upmanship that incentivizes effort in such a way that workers are less concerned with wage increases appears unlikely. Thus, this work suggests the efficiency cost of taxation can not be altered directly by relative earnings information disclosure policies - though since more revenue at a given tax rate could be raised there is a potential indirect improvement in the efficiency cost of taxation from such policies if governments are inclined to lower the tax rate because revenue goals are easier to meet thanks to the increased productivity brought about by the policy. It should be noted, however, that very high income earners - who have less use for the consumption value of additional earnings and who, more than others, may be thought to earn simply to compete for status - might behave differently than the workforce under study here, which has an average reported income below the national median. As such, we are somewhat hesitant to completely rule out the possibility that relative earnings information could indeed affect labor supply elasticity among high earners, and, thus, affect revenue-maximizing and optimal marginal rates.

It should also be noted that the null result regarding the usefulness of relative earnings revelation policies in lowering the efficiency cost of taxation does not mean preferences over relative earnings are irrelevant for optimal taxation calculations. Status concerns may affect optimal income tax rates in two ways, as determined by the two effects these concerns may have on a worker. First, there is the direct effect on worker utility whereby holding a worker’s earnings-leisure decision constant, utility is directly lowered (in the case of competitive preferences) by increases in peer earnings. Second, preferences over relative income may also affect labor supply decisions as workers jockey for position. Both of these effects can be represented in the optimal tax formulas presented by Saez (2001). For instance, the optimal linear tax rate takes the form

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + \epsilon}$$

where $\bar{g}$ is the normalized social marginal welfare weight averaged across individuals and weighted by individual pre-tax earnings, and $\epsilon$ is again the uncompensated elasticity of aggregate pre-tax earnings with respect to the net of tax rate. The direct effect, which would account for the negative externality imposed by higher peer earnings that is postulated by Frank (1985) among others, can be corrected and accounted for by appropriate selection of the normalized marginal social welfare weight ($g_i$) for individual $i$ in the optimal tax formula (for those whose income increases lower the utility of others, ($g_i$) can be reduced). There is still, thus, a role for status considerations in the formation of optimal tax rates. What we show here, however, is that the indirect effect of status considerations through labor supply decisions manifest themselves in a way (as an increase in labor supply level and not changes to labor supply elasticity, or, the efficiency cost of taxation) that does not matter for setting the optimal rates.

---

20Such behavior would obviously have to come about from preferences other than those presented in Section V, offered to explain the findings in this experiment including the lack of any difference in elasticity among those in the treatment and control groups.

21See Piketty and Saez, 2013.
Finally, the peer earnings disclosure presented here is considerably more narrow than the all-access disclosure systems currently enabled by government policy and newspaper databases. Clearly, workers at the same firm (even distant and anonymous ones) seem to form a peer group that motivates competitive preferences, as demonstrated here, but it is unclear how worker effort would respond when peer group selection is chosen by workers rather than exogenously imposed. While governments could easily control access to peer earnings information in a fashion similar to that outlined in this field experiment, we imagine firms would have an easier time imitating the peer disclosure policies presented here.

**Implications for Firms**

The positive implications of this work for firm personnel management echo those outlined above for governments: worker productivity could be increased significantly through implementation of a policy disclosing co-worker earnings, or, performance. Such an incentive scheme will obviously be attractive to firms in comparison to the costly alternative of additional monetary incentivization (in the present study, the relative earnings information treatment increased output by approximately as much as a doubling of the average wage). Moreover, more so than governments, firms have the ability to manipulate who the comparison group of workers is with few constraints. This means a firm could choose to compare most workers to low-performing peers (perhaps from another site of operation) simulating our inflated rank treatment, generating an even higher increase in productivity than the 10% jump observed here across all treated workers. Such a selection of peer groups would have the additional advantage of avoiding obvious welfare costs to workers from the relative earnings disclosure. Card et al. (2012) find that while those who learn that they are in the lower half of the co-worker earnings distribution report lower job satisfaction, those who learn they are in the upper half of the distribution report no decrease in job satisfaction. Thus, by manipulating the peer groups to inflate worker relative performance, firms can both additionally boost worker output and minimize potential utility loss from the peer disclosure. While some creativity might be needed to implement a functional and sustainable policy of this sort so that workers do not realize that they and their office mates all seem to be ranking high, it seems in principle doable (for instance by anonymizing the peer groups and by drawing the comparison group of workers from workers in a different site of the firm’s operation). The results further suggest that for workforces made up primarily of women, workers will respond positively relative to a status quo of no information whether the ranking information is good or bad. But the dynamics of this result need to be studied further to ascertain whether those workers who find themselves continually near the bottom of the pack will continue to increase their effort. An additional related question is whether the non-linearity observed in male workers’ response to high or low rank will persist when relative earnings information is publicly known by one’s office mates, for instance, and not

---

22Recall that given the piece rate compensation scheme in this setting earnings are isomorphic with productivity
anonymized. In this case those believing themselves to be of low relative standing may want to work especially hard to avoid the public shame of being so far behind (suggesting that in such a setting $\lambda$ may not be greater than $\lambda^1$ for men).

While worker productivity is a first order consideration for firms, there may be other potential effects from a peer earnings disclosure regime that also matter. For instance, Card et al. (2013) also find that workers who compare unfavorably to their co-workers in the earnings distribution report an increase in their reported intention to look for another job. If disclosure of peer earnings leads some workers to quit then there may be additional costs to firms from searching for and retraining replacement workers. This downside can be moderated, as above, by careful selection of peer comparison groups to simulate the inflated rank treatment groups (since the Card et al. study also demonstrates no increase in intention to look for another job for those who compare favorably to their peer group). Firms may also be concerned that worker earnings disclosure may lead to changes in wage bargaining with workers and increased peer compensation demands that could be costly to the firm (see Strauss, 1955, for one such example, and Gartenberg and Wulf, 2013, for related discussion of the effects of peer earnings disclosure on wages). These represent just one type of “social comparison cost” described by Nickerson and Zenger (2008) which may motivate managerial diseconomies of scale and scope, and, could potentially be exacerbated by emphasizing social comparisons via relative earnings disclosure. Such considerations again suggest the importance of studying the effects of relative earnings disclosure over the long-term, so, as in this case, to ascertain the net result of such a policy.

1.7 Conclusion

To the author’s knowledge, this is the first work using an experimental identification strategy to show that worker effort in the field is positively affected by exposure to relative earnings information, and, this effect is found to lift worker effective effort by 10% of average output. This result can be interpreted as demonstrating intrinsic competitive preferences and concern for relative standing as there was no direct financial reward from doing better than one’s peers and indirect future employment considerations were unlikely given that workers knew at the outset that the employment was one-off. The findings can also be interpreted as the first evidence of peer effects with absent peers, suggesting that having peers present in the workplace may be beneficial not because their physical presence is important in and of itself but because it allows a worker to observe a rough measure of peer performance or earnings and this inspires a competition for status in the worker. This is also the first work to use the random assignment of high and low performing peer groups to exogenously change worker rank and as a means to avoid the endogeneity problems that beset analysis of whether it is those who learn they performed favorably or unfavorably to their comparison group who supply more effort. The results suggest it is those workers who learn they ranked high relative to their peers in the previous round of work that especially exert more effort in the subsequent round of work, thus, driving the observed positive labor effect of relative earning
information. This heterogeneity in the effect of relative earnings information is caused by men who learn they occupy a low relative standing failing to exert significantly more effort than those who are not exposed to any relative earnings information. Among women, there is no significant difference in performance between those who learned they ranked high and those who learned they ranked low relative to their peers, and, both work significantly more than those not exposed to relative earnings information. Additionally, this work is able to test if relative earnings information exposure affects comparative statics results, in particular, labor supply elasticity with respect to the net of tax wage, shedding light on the limits of peer earnings disclosure policies as an instrument for optimal tax policy and the possibility of using such policies to minimize the efficiency cost of taxation. A model that incorporates worker concern for status can explain all of these findings. The implication of this work is that governments can potentially use relative earnings information to grow the tax base - but not to affect optimal labor income tax rates - and that firms can generate significant productivity boosts simply by providing workers with information about the earnings of their peers.
1.8 References


CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY


CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:  
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY


### Table 1.1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.561</td>
<td>0.497</td>
<td>0</td>
<td>1</td>
<td>547</td>
</tr>
<tr>
<td>Age</td>
<td>30.42</td>
<td>9.673</td>
<td>18</td>
<td>66</td>
<td>547</td>
</tr>
<tr>
<td>White</td>
<td>0.778</td>
<td>0.416</td>
<td>0</td>
<td>1</td>
<td>546</td>
</tr>
<tr>
<td>Income</td>
<td>29460</td>
<td>23601</td>
<td>0</td>
<td>125000</td>
<td>549</td>
</tr>
</tbody>
</table>

Notes: The table reports the summary demographics for the workers analyzed in this paper who provided answers to the demographic survey questions. This sample includes those workers who engaged in work in both rounds of work (as measured by production of at least one correct bibliographic journal entry) and excludes those observations that come from the same IP address (to avoid possible repeat users) and observations that report discernibly inaccurate demographic answers (2 observations that report an age of 0 plus another 2 outlier observations that report their income as over $150,000. Including these outlier observations does not significantly change any of the main results).
Table 1.2: Test for Balanced Treatment and Control Groups

<table>
<thead>
<tr>
<th></th>
<th>(1) Mean of control group</th>
<th>(2) Mean of treatment group</th>
<th>(3) Rank-Sum Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent male</td>
<td>55.7</td>
<td>55.4</td>
<td>z = 0.094</td>
</tr>
<tr>
<td></td>
<td>N=339</td>
<td>N=343</td>
<td>p = 0.9250</td>
</tr>
<tr>
<td>Percent white</td>
<td>77.4</td>
<td>77.4</td>
<td>z = -0.021</td>
</tr>
<tr>
<td></td>
<td>N=340</td>
<td>N=341</td>
<td>p = 0.9835</td>
</tr>
<tr>
<td>Annual income</td>
<td>28.9</td>
<td>31.1 (28.9)</td>
<td>z = 0.107 (0.241)</td>
</tr>
<tr>
<td>(in $1,000s)</td>
<td>N=336</td>
<td>N=338 (340)</td>
<td>p = 0.9144 (0.8096)</td>
</tr>
<tr>
<td>Mean age</td>
<td>30.6</td>
<td>31.2</td>
<td>z = -0.537</td>
</tr>
<tr>
<td></td>
<td>N=340</td>
<td>N=343</td>
<td>p = 0.5913</td>
</tr>
</tbody>
</table>

Notes: The table reports the mean values of demographic variables for the treatment (those exposed to relative earnings information) and control groups in columns 2 and 1, respectively. In Column 3, the z statistics and p-values are reported for a Wilcoxon-Mann-Whitney test comparing the underlying distributions of each variable for the control and treatment group. Results exclude 2 observations that report an age of 0. The value in parenthesis are the results excluding 2 outlier observations with income over $150,000.
Table 1.3: Test of Differential Attrition

<table>
<thead>
<tr>
<th>Dependent Variable: Attrition Status</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>-0.0026</td>
<td>-0.0139</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Deflated Rank</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( N )</td>
<td>694</td>
<td>343</td>
</tr>
</tbody>
</table>

Notes: Column 1 shows there is no statistically significant difference in the likelihood of dropping out (leaving the work website permanently) following exposure to information about only one’s own earnings in the previous round of work (those in the control group) rather than information about one’s own earnings and that of a peer group (those in the treatment group). Specifically, in Column 1 an indicator for attriting at the beginning of the second work period is regressed on an indicator for relative earnings information treatment status. In Column 2, this same outcome is regressed on an indicator for the type of relative earnings information treatment received (which takes a value of 1 for those assigned to a high performing comparison group and, thus, a lower average deflated rank) to test for differential attrition among the different types of treatment (using only treated workers in the specification). Huber robust standard errors are reported in parentheses. *\( p < 0.05 \), **\( p < 0.01 \), ***\( p < 0.001 \)
Table 1.4: Relative Earnings Information and Worker Output

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base Sample</td>
<td>Base Sample</td>
<td>Base Sample</td>
<td>MTurk ID</td>
<td>Base Sample</td>
</tr>
<tr>
<td>Wage</td>
<td>37.11***</td>
<td>39.62***</td>
<td>38.84***</td>
<td>39.32***</td>
<td>39.32***</td>
</tr>
<tr>
<td>Relative Earnings Information Trt.</td>
<td>2.433***</td>
<td>2.836***</td>
<td>2.819***</td>
<td>2.694***</td>
<td>0.176***</td>
</tr>
<tr>
<td></td>
<td>(0.906)</td>
<td>(0.898)</td>
<td>(0.902)</td>
<td>(0.910)</td>
<td>(0.0503)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.199***</td>
<td>-0.193***</td>
<td>-0.187***</td>
<td>-0.00612**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0411)</td>
<td>(0.0417)</td>
<td>(0.0425)</td>
<td>(0.00243)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3.357***</td>
<td>3.276***</td>
<td>3.444***</td>
<td>0.118**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.905)</td>
<td>(0.909)</td>
<td>(0.912)</td>
<td>(0.0500)</td>
<td></td>
</tr>
<tr>
<td>Log Income</td>
<td>0.00143</td>
<td>-0.00969</td>
<td>-0.0122</td>
<td>-0.00493</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.144)</td>
<td>(0.147)</td>
<td>(0.00767)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>-0.431</td>
<td>-0.410</td>
<td>-0.464</td>
<td>-0.0296</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.130)</td>
<td>(1.136)</td>
<td>(1.149)</td>
<td>(0.0615)</td>
<td></td>
</tr>
<tr>
<td>Log wage</td>
<td>0.166***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0364)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.040</td>
<td>0.101</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>N</td>
<td>561</td>
<td>542</td>
<td>542</td>
<td>535</td>
<td>542</td>
</tr>
</tbody>
</table>

Notes: The dependent variable in Columns 1 through 4 is the total number of correct articles entered in round 2 of work, the round of work following the revelation of one's own and (for treated workers) co-worker earnings information. The dependent variable in Column 5 is the log of this output. Relative Earnings Information Treatment is an indicator variable that takes the value of 1 if workers were randomly assigned to receive information about peer earnings in addition to the information about their own earnings that all workers receive. The Wage variable ranges from $0.02 to $0.16 per correct article entry, and the Log wage variable is the log of 100*Wage. Column 4 excludes those workers who failed to provide a valid MTurk ID. Huber robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Table 1.5: Deflated and Inflated Rank Treatments Exogenously Affect Rank

<table>
<thead>
<tr>
<th></th>
<th>(1) Inflated Rank</th>
<th>(2) Deflated Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Place</td>
<td>65%</td>
<td>11%</td>
</tr>
<tr>
<td>Second Place</td>
<td>6%</td>
<td>13%</td>
</tr>
<tr>
<td>Third Place</td>
<td>11%</td>
<td>12%</td>
</tr>
<tr>
<td>Fourth Place</td>
<td>8%</td>
<td>4%</td>
</tr>
<tr>
<td>Last Place</td>
<td>10%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Notes: The table reports the effect on own rank of assignment to the inflated rank treatment (being presented with the round 1 earnings of a peer comparison group where the co-worker earnings are drawn from low-earning peers) and the effect of assignment to the deflated rank treatment (i.e. being presented with the round 1 earnings of a peer comparison group where the co-worker earnings are drawn from high-earning peers). “First Place” indicates the worker learned prior to round 2 of work that s/he had finished at the top of the earnings distribution relative to a comparison group of co-workers, and similarly for “Second Place” and so on.
Table 1.6: High Rank Revelation and Worker Output

<table>
<thead>
<tr>
<th></th>
<th>(1) Base Sample</th>
<th>(2) Base Sample</th>
<th>(3) MTurk ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage</td>
<td>37.30***</td>
<td>39.91***</td>
<td>40.59***</td>
</tr>
<tr>
<td></td>
<td>(8.869)</td>
<td>(8.714)</td>
<td>(8.775)</td>
</tr>
<tr>
<td>Inflated Rank</td>
<td>3.102***</td>
<td>3.619***</td>
<td>3.459***</td>
</tr>
<tr>
<td></td>
<td>(1.030)</td>
<td>(0.994)</td>
<td>(0.995)</td>
</tr>
<tr>
<td>Deflated rank</td>
<td>1.817</td>
<td>2.122*</td>
<td>2.047*</td>
</tr>
<tr>
<td></td>
<td>(1.112)</td>
<td>(1.122)</td>
<td>(1.132)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.198***</td>
<td>-0.192***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0411)</td>
<td>(0.0418)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3.373***</td>
<td>3.534***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.905)</td>
<td>(0.909)</td>
<td></td>
</tr>
<tr>
<td>Log Income</td>
<td>-0.00780</td>
<td>-0.00607</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.145)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>-0.331</td>
<td>-0.392</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.138)</td>
<td>(1.150)</td>
<td></td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.040</td>
<td>0.102</td>
<td>0.102</td>
</tr>
<tr>
<td>$N$</td>
<td>561</td>
<td>542</td>
<td>535</td>
</tr>
</tbody>
</table>

Notes: Huber robust standard errors are reported in parentheses. The dependent variable in Columns 1 through 3 is the total number of correct articles entered in round 2 of work, the round of work following the revelation of one’s own and (for treated workers) co-worker earnings information. “Inflated Rank” indicates a dummy variable for treated workers who were randomly assigned to low-performing peer groups, resulting in an inflated rank for themselves, while “Deflated Rank” indicates those treated workers randomly assigned to a high-performing peer group, resulting in a deflated rank for themselves. Column 3 excludes those workers who failed to provide a valid MTurk ID. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Table 1.7: No Differential Elasticity Among Treated and Control

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treated</td>
<td>Control</td>
</tr>
<tr>
<td>Log Wage</td>
<td>0.166***</td>
<td>0.168***</td>
</tr>
<tr>
<td></td>
<td>(0.0448)</td>
<td>(0.0588)</td>
</tr>
<tr>
<td>Male</td>
<td>0.0598</td>
<td>0.181**</td>
</tr>
<tr>
<td></td>
<td>(0.0545)</td>
<td>(0.0853)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.00544**</td>
<td>-0.00748*</td>
</tr>
<tr>
<td></td>
<td>(0.00263)</td>
<td>(0.00438)</td>
</tr>
<tr>
<td>Log Income</td>
<td>-0.00348</td>
<td>-0.00762</td>
</tr>
<tr>
<td></td>
<td>(0.00896)</td>
<td>(0.0131)</td>
</tr>
<tr>
<td>White</td>
<td>0.00949</td>
<td>-0.0686</td>
</tr>
<tr>
<td></td>
<td>(0.0736)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.062</td>
<td>0.042</td>
</tr>
<tr>
<td>$N$</td>
<td>280</td>
<td>262</td>
</tr>
</tbody>
</table>

Notes: As in Column 5 of Table 4, the dependent variable is the log of round 2 output (see Table 4 for further description). Huber robust standard errors in parentheses. $^*$ $p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$
Table 1.8: Estimates of the Effects of the Placebo Treatment

<table>
<thead>
<tr>
<th></th>
<th>(1) Placebo</th>
<th>(2) Treatment</th>
<th>(3) p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>0.179</td>
<td>2.836***</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(1.191)</td>
<td>(0.898)</td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>40.32***</td>
<td>39.62***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.28)</td>
<td>(8.718)</td>
<td></td>
</tr>
<tr>
<td>Demographic Controls</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.105</td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>456</td>
<td>542</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The dependent variable in Columns 1 and 2 is the total number of correct articles entered in round 2 of work. Column 1 reports the results of placebo test where workers in the “treatment” group are the same as the control workers other than that they are also asked the second two earnings expectation questions about peer earnings presented in Figure 4. Column 2 repeats the results from Column 2 of Table 4, where the “treatment” workers are the relative earnings information treatment workers who received the same earnings expectation questions as the placebo treatment but also receive information about the earnings of a group of co-workers. Column 3 reports the p-value for the hypothesis that the placebo and treatment effect are equal. Huber robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Table 1.9: No Gender Differences Among Deflated Rank Workers

<table>
<thead>
<tr>
<th></th>
<th>(1) Inflated Rank</th>
<th>(2) Deflated Rank</th>
<th>(3) Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.169***</td>
<td>-0.0404</td>
<td>0.181**</td>
</tr>
<tr>
<td></td>
<td>(0.0630)</td>
<td>(0.0866)</td>
<td>(0.0853)</td>
</tr>
<tr>
<td>Log Wage</td>
<td>0.134***</td>
<td>0.212***</td>
<td>0.168***</td>
</tr>
<tr>
<td></td>
<td>(0.0444)</td>
<td>(0.0810)</td>
<td>(0.0588)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.00800***</td>
<td>-0.00350</td>
<td>-0.00748*</td>
</tr>
<tr>
<td></td>
<td>(0.00290)</td>
<td>(0.00433)</td>
<td>(0.00438)</td>
</tr>
<tr>
<td>Log Income</td>
<td>-0.00108</td>
<td>-0.00700</td>
<td>-0.00762</td>
</tr>
<tr>
<td></td>
<td>(0.00900)</td>
<td>(0.0141)</td>
<td>(0.0131)</td>
</tr>
<tr>
<td>White</td>
<td>0.0602</td>
<td>-0.0263</td>
<td>-0.0686</td>
</tr>
<tr>
<td></td>
<td>(0.0715)</td>
<td>(0.134)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>adj. (R^2)</td>
<td>0.134</td>
<td>0.047</td>
<td>0.042</td>
</tr>
<tr>
<td>(N)</td>
<td>133</td>
<td>147</td>
<td>262</td>
</tr>
</tbody>
</table>

Notes: Columns 1 and 2 decompose Column 1 in Table 7 into those treated with co-worker earnings information drawn from a low earning comparison group (Column 1) and those treated with co-worker earnings information drawn from a high earning comparison group (Column 2). Column 3 repeats Column 2 of Table 7 which reports results for control workers. Huber robust standard errors are reported in parentheses. * \(p < 0.10\), ** \(p < 0.05\), *** \(p < 0.01\)
Table 1.10: Average Output By Gender and Treatment Status

<table>
<thead>
<tr>
<th></th>
<th>Output is Number of Articles Correctly Inputed After Earnings Revelation</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Inflated Rank Treatment</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>31</td>
<td>28.2</td>
<td>27.8</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>26</td>
<td>26.5</td>
<td>22.7</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table reports the mean values of round 2 output (the number of correctly entered journal articles) by gender and treatment status. Inflated (Deflated) Rank indicates those treated workers who were randomly assigned to low (high) performing peer groups, resulting in an inflated (deflated) rank for themselves when shown information prior to the start of round 2 about their placement in the earnings distribution in the previous round of work. The earnings revelation for control workers includes only information about their own earnings in the previous round of work. Among men, the difference in output between the inflated and deflated rank treatments is significant and the difference between deflated rank and control worker output is not. Among women, both deflated and inflated rank workers perform significantly better than control workers but not significantly different than one another.
Table 1.11: The Labor Effect of High Rank for Men

Panel A: IV-2SLS (Dependent Variable: Worker Output After Relative Earnings Information)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>First Place</td>
<td>6.640***</td>
<td>-2.403</td>
</tr>
<tr>
<td></td>
<td>(2.507)</td>
<td>(3.003)</td>
</tr>
<tr>
<td>Wage and Demographic Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Panel B: First Stage (Dependent Variable: Indicator for When Told in First Place)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inflated Rank Treatment</td>
<td>Inflated Rank Treatment</td>
</tr>
<tr>
<td></td>
<td>0.576***</td>
<td>0.532***</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>Wage and Demographic Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: The table reports the 2SLS-IV estimates, for both men and women, of the effect of learning that one is in first place in the earnings distribution rather than any other place in the distribution. The sample in the table is of treated workers only. The Inflated Rank Treatment variable is an indicator that takes the value of 1 if a worker was randomly assigned to be compared to a low performing peer group (rather than a high performing one), resulting in an inflated rank for the worker when shown information prior to the start of round 2 about their placement in the earnings distribution in the previous round of work. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Entrepreneurship: Productive, Unproductive, and Destructive


Abstract

The basic hypothesis is that, while the total supply of entrepreneurs varies among societies, the productive contribution of the society's entrepreneurial activities varies much more because of their allocation between productive activities such as innovation and largely unproductive activities such as rent seeking or organized crime. This allocation is heavily influenced by the relative payoffs society offers to such activities. This implies that policy can influence the allocation of entrepreneurship more effectively than it can influence its supply. Historical evidence from ancient Rome, early China, and the Middle Ages and Renaissance in Europe is used to investigate the hypotheses.
FIGURE 2

Bibliographic data entry

We have a bunch of journal articles whose bibliographic information (article title, author, journal, year of publication, etc.) needs to be entered into our database. There are also a few worker survey questions we'd like you to answer to help us understand our workforce better. The worker survey questions should only take about a minute to complete and upon the end of all survey and data entry work you will be paid 1 dollar for the honest completion of the worker survey questions. For the data entry work, you will be given an extended time to work on as many articles as you can, and you will be paid a guaranteed bonus for each journal article you correctly enter. Most of our work is bonus work, and we are committed to always paying our workers promptly and as promised. Including the bonus pay, total take home pay for our workers is usually several times more than the posted flat rate of 1 dollar, and, in some cases up to ten times more. All together, the work requires no less than 45 minutes and at most an hour. An example of the data entry work can be found below.

<table>
<thead>
<tr>
<th>Title</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Author-1</td>
<td></td>
</tr>
<tr>
<td>Author-2</td>
<td></td>
</tr>
<tr>
<td>Author-3</td>
<td></td>
</tr>
<tr>
<td>Author-4</td>
<td></td>
</tr>
<tr>
<td>Author-5</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td></td>
</tr>
</tbody>
</table>
Instructions

READ THE INSTRUCTIONS CAREFULLY TO UNDERSTAND HOW TO GET PAID FOR YOUR WORK.

We have a bunch of journal articles whose bibliographic information (article title, author, journal, year of publication, etc.) needs to be entered into our database. There are also a few worker survey questions we'd like you to answer to help us understand our workforce better. The worker survey questions shouldn't take more than a minute and you will be paid a flat fee of 1 dollar simply for answering them honestly. For the data entry, you will be paid a 8 cent bonus for each article whose bibliographic information you enter correctly (any withholdings consistent with state law and required by Amazon will have already been applied, so the entire 8 cent bonus per correct article is yours, like an after-tax wage). You will be presented with a series of screens that look like this:

Scroll down for continued instructions
CHAPTER 1. USING STATUS TO MOTIVATE WORKERS:
A FIELD EXPERIMENT ON RELATIVE EARNINGS AND LABOR SUPPLY

FIGURE 3b

The Relationship between Corporate Entrepreneurship and Strategic Management


Abstract

Drawing on a sample of 283 subsidiaries in three countries, we investigate how headquarters' attention affects subsidiary performance. • Scholars have recently argued that top management's attention is the most critical, scarce and sought-after resource in organizations (Haas and Hansen 2001: Bouquet and Birkinshaw 2008). However, the question how headquarters' attention affects subsidiary performance remains largely unexplored. • Our study shows that subsidiaries which have a high level of strategic choice and receive attention from headquarters perform better than their peers. More specifically, we find that the interactions of subsidiaries' autonomy, inter-unit power and initiatives with attention increase subsidiary performance.

A correct entry consists of entering all requested bibliographic details as they appear for a given article and then hitting the “Submit” button (or pressing the Enter key). You will not be paid for an article if you submit the wrong information in any of the bibliographic fields for that article (though entering the wrong information for one article will not affect your pay for any other articles).

This means that when it comes to author names, for instance, you must enter the name of the author or authors as they appear (including middle initials) in the order they appear (for example, Bruce R. Barringer goes in the Author1 field above since his name comes first, and, Allen C. Bluedorn goes in the Author2 field since it comes second, and so on) without entering any extra commas, periods, or other punctuation in the fields. The same holds true for all other bibliographic fields, so that a correct entry may not attach additional commas, periods, or parentheses to the article or journal title, publication year, journal issue, volume or page number field. If there is only one author, leave the additional author fields blank.

You may enter the bibliographic details either by typing them in or by copying and pasting from the journal information presented. After you click the submit button or hit enter, a new article will be presented for you to enter into the database. You can work on as many articles as you can get through in 20 minutes, working at your own pace.

Scroll down for continued instructions
After 20 minutes of work at a pay rate of 16 cents per correct article, you will be informed of how much total money you earned from this data entry work. You will be asked to answer a few brief questions and you will be given further instructions about a second 20 minute round of work doing the same bibliographic data entry as before with new articles. Your payment per correctly entered article for this later round of work may be different than the first round of work, but it will be determined by our managerial needs and independently from your performance or pay or any other factor in the previous round of work. Following this second round of work you will be asked to give your Amazon Worker ID and then you will be given a unique payment code which will allow your earnings for all the data entry work and the worker survey questions to be transferred to your Amazon Mechanical Turk account. You must answer all worker survey questions and advance through both 20 minute rounds of data entry work to collect your payment, which you will need to collect your payment. You can, of course, do as much or as little work as you want and work at whatever pace you choose during the data entry work periods.
FIGURE 3d(Treated)

Instructions Continued

After 20 minutes of work at a pay rate of 8 cents per correct article, you will be informed of how much total money you earned from this data entry work. At this point, you will also learn how your earnings (and thus, your performance) compare to the earnings of others (a comparison group) participating in the same timed work under the same pay per article as you. You will learn if you earned/perform near the top, the bottom, or somewhere in the middle of this comparison group, but where you rank will not affect your payment since it will be determined, as stated above, by your performance alone (at a rate of 8 cents for each correct article submission).

Following the earnings information, you will be asked to answer a few brief worker survey questions and you will be given further instructions about a second 20 minute round of work doing the same bibliographic data entry as before with new articles. Your payment per correctly entered article for this later round of work may be different than for the first round of work, but it will be determined by our managerial needs and independently from your performance or pay or any other factor in the previous round of work. Following this second round of work you will be asked to give your Amazon Worker ID and then you will be given a unique payment code which will allow your earnings for all the data entry work and the worker survey questions to be transferred to your Amazon Mechanical Turk account. You must answer all worker survey questions and advance through both 20 minute rounds of data entry work to collect your payment code, which you will need to collect your payment. You can, of course, do as much or as little work as you want and work at whatever pace you choose during the data entry work periods.

Consent ->
FIGURE 4

Preliminary Questions

You are about to begin your work, for which you will be paid 12 cents per correct entry. Before you begin, please answer the following:

How much do you expect to earn in the next twenty minutes of work (working under the instructions and pay as just described)?

$ 0.00

How much do you expect a comparison group (those performing the same journal data entry work at the same pay as you) to earn on average in the first twenty minutes of work?

$ 0.00

Compared to this comparison group of fellow MTurk workers performing the same work at the same pay as you, how do you expect to rank in terms of the income you will earn in the next round of work?

〇 〇 〇 〇 〇
Bottom 1/5th Next 1/5th Next 1/5th Next 1/5th Top 1/5th

Start ->
FIGURE 5

Earnings from the last 20 minutes of work

Your earnings below will be transferred to your Mechanical Turk account as a bonus
(to collect, submit the payment code given after the 2nd round of work)

Number of correct entries in the last 20 minutes of work: 16
Earnings for the last 20 minutes of work: $1.92

Next →
FIGURE 6

Earnings from the last 20 minutes of work

Your earnings below will be transferred to your Mechanical Turk account as a bonus
(to collect, submit the payment code given after the 2nd round of work)

Number of correct entries in the last 20 minutes of work: 9
Earnings for the last 20 minutes of work: $0.18

Ranking

<table>
<thead>
<tr>
<th>Rank</th>
<th>ID</th>
<th>Correct Entries</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>24</td>
<td>$0.48</td>
</tr>
<tr>
<td>2</td>
<td>D</td>
<td>22</td>
<td>$0.44</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>21</td>
<td>$0.42</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>15</td>
<td>$0.30</td>
</tr>
<tr>
<td>you 5</td>
<td>A</td>
<td>9</td>
<td>$0.18</td>
</tr>
</tbody>
</table>

Next ->
Chapter 2

Quiet Riot: The Causal Effect of Protest Violence

2.1 Introduction

The decision to engage in violent forms of political protest comes at significant personal and economic cost. Risk of arrest, injury, and even death are high and the associated negative economic consequences are often lasting.\(^1\) Yet, in the face of these steep costs, we regularly observe political actors turning to violence, even in economically developed and relatively democratic societies where other modes of political influence are readily available.

In light of this, one important question is whether political violence actually works. The answer to this question is of intrinsic interest to empirical researchers and critical to our understanding of why individuals engage in costly political violence at all. While theoretical and empirical work has made headway explaining the causes of political violence (Walter, 1997; Collier and Hoeffler 1998, 2004; Fearon and Laitan 2003; Powell, 2003, 2006; Dal Bo and Dal Bo 2004; Sambanis 2004; Garfinkel and Skarpedes 2007; Blattman and Miguel, 2010), existing scholarship regarding its effectiveness, spanning the disciplines of political science, sociology, and economics, offers conflicting empirical results. Early work on labor unrest by Shorter and Tilly (1971) finding a positive relationship between violent protest and protest success has since been complemented with positive findings in the context of anti-war protests (McAdam and Su, 2002) and urban rioting and the expansion of the welfare state (Colby, 1975; Jennings, 1979; Hicks and Swank, 1983; Iris, 1983; Fording, 1997, 2001).\(^2\) Chenoweth and Stephan’s (2009; 2011) work studying global resistance movements is among

\(^1\)For instance, Collins and Margo (2007) show a persistent and statistically significant depression in long-term black home owner property values in areas hardest hit by "race riots" in the 1960s. See also Collins and Margo (2004) and Abadie and Gardeazabal (2003).

\(^2\)Many of these works on American urban riots of the 1960s derive their conclusions not from a comparison of violent and non-violent protests but by comparing the degree of violence among a set of violent protests. Works explicitly assessing the effects of protest violence relative to non-violent protest will be discussed further in Section II.
the most recent to find the opposite relationship, and is preceded in this negative finding by Franklin (2009) (studying Latin American protest movements), Snyder and Kelly (1976) (studying labor unrest), and Welch (1975) (also studying urban riots but challenging the Piven and Clawford (1971, 1977) social control thesis). Still, other work has found no net effect of protest violence on protest success (Kelly and Snyder, 1980; Frey Dietz and Kalof, 1992, who reanalyze Gamson, 1975) or conflicting results that lend support to both a negative and positive interpretation of the efficacy of protest violence (Button, 1978; Isaac and Kelly, 1981). Related literature focusing on the efficacy of terrorism (Pape, 2003; Abrahms, 2006, 2012; Gould and Klor, 2010) has also found mixed results.

Clearly, when it comes to the question of whether political violence is an effective means of achieving concessions, the existing empirical literature has not produced a consistent answer. This may be due to the lack of clear causal identification in existing work. Specifically, preceding papers on the efficacy of protest violence have not adequately addressed confounding issues of omitted variable bias or the endogeneity of violent forms of protest to prospects of protest success, possibly explaining the wildly different reported estimates and casting doubt on any inferred causal relationship.

This paper is the first to look at the efficacy of protest violence using an identification strategy that allows for credible causal inference. Specifically, we use a set of weather and school holiday measures as instrumental variables for protest violence, a design made possible by disaggregated, micro-level protest event data constituting the universe of reported French protests from 1980-1995 and including an indicator for property destruction. Our results suggest that protest violence lowers the incidence of obtaining a concession. Moreover, this statistically significant negative relationship is more pronounced than revealed by naive regression analysis that does not take into account omitted variable bias or endogeneity. The significant negative estimate is robust to a variety of specifications and remains even when relying only on the more traditional weather instruments. As a result, this paper provides the first empirical evidence that protester violence has a negative causal effect (as distinguished from simple correlation) on the incidence of concession that protesters obtain.

Successful identification of the causal effect of protest violence is hard to achieve in the non-experimental world in which scholars of political violence find ourselves. Correlations found with ordinary least squares regressions on observational data may be biased downward by unobserved negative local economic shocks, for instance, which might both lower the flexibility of firms or governments to make concessions to protesters and lead to an increase in protest violence (if the opportunity cost of engaging in violence has lowered with the worsened economic conditions or if police forces have been cut in austerity policies responding to the downturn). Alternatively, OLS results may be biased upward by unobserved weakness of the protest target, as protesters may be more likely to use violence when there are fewer negative consequences to it from a weaker authority and a weaker authority may be more likely to concede from any protest, violent or not. Similarly, some policies may be perceived as going too far by the public and by elements of the governing coalition, thus, enflaming the public’s passions and independently leading to unreported internal channels of successful elite checks on the controversial policy. In short, any number of additional pathways exist
that might compromise the reliability of causal inference using observational data and the traditional OLS specifications used to date.

If able to conduct a field experiment to avoid the pitfalls inherent in the use of observational data, one might randomly purchase bus tickets for black-bloc anarchists to attend some protests and not others, and then assess whether there are differential success rates at those protests with and without the black-bloc participants. Instead of this notional experimental ideal, however, we must rely on non-experimental data in the real world. As such, researchers interested in causal identification are left with the difficult task of finding a source of naturally occurring, random-like variation in the violence associated with political protest. Instrumental variable methods offer one such solution. Specifically, we seek instrumental variables which are uncorrelated with the potential outcome (exogeneity assumption) but which are correlated with protest violence (relevance) and which have no effect on the protest success except through their effect on violence (exclusion restriction)(Angrist et al., 1996; Woolridge, 2002; Angrist and Pischke, 2009).

We argue that the instrumental variables used here - precipitation, maximum daily temperature, and secondary school holidays - observe these properties and provide us with random-like variation in protest violence that allows us to test the efficacy of political violence. Rainfall and temperature on the day of a protest are thought to affect the incidence of violence at the protest through the relationship between these variables and physiological or psychological states that incline one to violence. In the case of temperature, the general folk notion of hot temperatures begetting hot tempers has basis in social psychology and criminology literatures. For instance, Rotton and Cohn (2000) find disorderly conduct calls increase (though not monotonically) with temperature. Other work shows aggressive behaviors increasing with temperature (for a linear effect of temperature see Kenrick and MacFarlane, 1984; Reifman, Larrick, and Fein, 1991; and the survey by Anderson et al., 2002; for evidence of an inverted U see Baron, 1972; Bell and Baron, 1976; Rotton and Cohn, 2000a, 2000b). Hsiang, Burke, and Miguel’s (2013) meta-study finds rising temperatures associated with increases in a variety of violent behaviors, such as domestic violence, other violent crime, armed conflict, and rioting. In the case of precipitation, getting wet lowers the local body temperature, causing protesters to conserve energy and lowering the likelihood they will engage in physically taxing, caloric-intensive violent behaviors. This paper follows a series of recent papers that have used rainfall as an instrument. In the most closely related, Collins and Margo (2007) use rainfall as an instrument for riot severity and

\[\text{Rainfall has been used as an instrument in several papers since Miguel et al. (2004) pioneered its use in the study of political violence. In their case, the authors sought to explain the onset, or existence of, civil war, treating political violence as the outcome of interest and using rain to instrument for economic growth in a study of the effect of income shocks on the outbreak and incidence of civil war. Bruckner and Ciccone (2011) similarly use rainfall as an instrument for economic shocks to assess the effect of income shocks on the degree to which a country possesses democratic institutions. Unpublished work by Madestam et al. (2012) makes use of rain as an instrument for protest size, not violence, a potential concern for our identification strategy. However, we do not find a significant effect on size in this work, discounting the possibility that there is an exclusion restriction violation from rain affecting protest success via weather variables, as detailed below.}\]
cite extensive media reports in the U.S. in the 1960s making the link between rainfall and riot severity.

The rational for the school holiday instrument comes from the incapacitation effect of school. Jacob and Lefgren (2003) document that the level of property crime (the measure of violence used here) committed by juveniles in the United States decreases by 14 percent on days when school is in session. In our context, we theorize the composition of protests held on school holidays will change to include an increased share of teenage participants (who are no longer incapacitated by school) and the greater destructive tendency of these younger protest participants will lead to greater incidence of property destruction associated with these protests. Ethnographic studies of soccer hooliganism find evidence of rioters rioting for the thrill of the riot (Buford, 1991; Wilkinson, 2009) and something similar may be going on with French youth, though we are agnostic as to whether this is a more appropriate explanation than criminal burglary motives, or general intemperateness among hormonal teens. While this work constitutes the first time any of our three instruments (or any others) have been used to estimate protest violence’s effect on protest success, rainfall and temperature instruments have been used in other studies on violence. There has been no previous use (to our knowledge) of the school holiday instrument, however (though, as stated, the results are robust to use of only the more traditional instruments as well). We use all three of these instrumental variables to identify plausibly exogenous variation in protest violence, and hope this effort brings a newfound interest to clean identification of causal effects in the study of political violence and civil conflict.

In the section that follows we present a survey of related literature regarding political violence. In Section III we review the data sets and data construction methods. Section IV presents the estimation framework. Section V presents the main results and considers their robustness. In Section VI we briefly review the consistency of candidate modeling assumptions with the results presented here, and Section VII concludes.

2.2 Existing Literature

In recent years economists and political scientists have paid renewed attention to the study of violent civil conflict and political violence. Much of this work has tried to explain the outbreak of civil war (Collier and Hoeffler 1998, 2004; Fearon and Laitan 2003; Sambanis 2004). Contest models have modeled the decision to go to war as a choice between production and predation, with the probability of war increasing as the opportunity cost of fighting decreases (Dal Bo and Dal Bo 2004; Garfinkel and Skarpedes 2007). Models of asymmetry of information (Powell 2002) have predicted outbreaks of civil war when opposing sides fail to understand the true probabilities of success, and commitment problems have been proposed as further explanations of violent breakdowns in bargaining (Walter 1997; Powell 2006). The consequences of violence have also been studied in the context of this recent renaissance in the study of civil war, with attention paid to the effects of war on poverty, growth, human capital, and other macro-level indicators (see Blattman and Miguel 2010 for
a detailed summary).

The focus of this paper departs from this emerging civil conflict literature in two main ways. First, while many recent works study political violence from the vantage of large-scale civil wars, here, we are interested in lower-level civil unrest in the form of popular protest, including demonstrations, occupations, strikes, and other sub-military political protest acts in which protest may be used to exact short term, issue-specific policy concessions rather than to overthrow an entire regime (non-displacement goals) and where opposition to a certain policy may or may not take a violent form. Secondly, in contrast to much of this work, which focuses on providing explanations for the causes of civil conflict, I look at civil unrest’s effects - specifically the effect of protest violence on the achievement of policy concessions that are consistent with protester demands.

Such a research focus has precedent in both the sociological and political science literatures. Classic work in sociology by Shorter and Tilly (1971) finds a positive correlation between violence and historical strike success in France (using primarily bivariate cross-tabular analysis), while Snyder and Kelly (1976) find the opposite with respect to industrial violence in Italy after controlling for other protest characteristics in a linear probability model. In a multivariate re-analysis of Gamson’s (1975) influential work studying the efficacy of a broad group of American social movements, Frey, Dietz, and Kalof (1992) find a weakly significant (only 10% significance level at best) negative effect of violence on the achievement of concessions. Giugni (1998) offers an overview of the mixed empirical evidence regarding the ability of urban rioting to effect policy and lead to protester gains. To take one recent example in this strand of work, McAdam and Su (2002) find that property damage associated with anti-Vietnam war protests increase the ”pro-peace” vote share in Congress while, also, slowing the frequency of Congressional votes about the war.

In political science, Franklin (2009), conducting a multinomial logit analysis, concludes that disruptive but non-violent contentious challenges to Latin American governments were more effective at gaining concessions and that violent challenges were likely to lead to repression. Chenoweth and Stephan (2009, 2011), using a combination of case studies and regression analysis with a detailed international and historical data set, find that violent tactics diminish the chance of a movement’s eventual success. Other work, yielding varied results, has focused on the outcomes of riots without making comparisons directly to non-violent protest (Welch, 1975; Hicks and Swank 1983; Fording 1997, 2001).

Inherent in all of this work is the lack of exogenous variation in violence (no small, and no easily corrected matter), leaving open questions as to whether endogeneity or omitted variable bias are driving any of the above (conflicting) results. To the author’s knowledge, no existing studies on the efficacy of protest violence have used an instrumental variables strategy or other quasi-experimental methods to effectively or convincingly control for the endogeneity or omitted variable bias affecting violence’s occurrence. Chenoweth and Stephan in the notable book version of their work are the first to make an attempt to implement an IV strategy in a related context, but their ability to do so effectively is hindered by the difficult setting they work in, one which lacks obvious candidates for valid instruments applicable to
the geographically and temporally varied protests under study\(^4\) (they focus on maximalist protest campaigns - anti-occupation, secessionist, and regime change movements - that fall outside the scope of the current study). Though not pursuing an IV strategy, two additional papers deserve mention for research designs that share the spirit of this work’s pursuit of random-like variation in the incidence of political violence and use of micro data. One, by economists Eric D. Gould and Esteban F. Klor (2010), looks at the effectiveness of political violence - however, in this instance in the context of terrorism - by using the geographic variation in terrorist attacks as an identification strategy to determine that more violent terrorist acts, up to a point, effectively shift the opinions of the targeted population to a more conciliatory position (they do not make a comparison with non-violent tactics). The other paper, by political scientist Jason Lyall (2009), focuses on the effects of state violence rather than protester violence, positing allegedly indiscriminate artillery shelling as a random-like distribution of violence, finding that the areas affected by the shelling were less likely to host future insurgent attacks. While the identification strategies of these papers hinge on the presumption that Palestinian militants and the Russian military, respectively, choose the location of their targets randomly, and may thus be open to challenge, we admire their attention to causal identification and think they present some of the more interesting initial approximations to date of Susan D. Hyde’s (2010) entreaty to extend field experimental methods to the study of political science (in her case international relations, in ours political violence). In the study of political violence, approximations may be the best we can do, as the notional political violence field experiment raises obvious legal, ethical, and operational hurdles. It is the author’s opinion that the instrumental variables design offered here presents the next-best thing to the randomized trial gold standard, and that this work takes us a large step forward in the pursuit of successful identification of the causal effect of political violence.

2.3 Data and Measurement

Protest Data

All explanatory variables and outcome measures are derived from the European Protest and Coercion Data (EPCD) database, created by political scientist Ronald Francisco.\(^5\) This micro-foundation data set contains a rich set of variables characterizing protests, including date, day, location, protester identifier, protest target, issue under protest, protest size, measures of protest violence, and protest form/action type (e.g. rally, strike, occupation, occupation).

\(^4\)In Table 3.3 of their book, Chenoweth and Stephan report the results of three IV specifications. All specifications rely on an instrument for whether the protest campaign is secessionist. While there may be first stage relevance for this variable, it poses a clear threat to the exclusion restriction (secessionist campaigns are likely to have a direct effect on the incidence of concession and the choice to make a political campaign explicitly secessionist may be endogenous to the chance of success).

\(^5\)Data is available on Ronald Francisco’s website [http://web.ku.edu/~ronfran/data/index.html](http://web.ku.edu/~ronfran/data/index.html). For discussion of the benefits of protest event analysis data sets of this type see Nam (2006) and Koopmans and Rucht (2002).
CHAPTER 2. QUIET RIOT: THE CAUSAL EFFECT OF PROTEST VIOLENCE

hunger strike etc.). The data covers the years 1980-1995 and provides a daily record of “all reported protests” in France with identifiable date and location that appeared in the newspaper record of some 60-plus French and international newspapers and wire services. Analysis here is limited to the French mainland and excludes observations whose issue of protest/grievance is coded as various forms of “separatism” (the frequent recourse to violence and lack of success among Corsican, Breton, and other separatists of the era suggests that their inclusion would only make OLS estimates more negative, but we leave an IV analysis of the impact of violence in separatist movements, which are likely governed by a different dynamic than the non-maximalist campaigns studied here, for future work).

The dependent variable is derived from a binary measure of whether the abovementioned news sources report that protester demands were met with a concession by the protest’s target (coded as 1) or not (coded as 0). To the author’s knowledge, other existing, multi-issue protest data sets do not allow for the present analysis of the effect of protest violence on protest success since they do not provide the information needed to code outcome measures while at the same time using daily protests with specific locations as the unit of observation, a necessary requirement for the IV strategy pursued here. While the EPCD does do so, this clear advantage comes with certain caveats. For one, the outcome measure almost certainly undercounts protest successes, particularly when protests indirectly affect policy changes and when the policy changes occur in the long run, as the reported concessions that constitute a success in our metric require journalists to some way link the protest and the protest target’s subsequent changed behavior. The more indirect, second-order, or temporally distant any such change is from related protests the less likely news reports are to link policy changes to protesters’ actions. Additionally, while this effect of temporal distance would be true in any event, as explained in the next subsection it is necessarily the case here given the data construction strategy outlined below. Thus, concessions in this paper should be understood to be attributed concessions that end a protest or take place during its span or in a short window following the protest’s end (specifically within seven days, as detailed below), and it is entirely possible that longer-run policy changes may be impacted differently by protest violence than the short-run changes assessed here. Taking account of these longer-run effects would presumably increase the rate of total concession beyond the relatively low rate observed in the data set overall (as seen in Table 1).

Secondly, an outcome coded as a concession should not be interpreted as an absolute victory for protesters, but, rather, as arrival at some point along a spectrum ranging from

---

That news reports are more likely to miss the indirect, second-order, or temporally distant effects of protests on policy highlights how concessions here, as with any data relying on media sources, should be understood to be attributed concessions (with the act of making such attribution always subject to a degree of subjectivity and possible deviation from real happenings, whether from reporters’ judgments (Davenport, 2010) or those of data coders) Even if journalists are not taken as the mediators of information, Lipsky (1968) and Tarrow (1998) point out the inevitable difficulty in attributing the full range of successes that are in part due to popular protest, as the more numerous indirect and non-proximate successes are very difficult to trace. See O’Brien and Li (2005) for a discussion of some of the indirect channels of protest efficacy in a Chinese context.
winning everything demanded to winning something more than nothing. In many cases, the data set describes a newspaper account of a concession that clearly marks a complete capitulation to protester demands by the protest target (“management agrees to demands”, or, workers “end strike after winning pay increase”), but, in others, the extent of the protesters’ win is exceedingly modest (“management agrees to reduce number of layoffs from 270 to 192”). Protests that win concessions should be seen then as an improvement beyond the expected status quo absent the protest.

Protest Event Groups

As indicated, the EPCD database records daily protests. However, protests over successive days are often linked and are reasonably thought of as a unified event (e.g., an ongoing strike, an encampment in the main city square, a building occupation, an organized week of action). Thus, the unit of observation is taken to be such unified, ongoing protests, termed protest event groups. Specifically, these protest event groups consist of protests that share the same protester and target and protest issue and that are separated by no more than a week from another protest in the group. Those protests without a defined issue were dropped on the grounds that we are interested only in protests with a clearly defined claim, and, those protests which coders could not classify into issue categories may be thought to lack an identifiable focus. Data is then collapsed by these groupings, with the protest event groups taken as the unit of analysis. A seven day window between same protester-target-issue protests is used to account for the fact that reporting on ongoing protests can drop off in the dataset and then reappear days later in another news story, and we want to avoid classifying such protests as separate events rather than one protest event with a long span. However, this decision also imposes an upper limit of seven days from the final day of protest by when a concession must materialize to be counted in the analysis, since a protest event group is assigned a concession as the value for its outcome variable if any of the individual daily protest reports composing the protest event group reported a concession. In this sense, the consequences of political violence assessed here truly are short-term consequences for protest success.

With these protest event groups as the unit of observation, a word on the construction of the explanatory variable of interest and other explanatory variables is in order. In the raw data, protest violence is characterized by an indicator variable that takes a value of one in the event of reported property destruction attending the protest. To measure the degree to which such property destruction defines the protest event group, in the collapsed data the Violence variable used in the following specifications is the fraction of newspaper reports on the protest throughout the protest’s span that report the occurrence of property damage, giving us a measure of the the degree to which a protest event group is characterized by

---

For a more detailed summary of the algorithm used to group protests and a description of the manual checks using basic article descriptions of protests that were undertaken to avoid false matches or false splits of protest event groups, contact the author. It is encouraging to note the results reported do not depend on the manual checks and are robust without them.
violence. Additional explanatory variables used in some specifications include Paris, the number of days that protests took place in the capital in the protest event group; Size, an indicator variable \(1[\text{average reported protest size in the protest event group } \geq 1000]\); Election, an indicator variable equal to 1 if part of the protest span is between thirty days prior to the first round of a presidential election and the first round election (or, the second round election, if applicable); and Duration, the number of days from the start to finish of the protest event group.

Table 2.1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concession Made to Protesters</td>
<td>0.021</td>
<td>0.142</td>
<td>2087</td>
</tr>
<tr>
<td>Violence (% violent days of protest)</td>
<td>0.182</td>
<td>0.371</td>
<td>2087</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instruments</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protests with Rain on Some Day</td>
<td>0.51</td>
<td>0.5</td>
<td>2087</td>
</tr>
<tr>
<td>Protests with Avg. Temp. Above 60 and Below 75</td>
<td>0.338</td>
<td>0.473</td>
<td>2087</td>
</tr>
<tr>
<td>% of Days in Protest on School Holiday</td>
<td>0.189</td>
<td>0.38</td>
<td>2087</td>
</tr>
</tbody>
</table>

Note: Violence is measured as the fraction of newspaper reports on a given protest that report the occurrence of property damage.

Weather and School Holiday Data

Data for the instruments is taken from two sources. For exogenous weather variation, precipitation and temperature data comes from the National Climatic Data Center’s (NCDC) Global Surface Summary of Day database.\(^8\) The NCDC data provides daily readings from more than 350 weather stations across France (224 of which provide usable data during our period of study), reporting maximum daily temperature (in degrees Fahrenheit) and total daily precipitation (in inches). Other common precipitation data sets, such as the Global Precipitation Climatology Project, are inadequate for the present analysis as they present monthly rather than daily readings. For school holidays, the moving annual calendar of French secondary school holidays, standardized nationally by educational zone, is taken from French Ministry of Education records.\(^9\) Traditionally, French secondary schools see several week-long holidays at intervals of a month to a month and a half throughout the year, as well as an extended holiday in the later part of July and August, providing temporal variation in the holiday measure. The educational zones rotate from year to year so that

---

\(^8\)Data can be found at [http://www.ncdc.noaa.gov/cgi-bin/res40.pl](http://www.ncdc.noaa.gov/cgi-bin/res40.pl)

educational zone A and educational zone B may be on the same holiday calendar one year, but not the next, providing additional regional variation in the holiday measure as well.

Figure 2.1: Plot of unique protest locations (in blue) and weather stations

Using ArcGIS global imaging software, each individual daily protest is geographically plotted. For the assignment of weather variables, each daily protest is matched with the nearest French weather station containing non-missing data for the day. See Figure 1 for an ArcGIS map of protest and weather station locations. Notice the distribution of unique protest sites and weather station sites is spread across the country with observations spread throughout the French administrative regions. For protests taking place throughout entire administrative regions, departments, or geographic areas, the average maximum daily temperature and average daily precipitation level for weather stations in the relevant area is used for the daily record. Upon collapsing data into protest event groups, the principal precipitation measure used is a dummy variable, Precipitation, that takes the value of 1 if it rained or snowed at least one day of the protest’s duration at one of the protest sites, while the principal temperature measure used Temp is a dummy variable that equals one when the maximum temperature averaged across days and sites of the protest falls in the range of 60 to 75 degrees Fahrenheit. This range is chosen primarily because it gives a strong first stage consistent with theory, but there are also findings in social psychology literature indicating aggressive behavior and disorderly conduct peak in warm temperatures in this range (See
Figure 2, from Rotton and Cohn, 2000a) rather than rising monotonically, suggesting the conventional belief that with rising temperatures come rising tempers holds only up to a point (Baron, 1972; Bell and Baron, 1976; Rotton and Cohn, 2000a, 2000b). Alternative temperature measures, including a linear temperature measure and a dummy for extreme hot or cold temperatures, are tested and arrive at qualitatively similar results for violence as many of the specifications using the Temp variable presented below. Similarly, results are robust to increasing the threshold for the precipitation dummy to the mean of average precipitation across all protests and using this instrument instead.

Figure 2.2: Disorderly Conduct Calls as a Function of Temperature (reprinted from Rotton and Cohn, 2000a)

Once protests are plotted in their administrative department, the moving annual calendar of French secondary school holidays is used to determine if secondary school students at the protest location were on extended holiday on the day of the protest. For protests reported to span over multiple school zones, that protest is considered to take place on a school holiday if some of the protesters protest in an area on holiday (results do not differ significantly when they are assigned the holidays common to the spanning school zones). The Holiday variable then is defined as the proportion of days at protest event group locations during which students are on extended school holidays.
2.4 Estimation Framework

Ordinary Least-Squares Regressions

Table 2 reports ordinary least squares (OLS) regressions of protest success on protest violence and additional protest characteristics. The linear regressions are for the linear probability model

\[
y_i = b + \alpha \text{Violence}_i + X'_i B_1 + \epsilon_i
\]

where \(y_i\) is a binary measure of whether protesters succeeded at winning a concession from the target of the protest (as reported by the press), \(\text{Violence}_i\) is the fraction of newspaper reports on the protest throughout the protest’s span that report the occurrence of property damage, \(X_i\) is a vector of other variables characterizing the protest which are used in some specifications (enumerated in Section III), and \(\epsilon_i\) is a random error term, where these errors are clustered by protest target to account for correlations in probability of success among those protests with the same target. Such correlation is reasonable when successive demands upon a protest target affect the target’s willingness to concede to future demands.

Column (1) reveals a small but significant negative correlation between protest violence and protest success. Going from a protest event group characterized by no violence to one characterized by violence on every day of its duration lowers the chance of achieving a concession by 2%. Controlling for annual time trends in Column (2) leaves this estimate unchanged. Similarly, adding other explanatory variables that have theoretical foundation in the protest and social movement literature has no effect (results for these other variables, including duration, location, size, and election cycle, are reported in more detail in the Appendix).

Overall, the results in Table 2 demonstrate that violent protests are significantly less likely to result in a concession. There are many reasons, however, not to interpret this correlation as a causal relationship, as already indicated in the discussion in Section I. To obtain a causal effect of political violence on protest success we now turn to an instrumental variables strategy.

Instrument Relevance

Our identification strategy makes use of a combination of three instrumental variables for protest violence, which is treated as an endogenous regressor from here on. To ascertain the strength of these instruments, violence is modeled as

\[
\text{Violence}_i = a + \delta \text{Precipitation}_i + \gamma \text{Temp}_i + \mu \text{Holiday}_i + X'_i B_2 + \epsilon_i
\]

where \(\text{Precipitation}, \text{Temp}, \text{and Holiday}\) are defined as at the end of Section III, \(\text{Violence}_i\) and \(X'_i\) are as defined in the previous subsection, and \(\epsilon_i\) is a random error term allowed to be correlated across observations with the same target of protest.
Table 2.2: Protest Violence and Concession to Protesters (OLS Results)

<table>
<thead>
<tr>
<th>Dependent Variable: Incidence of Concession</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>-0.0261***</td>
<td>-0.0247***</td>
<td>-0.0222***</td>
</tr>
<tr>
<td></td>
<td>(0.00755)</td>
<td>(0.00528)</td>
<td>(0.00543)</td>
</tr>
<tr>
<td>Time trends</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Additional Explanatory Variables</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.004</td>
<td>0.009</td>
<td>0.025</td>
</tr>
<tr>
<td>$N$</td>
<td>2087</td>
<td>2087</td>
<td>2087</td>
</tr>
</tbody>
</table>

Note: Standard errors are reported in parentheses, adjusted for protest target level clustering.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first-stage relationship between the instruments and protest violence is significant and robust to the inclusion of additional control variables, as can be seen in columns (1) thru (3) in Table 3. The Precipitation variable is significant at the 1% levels in all specifications (with the t-stat being particularly large) and the Temp and Holiday variables are significant at the 5% significance level or better in all three specifications. The robust F statistic ranges from 13 to 23, suggesting that there is not a weak instrument problem (Stock et al. 2002; Stock and Yogo 2002).\(^{10}\)

Moreover, the signs on the instruments’ coefficients are all in the direction expected by theory: rainfall leads to less violence, warm temperatures lead to more violence, and violence spikes when students are not incapacitated by school. Results consistent with our theory are present in the underlying uncollapsed data as well, and, when the first stage is run with one instrument at a time the coefficients and standard errors are very similar to those in Table 3 (results unreported). Alternative instruments using the underlying weather and holiday data (including a dummy for rainfall above the median and a linear temperature term) also yield similar results.

### 2.5 Main Empirical Results

#### Initial Estimates With Violence as Endogenous Regressor

Baseline instrumental variables results using Precipitation, Temp, and Holiday as instruments for Violence can be seen in Table 4. In the presence of non-iid errors (as is the case here) GMM estimation provides improvements in asymptotic efficiency relative to 2SLS in

\(^{10}\)The tables for weak instrument tests presented in the Stock et al. works are not directly applicable in this context, and to our knowledge there is not a comparable work for gmm estimation with non-iid errors. The usual rules of thumb should be seen as suggestive as a consequence.
over-identified models (Baum, 2006), and Table 4, thus, presents IV-GMM estimates correcting for endogeneity and omitted variable bias in the OLS estimates from Table 2.

Table 2.3: Weather, School Holidays and Protest Violence (First-Stage)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>-0.0538***</td>
<td>-0.0621***</td>
<td>-0.0520***</td>
</tr>
<tr>
<td></td>
<td>(0.0135)</td>
<td>(0.0124)</td>
<td>(0.0148)</td>
</tr>
<tr>
<td>Temp</td>
<td>0.0357**</td>
<td>0.0353**</td>
<td>0.0347**</td>
</tr>
<tr>
<td></td>
<td>(0.0167)</td>
<td>(0.0162)</td>
<td>(0.0151)</td>
</tr>
<tr>
<td>Holiday</td>
<td>0.0754***</td>
<td>0.0709***</td>
<td>0.0549**</td>
</tr>
<tr>
<td></td>
<td>(0.0240)</td>
<td>(0.0240)</td>
<td>(0.0242)</td>
</tr>
<tr>
<td>Time trends</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Additional Explanatory Variables</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.012</td>
<td>0.031</td>
<td>0.065</td>
</tr>
<tr>
<td>Robust F</td>
<td>19.12</td>
<td>23.07</td>
<td>13.30</td>
</tr>
<tr>
<td>$N$</td>
<td>2087</td>
<td>2087</td>
<td>2087</td>
</tr>
</tbody>
</table>

Note: Standard errors are reported in parentheses, adjusted for protest target level clustering.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Protest violence is found to be negatively related to the incidence of concession, with the estimates being highly significant. In Column 1 of Table 4, the point estimate on the violence variable is -0.16 (standard error of 0.05), with better than a 99% confidence level. This means that going from a protest characterized by no property destruction at all to one characterized by property destruction on every day of the protest lowers the probability of concession by almost 20% on average. Since we have instrumented for protest violence we interpret this negative relationship between protest violence and the incidence of concession as a causal one.\textsuperscript{11} In column 2 we add time trends and in column 3 we present the results from adding other explanatory variables (Size, Duration, Paris, and Election). It is common in the social movements and protest violence literature to include other protest characteristics such as these in the regression to assess their relative influence on protest outcomes. The degree to which these additional explanatory variables are actually appropriately considered exogenous control variables varies, as some may be more reasonably considered exogenous than others. In general, we think such variables may not be properly considered control variables and they are included here (with expanded results in the Appendix) primarily for comparability with existing work. In any case, as can be seen, adding these additional explanatory variables as controls does not dramatically change the estimate on Violence.

\textsuperscript{11}Obviously, such inference is conditional on the validity of our instruments, something discussed further later in this section.
CHAPTER 2. QUIET RIOT: THE CAUSAL EFFECT OF PROTEST VIOLENCE  

Table 2.4: Effect of Protest Violence on Concession (IV-GMM Results)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>-0.162***</td>
<td>-0.157***</td>
<td>-0.145**</td>
</tr>
<tr>
<td></td>
<td>(0.0506)</td>
<td>(0.0590)</td>
<td>(0.0621)</td>
</tr>
<tr>
<td>Time trends</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Additional</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Explanatory</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value (from</td>
<td>[0.18]</td>
<td>[0.19]</td>
<td>[0.30]</td>
</tr>
<tr>
<td>chi-squared test)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2087</td>
<td>2087</td>
<td>2087</td>
</tr>
</tbody>
</table>

Note: Standard errors are reported in parentheses, adjusted for protest target level clustering.

* * p < 0.10, ** p < 0.05, *** p < 0.01

The large difference in the size of the IV and OLS estimates suggests that the OLS estimates are afflicted by endogeneity and omitted variable bias, as suspected. Some omitted variable or endogeneity bias stories can be ruled out as dominant, while others are consistent with the results. For instance, unobserved negative economic shocks may be presumed to increase the tendency to resort to violence by protesters (if opportunity costs of violence are lowered or police budgets have been cut under austerity policies, for instance) while also leaving firms or government less likely to reverse layoffs or program cuts. However, this downward bias of OLS results is inconsistent with the more negative coefficient on Violence in the IV results relative to OLS. On the other hand, unobserved weakness or divisions among the governing authority would predict an upward bias on the OLS estimates, as protesters may be more likely to resort to violence and authorities may be more likely to capitulate independent of the violent tactic employed. One can think of other channels of bias that are similarly consistent with the findings.

Considering the Exogeneity Assumption and Exclusion Restriction

The credibility of the IV estimate for violence depends on the degree to which we believe our instruments are exogenous. The argument for the exogeneity of the weather instruments strikes us as obvious. However, the exogeneity of the school holiday instrument, while intuitively plausible, may stand on weaker ground. Though the school holiday calendar is chosen years in advance of time-t protests out of consideration of factors entirely independent of the current events motivating protests and their success at time t, it is possible that those

---

12 The possibility that the IV estimates reveal and correct for attenuation bias on the violence measure in the OLS specifications presumes the existence of classical measurement error in the underlying property damage incidence variable used to construct the Violence variable, something that is not the case given that this variable is dichotomous.

13 All rain dances aside.
engaged in the protest-concession dynamic make their strategic decisions with the school calendar in mind, causing possible violations of the exogeneity assumption. For instance, policy makers, anticipating either differential violence or youth participation in protest, may schedule the most controversial policy changes in accordance with the school calendar. If these controversial policy changes are also those which policy makers are most intent on implementing, yielding the lowest chance of successful protest, then there may indeed be a violation of the assumption that the school holiday calendar is independent of potential outcomes. Alternatively, protest organizers, with advance knowledge of the local school calendar, may shift the day of their protest by a few days to affect the composition of the protest and this decision may be correlated with underlying perceptions of the incidence of concession.

Such concerns are most plausible with respect to education policy changes, an issue of protest most likely to engage students. Column 1 in Panel B of Table 5 includes the first-stage results from an IV regression that excludes protests in which the issue is education. The first-stage results are unchanged from the base sample, and, as can be seen in the corresponding column in Panel A of Table 5, the IV estimate on violence is also virtually unchanged.

Column 2 in Table 5 reports the first-stage and IV estimate without use of the school holiday instrument. Using only rainfall and temperature instruments we continue to find a highly significant (over 99 percent confidence) negative IV estimate of violence on the incidence of a concession (the violence coefficient increases in size to -0.44). These results give us confidence in the credibility of the baseline negative estimates and serve to minimize concern that strategic agenda setting on the part of policy makers or protest planning in accordance with the school holiday calendar is driving the result, though we can not rule out with absolute certainty the possibility that the Holiday instrument may not be completely exogenous.

The fact that there is not comparable advance knowledge of, or certainty about, future weather conditions as there is with the school calendar, makes the exogeneity assumption with respect to rain and temperature compelling and the strategic concerns raised above with respect to the Holiday instrument unlikely to affect our weather instruments. In a pre-social media age, protest organizers need lead time to publicize their protests, and, once a flyer is put up or word is otherwise spread there is no easy way to alert all who have been informed not to attend. Moreover, weather forecasts until very recently had been notably bad. It is also implausible that we are observing reverse causality or a common cause of both rain and the incidence of concession. However, the exclusion restriction must not be violated in other ways, such as by either of the weather variables having an effect on the protest

14We thank Jonah Levy for fruitful discussions about the nature of French youth participation in protest.

15For instance, Silver, 2012, recounts how near the start of our data period weather forecasts three days out were off by six degrees Fahrenheit on average. One study regarding weather predictions in this author’s hometown (Kansas City) found that when TV meteorologists predicted a 100% chance of rain it failed to rain at all one third of the time, making it unlikely that protest organizers can confidently make plans based on the forecast even if able to call off protests at the last minute.
Table 2.5: Weather-only Instruments and Overidentification Tests

Panel A: IV-GMM (Dependent Variable: Incidence of Concession)

<table>
<thead>
<tr>
<th></th>
<th>(1) No Educ.</th>
<th>(2) Weather only</th>
<th>(3) Temp as Exog.</th>
<th>(4) Precp as Exog.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>-0.141**</td>
<td>-0.444***</td>
<td>-0.490***</td>
<td>-0.277**</td>
</tr>
<tr>
<td></td>
<td>(0.0612)</td>
<td>(0.153)</td>
<td>(0.159)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>Temp as Control</td>
<td>0.00695</td>
<td></td>
<td>0.00671</td>
<td>0.0135</td>
</tr>
<tr>
<td>Precipitation</td>
<td></td>
<td></td>
<td></td>
<td>(0.00869)</td>
</tr>
</tbody>
</table>

p-value (chi-squared test) [0.19] [0.29] – –

N 2045 2087 2087 2087

Panel B: First Stage (Dependent Variable: Violence, % violent days of protest)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>-0.0612***</td>
<td>-0.0631***</td>
<td>-0.0631***</td>
<td>-0.0631***</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0126)</td>
<td>(0.0126)</td>
<td>(0.0126)</td>
</tr>
<tr>
<td>Temp</td>
<td>0.0355**</td>
<td>0.0326**</td>
<td>0.0326**</td>
<td>0.0326**</td>
</tr>
<tr>
<td></td>
<td>(0.0164)</td>
<td>(0.0161)</td>
<td>(0.0161)</td>
<td>(0.0161)</td>
</tr>
<tr>
<td>Holiday</td>
<td>0.0720***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0240)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

adj. $R^2$ 0.031 0.026 0.026 0.026

Robust F 24.68 24.67 25.03 4.08

N 2045 2087 2087 2087

Note: Standard errors are reported in parentheses, adjusted for protest target level clustering. Columns 3 and 4 report results from regressions in which Temp and Precipitation are included as exogenous variables, respectively, while using the other variable as the instrument for Violence. All regressions also include time trends.

outcome directly or through a channel other than violence. While such possibilities strike us as unlikely, an overidentification test can be performed to attempt to address the concern. Formally, the overidentification test will reject the validity of the weather instruments if one of the instruments has a direct effect on the incidence of concession, or if they effect incidence of concession through another variable than violence, or, if there is not a constant coefficient $\alpha_i$ in the structural equation (1). The p-value for Hansen’s overidentification test in Table 5 Column 2 indicates that we can not reject the validity of the instrument (as do the overidentification test p-values in Table 4, with the Holiday instrument included). Columns 3 and 4 in Table 5 present an explicit form of the overidentification test. In these columns one weather instrument is presumed to be valid and the other is included as an exogenous regressor to determine if it has a direct effect on incidence of concession. If so, we would expect to see a significant coefficient on the instrument treated as exogenous, but instead, for both instruments their coefficients are very small and statistically insignificant upon their turn being treated as exogenous.

One obvious related, specific concern is that the weather instruments may be affecting the outcome not through their effect on violence but rather through an effect on protest size. It may be that, just as rain and temperature affect violent behavior at a protest they also affect turnout to the protest and impact concession via that channel instead, confounding
our preferred causal interpretation. More generally, it is known that the inclusion of a second endogenous regressor as a control may bias the coefficient on the endogenous regressor being instrumented (see, e.g, Appendix A in Acemoglou et al., 2001). The concern is acute when the instruments used for one endogenous variable are strongly correlated with the second endogenous regressor included as a control. In the case at hand, the protest size variable, included above in some specifications, may reasonably be thought to be endogenous for many of the same reasons as protest violence (e.g. an omitted indicator of what policies are perceived as going too far by the public and by elements of the governing coalition; intransigence or weakness of the decision making authority). However, OLS regressions of Size on the instruments reveal that there is no significant relationship between the variable and Temp and Precipitation. Both instruments have very small coefficients (0.0002 for Temp and 0.017 for Precipitation) that are not significantly different from zero (the t statistics for Temp is 0.01 and 0.63 for Precipitation. Taken all together, this evidence gives us confidence in the view that our instruments are valid (never something that can be proven absolutely), and that we are in fact identifying a negative causal relationship between protest violence and protest success.

2.6 Implications for Future Modeling

While the contribution of this paper is primarily empirical, the findings also shed light on the appropriateness of certain assumptions in modeling protest violence. We will now briefly discuss what the negative estimate on violence implies for future modeling of violent protest.

The obvious question raised by the negative finding is why so many protesters still choose to engage in costly violent protest if resorting to violence lowers their chance of success? Is such a decision irrational? Seen from the perspective of a basic two-party bargaining model with the protester as a unitary actor with rational expectations using violence to win concessions, this conclusion makes sense: if fully informed about the expected probability of gaining a concession, protesters would be acting against their interest to resort to violent tactics that in fact lower their chance of achieving their aims.

However, one need not throw away the assumption of rational actors to explain the findings. If instead of hewing to a unitary rational actor model, we allow for two types of protesters and the existence of political agency problems the negative result can be given a rationalist explanation. For instance, some protest participants may only partly share the objectives of protest organizers. In addition to deriving utility from exacting a concession, these protesters may additionally derive utility from the looting so often associated with protest violence, thus, putting their personal gain from the spoils of rioting in conflict with the collective loss experienced from protest violence lowering the prospect of achieving a concession for the issue under protest.

The motivational heterogeneity among participants in political violence has already been emphasized in previous work. Weinstein (2005, 2007), for instance, makes the distinction between opportunistic and ideological joiners to armed rebel groups, divisions which, to a
lesser extent, almost certainly exist in sub-military political protests as well. Powell (2006)
explains the decision to fight rather than to agree to a settlement by going beyond unitary
actors, allowing one side to be composed of competing factions that can not commit them-
several to a future division of their side’s gains. In work in progress by this author, we sketch
a simple signalling model that employs a biforcation in the protester types, or, alternatively,
in the perception of protester types held by the stakeholders to whom the protest target
answers (the voting public, shareholders). The upshot is that with complete uncertainty
about protester type, the government or firm will react to violence with no concession (and
to non-violence with a concession) when stakeholders sufficiently value concessions to nonvi-
olent protesters whom they perceive as principled, and, the net private gain to violence for a
criminal type of protester outweighs his net loss in the protest-issue-space from a failed but
violent protest (relative to what could have been won via a concession).

Of course, one does not need to abandon a unitary actor model to explain the results if
willing to allow for agents governed by “behavioral” features such as emotion or overconfi-
dence. Introducing emotion, for instance, into agent’s decision-making goes against the basic
rationalist model whereby decision making is seen as a cognitive process in which “decision
makers [are] assumed to evaluate the potential consequences of their decisions dispassion-
ately and to choose actions that maximize the “utility” of those consequences” (Loewenstein
and Lerner, 2003). This explanation for why political violence occurs would diverge from the
tradition in political science, influenced by Schelling (1960), which sees recourse to violence
as a strategic choice undertaken to coercively increase the chance of victory. While in inter-
state or insurgent-state conflicts the strategic model may fit, the non-hierarchical nature of
most protests and their diverse and open participation admits for participants who, unlike
leaders of states or armed forces, have not been selected for their ability to successfully de-
liver results or maximize societal, or factional, welfare functions. This would seem to allow
for a greater role for “behavioral” participants. Emotion can be incorporated with otherwise
rational elements (see for example, Passarelli and Tabellini, 2013) and it need not be the
only behavioral element admitted. Indeed, the same selection effect referenced might allow
those leading states or armies to have a better grasp of their odds of victory, while protests of
the masses may well include participants who are overconfident in their assessment of their
likelihood of success or in their assessment of how the protest target (and its constituents)
will respond to the tactic of violence.

The above explanations parallel some of the those offered by Fearon (1995) in his typology
of the explanations for the inefficiency puzzle of war. Just as in that instance, it is also
possible when explaining protesters’ recourse to ineffective violence to leave aside either of
the above explanations, and, to instead envision explanations that involve a strictly rational
and unitary actor. For instance, it may be that protesters are choosing violence to maximize
their returns over their life cycle. Even if violence loses them the chance of concession in the
current period it may decrease the chance in the future that additional hostile moves are taken
against the protesters by the protest target. This may make sense if facing a target who does
not want to be seen as weak from capitulating to protester coercion, and, so, who will not
reverse the policy under protest, but, who will be less likely to enact similarly controversial
policies in the future out of fear of additional violent reprisal. Thompson’s (2003) study of French agricultural protests by Coordination Rurale and other farmer groups opposed to EU trade and agricultural policy in the 1990s - protests that often led to violence in our data set - is consistent with this reading. It is also possible that the lower chance of getting a win with violent protest is compensated for by a more generous split of the pie in the event that one does win. Such an explanation could be ruled out by more fine-grained measures of outcome variables (going beyond simple binary measures of success). Distinguishing between the competing explanations will need to be done in future work that pays more careful attention to formal modeling and the collection of higher-quality data for empirical tests.

2.7 Conclusion

In this paper we attempt to address a main methodological problem in work that studies the efficacy of political violence, namely, the possible endogeneity of violence. We aim to go beyond previous correlational studies that yield conflicting conclusions about whether turning to violence pays off for protesters, and, we identify a causal effect of protest violence on the incidence of concession made to protesters in the short run. To do so, we use instrumental variables methods, with rain, temperature, and school holidays as instruments for the degree of violence, so as to identify plausibly random-like variation in violence and avoid omitted variable and endogeneity bias. We find a significant and negative relationship between property destruction associated with protests and the chance of near-term success in changing policy, and we interpret this IV estimate as indicative of a negative causal effect of protest violence on the incidence of concession. Our IV point estimates are significantly larger in magnitude than those from a comparable OLS regression, indicating bias from endogeneity or omitted variables in OLS specifications.

We hope this work demonstrates one avenue by which the study of the effects of political violence can be pursued with careful attention to credible causal identification. We readily admit that “a single estimate is unlikely to provide a definitive and comprehensive basis for informing policy” and what we need is multiple studies based on different populations and different settings to understand the causal mechanisms at work (Imbens, 2009). We are optimistic about the applicability of this particular identification strategy to other countries, as the mechanisms by which our instruments are assumed to affect violence are fairly general. The rain and temperature instruments are presumed to affect violence primarily through physiological processes that we would expect to be fairly universal (though there may be some variation in the range of sensitivity to bodily temperature changes at different latitudes). The holiday instrument similarly rests on the general tendency of youth to be more violent than average. The extent to which this is true will undoubtedly vary from culture to culture, but as long as youth are at least somewhat more violent than average then school holidays may prove to be a useful instrument in other contexts. More of a challenge to cross-country application of the holiday instrument is the degree to which compulsory schools laws are implemented and enforced: in areas where enforcement against truancy is so lax that many
students (possibly the students most likely to engage in property violence) do not regularly attend school, there will be little likelihood of a strong correlation between the nominal days off from school and broken windows at a protest. Thus, the current identification strategy will be easiest to replicate in nations like France, with obligatory participation in secondary education and strict enforcement of truancy laws. Youth participation in protest is also very high in France and the strength of the school holiday instrument also rests on the degree to which the young participate in protests when not incapacitated by school. Existing work suggests that France is not anomalous in this respect (see Resnick and Casale, 2011, for a survey of African nations).

While the identification strategy may be generalizable, it is an open question as to whether the negative estimates for protest violence will be observed in other countries. While many have noted the unusually high frequency of protest in France (something confirmed by cross-country comparisons over the same time period in EPCD data on other European nations), it is unclear that French protest targets should respond uniquely negatively to violent protest. In fact, if indeed protest targets choose how to respond to violent protest on the basis of whether these targets’ constituents perceive the protesters to be in the right, we may actually see even larger negative results in countries such as the United States where cultural attitudes towards protest and especially violent protest are more hostile than in France and where there is little to be gained from conceding. In addition to culture, institutions and forms of government also may play a critical role in mediating the response to violence. While non-democratic regimes may be thought to be most concerned with being seen as weak (and so may refuse to concede in the face of violence), in relatively democratic France we find that concessions are also not often granted in the face of protest violence. In weakly institutionalized environments where exerting political influence via established channels of voting and lobbying is less common, and where there is less disapproval for going outside these channels, we may actually see violence playing a more effective role in bargaining. To test any comparative differences more and better data is needed. Especially important for future micro-empirical analysis is the creation of comprehensive data sets that attempt to record not only the characteristics of protest, but, that also do the hard work of recording protest outcomes, both in the short run and long run. As in many cases, these data improvements will enable more sophisticated empirical - and we hope causal - analysis.
2.8 References


Baum, Christopher F. 2006. An Introduction to Modern Econometrics Using Stata. STATA Press.


Hsiang, Solomon H., Marshall Burke, and Edward Miguel. “Quantifying the Influence of
Climate on human Conflict.” Science. 10.1126/science.1235367


CHAPTER 2. QUIET RIOT: THE CAUSAL EFFECT OF PROTEST VIOLENCE


### 2.9 Appendix

Table 2.6: Expanded OLS, IV, and First-Stage Results with Additional Variables

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) IV-GMM</th>
<th>(3) First stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>-0.0222***</td>
<td>-0.145**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00543)</td>
<td>(0.0621)</td>
<td></td>
</tr>
<tr>
<td>Size(,1000)</td>
<td>-0.00657</td>
<td>-0.0377****</td>
<td>-0.151***</td>
</tr>
<tr>
<td></td>
<td>(0.00421)</td>
<td>(0.00898)</td>
<td>(0.0290)</td>
</tr>
<tr>
<td>Duration</td>
<td>0.000656***</td>
<td>0.000664***</td>
<td>-0.000700***</td>
</tr>
<tr>
<td></td>
<td>(0.000131)</td>
<td>(0.000177)</td>
<td>(0.000123)</td>
</tr>
<tr>
<td>Election</td>
<td>0.00401</td>
<td>0.0173</td>
<td>-0.0313</td>
</tr>
<tr>
<td></td>
<td>(0.0153)</td>
<td>(0.0113)</td>
<td>(0.0333)</td>
</tr>
<tr>
<td>Days in Paris</td>
<td>0.00184</td>
<td>0.00228***</td>
<td>-0.00252***</td>
</tr>
<tr>
<td></td>
<td>(0.00118)</td>
<td>(0.000619)</td>
<td>(0.000786)</td>
</tr>
<tr>
<td>Precipitation</td>
<td></td>
<td>-0.0520***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0148)</td>
<td></td>
</tr>
<tr>
<td>Temp</td>
<td>0.0347**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0151)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holiday</td>
<td>0.0549**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0242)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time trends</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.025</td>
<td>.</td>
<td>0.065</td>
</tr>
<tr>
<td>N</td>
<td>2087</td>
<td>2087</td>
<td>2087</td>
</tr>
</tbody>
</table>

Note: Standard errors are reported in parentheses, adjusted for protest target level clustering

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$