University of Tennessee, Knoxville
Trace: Tennessee Research and Creative Exchange

Doctoral Dissertations Graduate School

8-2015

The Effects of Trade Competition on Health, and Determinants of Workplace Behavior

Thomas Clayton McManus
University of Tennessee - Knoxville, tmcmanus@vols.utk.edu

Recommended Citation
McManus, Thomas Clayton, "The Effects of Trade Competition on Health, and Determinants of Workplace Behavior. " PhD diss., University of Tennessee, 2015.
https://trace.tennessee.edu/utk_graddiss/3446

This Dissertation is brought to you for free and open access by the Graduate School at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a dissertation written by Thomas Clayton McManus entitled "The Effects of Trade Competition on Health, and Determinants of Workplace Behavior." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Economics.

Georg Schaur, Major Professor

We have read this dissertation and recommend its acceptance:

William S. Neilson, Scott Gilpatric, John Bell

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

The Effects of Trade Competition on Health, and Determinants of Workplace Behavior

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Thomas Clayton McManus
August 2015

Copyright © 2015 by Thomas Clayton McManus
All rights reserved.

ACKNOWLEDGEMENTS

I would like to thank the many people who have helped me through this journey in so many different ways. I am forever appreciative.

First, I would like to thank my parents and most influential teachers, Bob and Holly, and my siblings, Mallary, Wade, and Sam, for their constant love and support over the years. They have shaped me into the person I am, and I would not be here without them. I would also like to especially thank my fiancée, Ashley, who has been my steadfast companion these past few years and a constant source of encouragement and confidence. For that, and for the sacrifices she has made so that I might pursue my goals, I am very grateful. I look forward to continuing this great adventure together.

I am indebted to all of the educators who have taught me so much and along the way cultivated my passion for learning and research and for sharing that with others. Chief among them, I would like to thank my advisor, Georg Schaur, who has invested incalculable time and energy in me and my dissertation. He has been all at once demanding supervisor, magnanimous mentor, patient teacher, exacting colleague, staunch supporter, and loyal and compassionate friend, and in each role he has meant more to me than I can say. I owe just as much to Bill Neilson and Justin Rao, who have both been fantastic teachers and mentors and some of my biggest advocates. Their commitment and investments in me have been crucial to my success.

Thanks also to the many other faculty at the University of Tennessee and elsewhere who have helped me in so many capacities, especially Scott Gilpatric, John Bell, Tibor Besedeš, Ernie Cadotte, Celeste Carruthers, Scott Holladay, Marianne Wannamaker, Christian Vossler, and Mike Price. Thanks also to Calvin Blackwell and Betsy Jane Clary, who instilled in me a passion for economics and research during my undergraduate studies. I would also like to express my appreciation to Donna Kemper, Susan McGee, Sherri Pinkston, and Kathy Zelem, the wonderful administrative assistants who have done so much for me over these past five years.

I would like to extend a special thanks to my fellow graduate students for their support and feedback, which has markedly improved this dissertation, and for their friendship, which has been valuable in so many other ways. I would like to especially thank Matt McMahon, Justin Roush, Jens Schubert, Jing Li, Yin Chu, Becky Davis, Jill Welch, Michael Craig, Annette Tower, and Lawrence LaPlue for their help. Thanks to the many others who have provided valuable feedback on this research, including conference participants and discussants at the EITI Conference, Midwest Meetings, SEA Meetings, and International Trade Workshop at Georgia Tech, and seminar participants at the BLS, ILAB, Mary Washington, Tennessee, Wofford, and Xavier.

Lastly, I thank the Russell Sage Foundation for funding the experiments in the first chapter of this dissertation (Award #98-13-07). I thank also the Department of Economics at the University of Tennessee for funding that research and for providing generous financial support throughout my graduate studies.

ABSTRACT

My dissertation consists of three essays related to workplace behavior. In the first paper, we design a controlled laboratory experiment to study image motives in a setting where decisions signal intelligence. The experiment results show that in some settings social scrutiny can discourage individuals from making choices that signal their intelligence, despite evidence that the signal was privately valuable. In the second paper, we study the effect of Chinese import competition on occupational safety and health at US manufacturers. We find that a change in US trade policy and Chinese import shocks significantly increases worker injury and illness rates in competing US industries, especially at smaller, less productive plants. This paper presents the first evidence that import shocks affect welfare through changes in worker health. Building on this, in the third paper we look at broader mental and physical health effects of import competition in local labor markets. We find that import exposure worsens overall mental, physical, and general health in the surrounding area. The effects are greatest for mental health and among the employed, consistent with theory from the health literature pertaining to the documented effects of import competition on wages, employment, and job security.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

This dissertation is comprised of three independent chapters. The first chapter is a behavioral experiment studying concerns for perceived intelligence. The latter two chapters study the effects of import shocks on physical and mental health outcomes for workers at import-competing firms and throughout the local economy.

In the first chapter, we study image motives in an environment where economic decisions signal intelligence and competency. Evidence in the social psychology literature explains that these traits are intrinsically and socially desirable, implying that actions which signal these traits are more attractive. We design a new laboratory experiment to study these effects empirically. Subjects attempt either an easier or harder set of verbal analogy questions and are paid for correct answers such that harder questions are worth more. Payments are calibrated so that high-ability subjects earn more attempting the hard questions while low-types make more with the easy questions. This makes choosing the hard questions a valid signal of intelligence. Sorting behavior is publicly revealed in our "audience" treatment, facilitating social signaling. In the "intrinsic only" treatment, the signaling mechanism was explained but decisions were kept private. In the control, there was no discussion of the signaling mechanism and all decisions were private. We find that intrinsic only subjects were significantly more likely to choose the high-type action than the control. In comparison, subjects were significantly *less* likely to choose the signal in the audience treatment, when doing so was publicly observed. The effects are more pronounced in males. The results suggest that social observation can demotivate individuals when decisions signal intelligence, despite evidence that the underlying trait was privately considered desirable.

The second chapter looks at the relationship between trade and occupational safety and health, an important determinant of worker welfare that has been largely ignored in the literature. Our theory explains that firms facing greater shut down risk reallocate resources to improve productivity at the expense of safety. Therefore, import shocks worsen safety conditions at firms which are at greater risk of shutting down as a result. We test this prediction with novel data on reported injuries at US manufacturers using growth in Chinese imports in the years 1996-2007 as a shock to competition. The data show that injury rates in the competing US industries increase over the short to medium run, particularly at smaller establishments. Back of the envelope calculations show that injury risk increases by 13 percent at the smallest establishments, costing workers the equivalent of a 1 to 2 percent reduction in annual wages. This chapter presents the first evidence to our knowledge that import shocks affect worker welfare through changes in health. These results have implications for trade and regulatory policy.

The final chapter of this dissertation builds on these findings by studying broader health measures, including mental health, and by looking at consequences throughout the economy, rather than just for workers who remain at import-competing jobs. We find that average mental, physical, and general health worsens in US local labor markets exposed to greater Chinese import competition between 2000 and 2007. The effects are greatest for mental health. Moving a region from the 25^{th} to 75^{th} percentile of import exposure corresponds to a 5.5% increase in the time individuals report suffering from poor mental health, adding about 0.18 days per month. The effects are greatest for the employed, consistent with theory from the health literature pertaining to the documented effects of import competition on wages, employment and job security. We also find that import exposure in a region significantly increases the share of people unable to afford necessary medical care. These estimates provide direct evidence that import exposure in local labor markets affects overall mental and physical well-being, as well as access to health care.

CHAPTER I
SIGNALING SMARTS? REVEALED PREFERENCES FOR SELF AND SOCIAL PERCEPTIONS OF INTELLIGENCE

This paper is collaborative work with Justin M. Rao at Microsoft Research. A version of this chapter was originally published by T. Clay McManus and Justin M. Rao:

T. Clay McManus, Justin M. Rao. "Signaling smarts? Revealed preferences for self and social perceptions of intelligence." *Journal of Economic Behavior and Organization* 110 (2015): 106-118.

Abstract

We design a laboratory experiment to test for image motives in a setting where decisions signal intelligence to a social audience. Money-maximizing behavior in the experiment sorts subjects by academic ability, as measured by performance on verbal analogy questions, across two levels of question difficulty. Sorting behavior is publicly revealed in our "audience" treatment, facilitating social signaling. In the "intrinsic only" treatment, the signaling mechanism was explained but decisions were kept private. In the control, there was no discussion of the signaling mechanism and all decisions were private. We find that intrinsic only subjects were significantly more likely to choose the high-type action than the control. In comparison, subjects were significantly *less* likely to choose the signal in the audience treatment, when doing so was publicly observed. The effects are more pronounced in males. The results suggest that social observation can demotivate individuals when decisions signal intelligence, despite evidence that the underlying trait was privately considered desirable. Audience effects have a less predictable impact on behavior in this setting as compared to the near universally positive findings from the altruism and trust literature. Our experimental design can be easily adapted to study image motives in a broad set of environments using revealed preferences.

1.1 Introduction

There is considerable evidence that individuals alter their behavior to signal personal traits to a social audience and even themselves (Glazer and Konrad, 1996; Harbaugh, 1998; Bénabou and Tirole, 2006; Ellingsen and Johannesson, 2008; Andreoni and Bernheim, 2009; Karlan and McConnell, 2013). Existing studies of image motives typically focus on qualities that are unambiguously positive, such as generosity, altruism and fairness (Soetevant, 2005; Dana et al., 2007; Alpizar et al., 2008; Koch and Normann, 2008; Ariely et al., 2009a; Funk, 2010; Lacetera and Macis, 2010; DellaVigna et al., 2012), and find, with near unanimity, a positive impact of public observation. In this paper we study behavior that signals academic ability, a trait that is no doubt desirable in many settings, such as a job interview, but potentially objectionable in others, such as socializing with one's peers.

We design a laboratory experiment that allows subjects to send credible, but potentially costly signals of intellectual ability and ambition.[1] Our main finding is that when observed by others subjects distort their behavior and avoid displaying these traits. Instead of projecting themselves as "high types", they opt for choices associated with "low-type" ability when observed by a social audience, despite evidence that the underlying trait was privately considered to be desirable. Our results suggest that observation by one's peers can be demotivating when making choices that connote traits related to intelligence and ambition.

At the center of our experiment is a mechanism for signaling intelligence. Subjects answered verbal analogy questions in a series of rounds and were paid a piece rate for each correct answer. Before each round, subjects chose to answer questions from either an easier set, taken from an *undergraduate* entrance exam (SAT), or a more difficult set, taken from a *graduate school* entrance exam (GRE), which had weakly higher piece rates. Incentives were designed – and empirically verified – so that subjects with relatively high academic ability maximized earnings by choosing the harder test and those with lower ability did so with the easier test. Choosing the hard test is therefore a credible signal that subjects may use to influence perceptions of their ability.

Two treatment protocols and a control were used to isolate signaling motives. The experiment was divided into three stages. All three conditions were identical through stage 1 and decisions were always private in this stage. In stage 2, the "audience" treatment allowed subjects to signal their ability to other subjects (and the experiment proctor) through a straightforward procedure: they stood and were recognized if they answered questions from the hard test. This recognition mechanism was explained in our "intrinsic only" treatment as well, but subjects were truthfully informed that it was used in "other sessions" and their choices would remain private. Information sets were thus identical across the two treatment conditions through the end of stage 2 when choices were revealed in the audience treatment. Choices in stage 3 were private for both treatment conditions. In the control, subjects completed all three stages in private and without any mention of the recognition procedure.

Cross-treatment comparisons of our three conditions identify the main causal effects of interest. We employ a difference-in-difference strategy such that the point of comparison across treatments is within-subject changes in behavior relative their idiosyncratic baselines.[2] Comparing audience to intrinsic only cleanly isolates the net effect of social observation, our main result. Subjects in intrinsic only differ from control in that they learn that subjects were elsewhere grouped by their test choices. We compare behavior in these protocols to identify the effects of this priming, including self-signaling and other intrinsic motives. Lastly, comparing the audience treatment to control gives an

---

[1] Social psychologists have substantiated that intelligence is an important determinant of social and self-esteem, making it a likely domain for audience effects. Intelligence is found to be highly socially desirable (Park and Kraus, 1992) and associated with greater social happiness (Solomon and Saxe, 1977) and mate appeal (Prokosch et al., 2009; Li et al., 2002). Murphy (2007) documents behavior manipulations people use to convey intelligence in interpersonal communication. See Gottfredson (1997) for a discussion on the pervasive importance of intelligence.

[2] The key takeaways are similar to a simple comparison of means.

estimate of the combined impact of social signaling and the intrinsic motives that arose with the mechanism's explanation and implementation.[3]

Our results are as follows. The greatest increase in likelihood of choosing the hard test in stage 2 (relative the stage 1 baseline) occurred in intrinsic only, followed by audience and then control. The likelihood of choosing the hard test in intrinsic only exceeded the control by 16.8 percentage points, a magnitude that is nearly 100% of the base rate in the control. Subjects were also 13.2 percentage points more likely to choose the hard test in intrinsic only than the audience treatment. Both of these differences are statistically significant beyond the 5% level. The audience treatment exceeded the control by a statistically insignificant 3.4 percentage points. The results suggest subjects placed greater value on the hard questions in private after learning of the social recognition procedure that took place in *other* sessions, but were hesitant to "show off" to their peers. The audience exerted a negative causal effect when information sets were held constant.

In stage 3, audience subjects were more likely to choose the hard test (in private) if they observed that a relatively high number of their peers had done so in stage 2. This pattern is consistent with intrinsic preferences for social conformity (Lindbeck et al., 1999). However, we cannot separately identify a taste for conformity from social learning about the money maximizing strategy and leave this as an interesting effect to explore in future work.

Overall, we believe the results cleanly show that signaling ability to one's peers is socially *un*desirable in some settings. Publicizing aptitude or ambition may be counter-productive, even if the underlying trait is privately considered to be desirable. These findings have implications for the effects of esteem rewards like "Employee of the Month" programs in the workplace and social recognition in the classroom. Aversion to "showing off" the otherwise valued trait may push workers or students away from the desired results of the program. While social recognition programs may not have a gross negative effect – the audience treatment was slightly more inclined to choose the hard test than the control – they could be improved by endorsing the underlying qualities with more private esteem rewards that minimize the potential scrutiny of one's peers (or perhaps a subset of one's peers). In this regard, our findings inform the growing literature in personnel economics on non-monetary incentives.[4]

Putting our results in the context of the literature, theoretical models suggest a broad role for audience effects to affect the behavior of economic agents, while existing empirical work has focused almost exclusively on communicating fairness and altruism in giving decisions. In deciding to donate money or reciprocate trust, audience effects, self-esteem motives, and other-regarding preferences have all been found to push individuals in the same direction, namely towards the decision consistent with the "good"

[3] Within-subject changes may also be affected by the dynamic effects of learning and experience in the experiment. The comparison to control accounts for these effects. We thank anonymous reviewers for this suggestion.

[4] In a recent field study, for example, Ashraf et al., (2014) compare the effects of different non-monetary rewards on the training performance of health workers in Zambia. Private information on relative performance hurt test outcomes, a result the authors attribute to information avoidance and concerns for self-perceptions of ability.

trait. The countervailing effect of social-signaling motives in our experiment on intelligence stands in contrast to these findings and suggests that audience effects influence behavior less predictably than we previously might have thought and can indeed be counterproductive in some settings.

Our final contribution is to the experimental design literature. We introduce a new mechanism for laboratory experiments that can be used to test other theories of sorting behavior. It is particularly well-suited for studying labor market decisions and workplace behavior with student subjects as these tests are commonly used to assess students' "job" performance and competency.[5] More broadly, our methodology can be adapted to study image motives in a wide range of economic settings. A researcher can design a sorting experiment where payouts are pinned to an index of some trait of interest and test for audience effects in that domain by comparing behavior when the task choice is publicly revealed to a private baseline.[6]

The remainder of the paper is organized as follows: Section 2 presents the experimental design, Section 3 contains the predictions for the experiment, Section 4 provides the results, Section 5 presents a discussion, and Section 6 concludes. Tables and figures are provided in Appendix A.1.

1.2 Experiment

The experiment took place in the University of Tennessee Experimental Economics Laboratory. Subjects in the experiment were students there who had registered to participate in economics experiments and were recruited through the Online Recruitment System for Economic Experiments (Greiner, 2004). They were recruited individually via email as our intention is to test for social signaling motives in a peer group without relying on pre-existing relationships or priming a group identity; where it existed, any familiarity between subjects in a session occurred either at random or in the event that individuals made plans to attend together.

We conducted 14 sessions with a total of 148 subjects, 4 intrinsic only (IO) sessions with 41 subjects, 6 audience treatment (AT) sessions with 63 subjects, and 4 control (C) sessions with 44 subjects. Each session had 8-12 subjects, seated at two rows of computer terminals. All decisions were entered into the computer, and the experiment was programmed and conducted in z-Tree (Fischbacher, 2007). Subjects randomly drew their station assignment and decisions were recorded using a station identifier only so that

---

[5] Performance on standardized test questions has previously been used as a real effort task to generate earnings (List and Cherry, 2000; Cherry et al. 2002; Oxoby and Spraggon, 2008). Our design adds to this paradigm with a mechanism that retains classical properties of a signaling game (i.e. individuals have private information of an ability measure and signaling cost is decreasing in ability).

[6] For example, an experiment might ask subjects to choose between a safe outcome and a risky gamble, where the probability of the high outcome depends on their gender or racial attitudes, measured by an Implicit Association Test, such that the value of the gamble was decreasing in the level of a subject's bias towards one gender or race. If subjects are concerned with appearing unbiased to a social audience, they would be more inclined to choose the gamble when doing so is publicly revealed. We thank a referee for pointing out this generalization and example.

individual decisions and performance were kept anonymous to the experimenter, whose observation may otherwise trigger unwanted audience effects in both treatment and control sessions.

Our design needed to possess the following features: 1) subjects must be able to assess their own ability within the experiment and make decisions consistent with their ability; 2) the signaling mechanism – choosing the harder test – must form a credible signal of ability; 3) subjects must understand (2). We'll now explain the details of how we met these requirements.

We chose a task, verbal analogy questions from standardized tests, that we believed subjects would associate with intelligence. While the strength of this association presumably varied across subjects, we find that choosing the harder test was on average a valid signal of academic ability (Tables A.1 and A.2) and that subjects correctly interpreted it as such (Table A.3) – the presence of some subjects who did not view it as a meaningful measure of intelligence will only bias us towards a null results.[7] In each round subjects were given four minutes to answer six multiple choice questions and earned a varying piece-rate for each correct answer. Earnings were denominated in tokens, with a maximum 5 tokens per correct answer, which were converted to dollars at the end of the experiment at a rate of 10 tokens to $1.00. Subjects were paid for every (correctly answered) question in the experiment. Before each round, subjects chose between answering from a harder set of questions, taken from the GRE (Graduate Record Examination), or an easier set of questions, taken from the SAT (Scholastic Aptitude Test). The SAT is taken by prospective undergraduate college students to assess their academic readiness for college, and the GRE is taken by post-baccalaureate applicants seeking degrees such as a Master's or PhD. The GRE is a more difficult test by construct and is designed to sort a set of higher ability types.

There were three piece rate "pairs", one of which was randomly selected each round for each subject based on a known probability distribution. For each, a subject chose which test to take if that wage pair was implemented. The harder test paid weakly more in each pair. Subjects observed their draw and then attempted a set of questions taken from the exam they selected for that wage pair. After each round, they were shown their own performance and earnings. To provide familiarity with the difficulty of the two exams, subjects first completed a practice round of questions for each exam, in random order, and were paid 5 tokens for each correct answer.

Figure A.1 shows the timeline for the experiment. The experiment was partitioned into three stages to isolate the behavioral channels of interest with the instructions for each stage directly preceding it so as to keep information sets separate. Each stage consisted of four decision rounds, with six questions in each. In control (C) sessions, test choices were kept private in every stage and there was no mention of recognition or any new information given during the experiment. In the two treatment protocols, stage 1 was a baseline and identical to control. At the beginning of stage 2, audience treatment (AT) subjects were informed that the end of the stage subjects would stand based on whether

[7] Non-native speakers (about 10% of subjects in our experiment), for example, may not have felt their performance or decision to take the harder questions to be a valid indicator of their intelligence.

they answered from the (harder) GRE or the (easier) SAT for a majority of rounds. The standing mechanism that took place in AT was also explained to subjects in intrinsic only (IO) at the start of stage 2, but they were (truthfully) told that choices in their session would remain private. After completing stage 2 and standing, subjects in AT completed the third stage with the instructions that their choices thereon would remain private. Subjects in IO and C also completed stage 3 privately, but did so without any knowledge of others' choices.

1.2.1 Sorting

It was necessary that money-maximizing behavior in the experiment be related to ability to facilitate sorting along the desired dimension and support our tests for signaling. The money-maximizing exam choice is determined by a subject's relative success rate: the likelihood of correctly answering questions from one exam relative the other. An individual will maximize earnings by choosing the GRE if her relative success rate is greater than the ratio of piece rates. With risk-neutrality, the decision rule for choosing the GRE is given by:

$$\frac{P_{GRE}}{P_{SAT}} > \frac{w_{SAT}}{w_{GRE}} \tag{1}$$

where P_j is the success rate on exam j and w_j is the piece rate offered for each correct answer. Risk-averse individuals, when expected earnings are equal, will maximize expected utility by choosing the easier SAT, as it has the lower variance in expected earnings.

The success of our design is underpinned by a single-crossing condition that makes choosing the GRE a signal of high academic ability: individuals with higher ability must have higher relative success rates. That is, it must be the case that individuals who do very well on the SAT tend to do well on the GRE, but individuals who do fair on the SAT tend to struggle considerably with the GRE. This makes sense from a test design point of view, as the GRE is designed to segment individuals within the higher end of the ability distribution. This condition was supported by simulations using publicly available data on test performance and the distribution of scores for our subject pool, University of Tennessee students. We empirically confirm signal validity within our experiment.

The degree to which a pair of piece rates separates individuals depends on the underlying distribution of relative success rates. The use of three piece rate pairs and a variant of the strategy method – with nature acting as first mover – segments subjects to a finer level than a single pair would allow. GRE piece rates were held constant at 5 tokens for simplicity and the three SAT piece rates were {3,4,5}. We henceforth refer to each pair of piece rates in ratio form, {0.6,0.8,1}. The wage pair was randomly selected with probabilities {0.25,0.5,0.25}.

A subject who chooses to answer from the GRE for a particular SAT piece rate should also choose the GRE for any lower SAT piece rate. Accordingly, the strategy set consists of four non-dominated strategies for the three piece rate ratios {0.6,0.8,1}. They are: {SAT, SAT, SAT}, {GRE, SAT, SAT}, {GRE, GRE, SAT}, and {GRE, GRE,

GRE}. Following the decision rule above, the money-maximizing strategies based on success rates are shown in Figure A.2.

The lower two piece rate ratios were chosen as targets to separate the subjects by ability based on simulations. Two rates were used to ensure adequate separation and hedge against the risk that subjects pooled into one strategy, which was of concern given this design had never been implemented before. The third pair was a one-to-one offer for the SAT and GRE piece rates, a purely wasteful signal if GRE questions are in fact more difficult for everyone.

1.2.2 Audience Effects

Subjects were read the initial instructions at the beginning of the experiment and told that additional instructions would be given prior to beginning each new stage. This ensured the observations in stage 1 were identically framed in all sessions and could serve as a baseline for comparison. Subjects in control sessions are not given any additional instructions throughout the experiment, and provide a point of comparison that accounts for learning and experience in the experiment but not the signaling motives introduced in the treatments.

We introduce social observation in stage 2 in AT, and test for audience effects using an AT-IO comparison of behavior in stage 2 relative the baseline. Before beginning stage 2, subjects in AT were told that at the end of the subsequent four rounds those who attempted questions from the GRE in two or more rounds would be asked to stand at their station. To control for any preferences for standing and gaining social attention that are independent of concerns for perceived intelligence, the complementary set of subjects—those that attempted the SAT in 3 or more rounds—then stood. A counter on the test selection screen displayed the number of SAT and GRE rounds that a subject had attempted thus far in the stage.

The explanation of the standing procedure in AT sessions at the beginning of stage 2 may introduce confounding influences that are unrelated to social observation. A shift to the harder test may instead reflect self-esteem concerns associated with meeting the benchmark implicitly endorsed by the experimenter in order to *self*-signal intelligence or accomplishment. To isolate audience effects in the AT-IO comparison, we explained the standing mechanism to subjects in the IO treatment also, while making it clear that standing would not take place in their session.

The sorting decision allowed individuals to sacrifice earnings in exchange for social recognition through choice of test, not performance. This distinction parallels giving experiments used to test for audience effects: a selfish dictator who has concerns for audience impressions can still signal pro-social preferences by giving without revealing that he is in fact selfish and would not give absent the social scrutiny.

The randomness in selecting the offered piece rate injects noise into the social signal as only the exam attempted, and not one's full strategy and the piece rate offered, is publicly revealed. As a result, an individual could choose the GRE for the 1:1 relative wage if his value for standing and signaling intelligence to the audience is sufficiently high. If piece rates were publicly revealed in addition to the exam chosen, choosing the GRE in this case would instead signal incompetence or braggadocio, as it is the less profitable choice for individuals of all abilities. The cost of this method is that noise added to the signal weakly decreases its value for individuals concerned with social recognition, which will bias towards a null result.[8]

After stage 2, all subjects in AT were asked to stand in place with their respective group for 8 seconds via a message displayed on their computer.[9] In stage 3, subjects in both treatments and in control all completed four rounds privately. We test whether subjects in AT altered their behavior in stage 3 in response to the observed actions of others in the group.

After finishing the experiment, subjects were asked to complete a brief questionnaire. The monitor who oversaw the experiment exited the room and a third-party paid subjects in sealed envelopes. The experiment lasted approximately 90 minutes and earnings ranged from \$9.40-\$38.00, with an average of \$23.66.

1.2.3 Mechanism Performance

Before presenting our signaling predictions and results, we appraise the design mechanism by assessing if subjects were sorted by underlying academic ability such that those who chose the GRE generally have greater ability and that, if so, subjects recognized this. The conditions for mechanism validity are addressed in lockstep below.

First, the results from the experiment confirm that GRE questions were in fact more difficult than SAT questions. The kernel density estimates and cumulative distributions for the success rates on both exams are shown in Figure A.3 and Figure A.4, respectively. This is confirmed by a test for first-order stochastic dominance (Davidson and Duclos, 2000), given in Table A.4. We can thus assert that there exists, for each subject, a piece rate between 0 and 1 that separates her money-maximizing exam choice.

Figure A.5 shows subjects' average success rates on both exams and average strategy choice in the experiment. The rays from the origin represent the three piece rate ratios offered, and give the regions for money-maximizing test choices (see Figure A.2). Most subjects in the experiment sorted in accordance with their estimated success rates. Inaccurate beliefs over true success rates may have caused deviations in earlier rounds and we will argue that choices in later rounds were impacted by social and self-signaling motives. We present the quintile-quintile version of these distributions in Figure A.6.

For the GRE to be a valid signal of intelligence, sorting must separate subjects by underlying academic ability. We define baseline sorting behavior by subject decisions in stage 1, before new information and signaling are introduced (all three experimental protocols are identical through this point). Using self-reported scores on the ACT, an undergraduate entrance exam, as an external yardstick measure of academic ability, we

---

[8] The stochastic element supports the parallel to recognition mechanisms like an employee of the month program. Consider a store employee and an experiment participant who respectively exhibit baseline behavior that assures they will not be recognized as employee of the month or a GRE-chooser. The participant knows the benchmark but is uncertain of the action that will result from his choices, while the worker knows her actions but is uncertain whether they are enough are enough to secure recognition, but in both cases the recognition outcome is stochastic and the likelihood is increasing in the cost of the action undertaken by the agent.

[9] All subjects (in AT) followed the instructions to either stand or remain seated at the appropriate time.

find that the mechanism successfully sorted subjects such that GRE-choosers on average had significantly greater academic ability than SAT-choosers (see Table A.1).

Finally, our design requires that subjects recognize decisions as a valid signal of academic ability. We affirm this using data on impressions of academic ability that were gathered in the post-experiment questionnaire. Subjects were asked to evaluate their academic ability as it compared to others in the session and indicate their relative position within the group.[10] An individual who believes that those around him are more intelligent should report a lower relative position for himself, all else equal. Shown in Table A.3, AT subjects who observed a greater share of others choosing the GRE had significantly lower self-evaluations of relative academic ability.

1.3 Predictions

In this section, we present our hypotheses for social and self-signaling intelligence in the experiment and the theory and related empirical evidence that supports them.

1.3.1 Social Signaling

We compare behavior across AT and IO to test for social signaling motives. Subjects in AT stood after stage 2 to reveal which test they attempted more often, while behavior in IO was kept private. The two treatment protocols are otherwise identical when subjects make their choices in stage 2. This frames a clean test for our audience effects hypothesis.

<u>Net Positive Audience Effects Hypothesis</u>: Subjects in AT are more likely to choose the GRE in stage 2 than subjects in IO.

The hypothesis is motivated by findings that individuals distort behavior so as to signal themselves to others as altruistic or fair (Dana et al., 2007; Andreoni and Bernheim, 2009; Lacetera and Macis, 2010), and evidence in social psychology that suggests intelligence, like altruism, is a socially desirable trait (Park and Kraus, 1992). The only difference between AT and IO is that choices were publicly revealed in the former, so that observed changes in behavior are attributable to the presence of a social audience. A difference-in-difference comparison allows us to test if the audience effect leads to a positive difference in signaling motives on net. We call this the *net* positive audience effects hypothesis to reflect the distinction that audience effects must be

[10] Specifically, we adapted Svenson's (1981) question on driving confidence: "We would like to know about what you think about your academic ability. We want you to compare your academic ability to the academic abilities of other people in this experiment. By definition, someone has the greatest level of academic ability in the room and someone has the lowest level of academic ability in the room. We want you to indicate your own estimated position compared to others in this experimental group. Of course, this is a difficult question because you cannot be entirely sure based only on your experiences and you do now know all the people gathered here today, but please make the most accurate estimate you can." Subjects were asked to indicate there position along a line of radio buttons from lowest academic ability in the room to highest, each representing a decile in the distribution.

stronger than any crowd-out of internal motivations (discussed below). There is evidence that extrinsic rewards, such as money or public recognition, can lessen the intrinsic value of signaling choices (Gneezy and Rustichini, 2000b; Meier, 2007; Mellström and Johannesson, 2008). If the public nature of the decision damages the intrinsic reward for choosing the hard questions, the net audience effect will only be positive if social rewards outweigh this moderating effect.

1.3.2 Self Signaling and Other Intrinsic Motives

Although our experiment is primarily designed to test for social signaling motives, the description of the standing procedure in IO serves as a prime that could excite internal motivations. Just as individuals may alter their behavior in our experiment so as to affect others' perceptions of them, evidence from psychology and related work in economics suggests self-image concerns may motivate subjects to choose the GRE in private when they recognize it as a signal of intelligence. Psychologists posit that individuals assess their own behaviors like they assess the behavior of others, adopting the role of outside observer in forming self-perceptions (Bem, 1972). Self-impression management theory explains that people are consequently motivated to show themselves in a positive light, even when they are the only member of the audience (Baumeister, 1998). Related economic models of self-image and ego explain that signaling actions affect agents' self-impressions when they are unable or unwilling to introspect the true motives behind them (Bénabou and Tirole, 2002, 2004, 2006; Bodner and Prelec, 2003; Koszegi, 2006). Giving experiments find evidence of self-signaling in that environment.[11] We compare behavior in IO in stage 2 to the control to test for self-signaling and other intrinsic motivations primed by the explanation of the signaling mechanism.

<u>Intrinsic Motivation Hypothesis</u>: Subjects in IO are more likely to choose the GRE in Stage 2 than subjects in control.

While we have motivated this hypothesis with a discussion of self-signaling, we give it the more general "intrinsic motivation" label because we are unable to independently identify self-signaling motives from other priming effects, such as endorsement or recommendation effects. Indeed it may well be impossible to distinguish some of these motivations from the self-signaling hypothesis because they are often implicit in any esteem reward mechanism. For example, publicly or privately conferring awards for hitting $\$x$ in sales in a month or scoring a grade of x on a school assignment makes x a more salient outcome and induces endorsement and demand effects, in addition to providing a vehicle for affecting self- or social-esteem. This tight link means we cannot cleanly disentangle these motives in our experiment.

Relatedly, the conformity literature suggests that self-impression concerns will push subjects towards the revealed "norm" after observing the decisions of others (Lindbeck et al., 1999). We use stage 3 choices in AT to identify how the distribution of

[11] Murnighan et al. (2001), Dana et al. (2007), Lazear et al. (2012), and Tonin and Vlassopoulos (2013) find evidence of self-signaling altruism or fairness, while Grossman (2012) finds contradicting results.

others' choices affects the (private) behavior that follows, but we cannot separate internal pressures to conform from social learning about their money-maximizing strategy since both may compel subjects to take the GRE when they observe others doing so. Accordingly, we do not list this as a formal hypothesis.

1.4 Results

The mean likelihoods of choosing the GRE in each stage are shown in Figure A.7. The middle piece rate is where all the interesting action is. The monetary incentives dominate for the other two piece rates. For Pair 1 ($w_{SAT} = 0.6w_{GRE}$), the strong majority of subjects chose the GRE in the baseline and behavior is very similar across all conditions. The monetary incentive similarly dominates in Pair 3 ($w_{SAT} = w_{GRE}$), where the GRE was suboptimal for all money-maximizers and the costliest signaling action. Recall that the middle piece rate was selected to be implemented 50% of the time, as it was our prior that this is the wage that would separate subjects. In AT a subject who selected the GRE for the middle piece rate each round had a 95% chance of "standing" (attempting 2 or more GRE rounds in stage 2), relative a 26% chance if selecting the GRE for the lower piece rate only.

The aggregate results for the middle piece rate suggest that subjects were motivated to choose the GRE when it was framed as a signal and decisions were private, but were not so motivated when doing so was socially recognized. We use pooled OLS to estimate a linear probability model of the likelihood of choosing the GRE as a function of treatment status and stage, and use a difference-in-difference approach to identify the treatment effects.[12] The probability of selecting the GRE for a particular piece rate is modeled as:

$$(2)\quad \Pr(\text{test} = \text{GRE}|\text{piece rate} = \text{r}) = \alpha_0 + \beta_1 S_2 + \beta_2 S_2 T_{IO} + \beta_3 S_2 T_{AT} + \beta_4 S_3 + \beta_5 S_3 T_{IO} + \beta_6 S_2 T_{AT} + \gamma \text{RSR}_{\text{ante}} + \delta 1_{\text{RSR}_{\text{ante}} \geq \text{r}} + \theta + u$$

where T_{IO} and T_{AT} are indicators for the two treatment groups, S_j is an indicator for a round in stage j, RSR_{ante} is the ratio of success rates for all questions attempted in prior rounds, and θ is an individual fixed effect. Also included is an indicator for whether the subject's relative success rate for the full experiment is higher than the piece rate (indicating the GRE to be the money-maximizing choice).

Shown in Table A.5, the results indicate that the likelihood of selecting the GRE was increasing in the relative success rate observed in earlier rounds. This suggests that subjects updated their beliefs based on observed performance and correctly responded to

[12] The results of a pooled logit model support our main results at the 5% level also, but we prefer the OLS specification given its consistency under unspecified heteroskedasticity and with subject fixed effects and a relatively small number of observations for each (36 pair-rounds). Of further appeal, the linear specification provides straightforward estimates for the marginal effects of treatment and stage on the probability of taking the hard exam.

this information. Also, subjects whose relative success rates exceeded the threshold were more likely to choose the GRE for the two separating piece rates, as designed.

We use the difference-in-difference estimates for the middle piece rate from our main specification (coefficients in column 2) to test for audience effects and intrinsic motives. These estimates are shown in the first column of Table A.6. In support of our intrinsic motivation hypothesis, the IO-C comparison for stage 2 reveals that subjects were 16.8 percentage points more likely to choose the GRE after the standing mechanism was explained but not implemented. The result is consistent with self-esteem concerns attached to meeting the benchmark associated with intelligence.

Subjects in AT, in contrast, were only weakly more likely to choose the GRE in stage 2 when social recognition was introduced than the control. Counter to our positive audience effects hypothesis, social recognition did not motivate subjects in the experiment to choose the GRE, despite evidence that doing so was a valid and recognizable signal of academic ability. In fact, subjects in AT were 13.4 percentage points *less* likely to choose the GRE in the second stage of the experiment than subjects in IO. The interpretation of this behavior is discussed below.

We estimate the main specification for each gender separately and present the difference-in-difference estimates in columns 2 and 3 of Table A.6. We find that the observed effects were greatest among male participants. Males were 18 percentage points less likely to choose the GRE in AT than they were in IO. Gender differences in these environments could be investigated more thoroughly in future work.

Lastly, we examine the impact of the publically revealed signals on future behavior. We use a linear probability specification similar to our preferred specification to test whether subjects in AT were on average more likely to choose the GRE if they observed that a greater share of other participants in the session had been doing so.[13] Without time-varying information on this belief measure, we cannot use individual fixed effects in the estimating equation. We include individual characteristics observed in our data to control for other determinants of behavior. The results are presented in Table A.7.

Subjects were significantly more likely to choose the GRE if they witnessed that a larger share of others has done so. We estimate that for a session of ten participants, seeing an additional person stand for taking the GRE increased the likelihood that an observer chose the GRE for the middle piece rate by 3 percentage points at the mean. As mentioned in the introduction, this difference could occur through self-image (conformity) channels or subjects updating their perception of the money-maximizing strategy through social observation.

1.5 Discussion

1.5.1 The Sorting Mechanism

Related work has looked at the effects of a social audience on performance (Mas and Moretti, 2009; Ariely et al. 2009b; Bandiera et al., 2013; Ashraf et al., 2014), but in these

[13] This measure is assumed to be orthogonal to a subject's own choices as decisions remained private information through stage 2.

protocol intentional distortions in behavior—such as those suggested by our experiment—are indistinguishable from involuntary changes in response to social scrutiny, such as anxiety. For example, Ariely et al. (2009b) finds that individuals solving anagrams for a piece rate perform worse when they complete the task in front of others than when they do so in private. They posit that this underperformance is the result of "choking" under the added social incentives of performing well. Our findings raise an alternative explanation, namely that individuals might underperform because they do not want to be socially observed as doing well on the task. Our mechanism gets around this problem by creating a wedge between signaling and task performance. Importantly, we retain signal credibility when creating this wedge – even though performance is never revealed to any other participant. This could be particularity useful in studying more sensitive topics such as racial attitudes.

Our results demonstrate that the mechanism can effectively sort subjects by academic ability and serve as a credible signal. The results can also help calibrate the choice of wages for future studies. For our subjects at a major state university a single piece rate ratio of 0.75 would evenly separate individuals across the two exams and maximize the number of subjects on the margin of the decision rule, who are the most likely to distort behavior in response to non-pecuniary incentives. This value can be tailored to any U.S. institution using publicly available data on admissions test scores (a procedure we followed and validated).

1.5.2 Audience Effects in the Experiment

We find empirical evidence that social observation has more nuanced effects on behavior than has previously been found. If social signaling is not unambiguously attractive, the presence of an audience may influence behavior in different ways than is described by findings from experiments focusing on giving and altruism. We find that social observation discouraged signaling intellectual ability and ambition in our experiment. In fact, the negative audience effect almost entirely swamped positive internal motivations excited by our experimental prime, such that the overall effect was only slightly positive but not significant.

Of course, the audience effects we observe are determined by the private and social value placed on signaling a particular dimension of ability – at Ivy League schools, for example, we might see the directionality of the audience effect reversed. Nonetheless, our findings suggest that in some settings social observation can undo rather than reinforce the desire to behave in manner consistent with the endorsed trait, and that this can occur even when the trait is privately considered to be desirable.[14] We'll now discuss three explanations for this result. These explanations can have very different implications, highlighting the nuanced impact an audience can have in these environments.

"Smarty Pants" Effect

[14] In an example of positive audience effects in a related dimension, the aforementioned Ashraf et al. (2014) finds that social recognition improved test performance for the Zambian health trainees.

A straightforward explanation of the result is that students at the University of Tennessee did not want to be seen by their peers as having high academic ability or as trying too hard in the experiment. Subjects also often turned in their answers early (63% of the time with at least one minute remaining); though doing so did not expedite the experiment, it could be inferred by others in the session whether one was working on the task or had finished.[15] This behavior is consistent with models of attention avoidance and conformity (Bernheim, 1994) if subjects believed ex-ante that the majority would choose the easier questions. Choosing the harder test questions could have instead been associated with arrogance or a type of intelligence that is socially undesirable. This explanation would be particularly important for "Employee of the Month" programs or public recognition programs in schools. An interesting aspect of this explanation is that knowing individuals' private assessment of a trait is not sufficient to predict how they will feel about publicly expressing it.

Social Preferences

Social preferences offer another explanation for the negative treatment effect observed in our experiment. Bandiera et al. (2005) studies the productivity of fruit pickers under both piece rate pay and relative incentives, and find that productivity is at least 50% higher under piece rate pay. They find that workers internalize the negative externality that their effort imposes on others when pay is determined by relative output. If subjects in our experiment recognize that standing for taking the GRE affects the self-esteem of others in the session and have social preferences over their utility, choosing the GRE would be relatively less attractive than it otherwise would. Decisions in IO sessions (and control) were private and therefore this externality was not present. These motives would not impact an "Employee of the Month"-type programs in which there is always one winner, but have an impact when the number of "winners" is endogenous.

Motivational Crowding Out

Introducing the social recognition reward may have altered the perception of choosing the GRE for subjects in AT as compared to IO. It's possible the internal motivations for choosing the GRE were not the same across the two treatment sessions because the social reward dampened the intrinsic value of attempting the harder questions, just as extrinsic rewards for giving have been found to "crowd out" intrinsic motivations (Frey and Oberholzer-Gee, 1997; Gneezy and Rustichini, 2000b; Meier, 2007; Mellström and Johannesson, 2008; Holmås et al., 2010) and image motivations (Ariely et al. 2009a). The negative AT-IO difference in stage 3 is consistent with Meier's (2007) finding that crowding out persists even after the extrinsic incentives are removed. An important note for future work is that motivational crowding out cannot lead individuals to choose the high-type action less when it's framed as a social signal than they otherwise would absent

[15] Recessed computers with privacy screens ensured decisions were private, but there were not dividers in the lab so it was observable to others whether or not a subject was interacting with the terminal. Subjects who finished early were observed emitting behaviors such as reclining in the chairs with their arms crossed or hands on their head.

any signaling motives, whereas the smarty pants effect and social preferences can have a gross negative impact. Given that the audience treatment and control were not statistically different we cannot speak definitively on this issue in this paper.

1.6 Conclusion

In this paper, we design a sorting mechanism for laboratory experiments that separates subjects by academic ability. Sorting decisions served as a valid and recognized signal of ability, and subjects could make costly changes in behavior so as to signal their ability. The experiment retains the important characteristics of sorting mechanisms in the real world, and may be viable for future tests of signaling theory. We use the experiment to test previously unanswered predictions on how audience effects impact signaling intelligence to one's peers.

We find that social observation discouraged subjects from signaling academic ability, despite evidence that the underlying trait was privately considered to be desirable by revealed preferences in our intrinsic only condition. Our findings illustrate the difference in audience effects in signaling ability to one's peers as compared to giving decisions, which have been the primary focus of prior work. In contrast to those settings, here the social value of signaling ability is not unambiguously positive, and social observation can actually be demotivating as compared to privately conferred esteem rewards. These findings have substantial practical importance as well, in particular for the use of recognition programs in schools and the workplace that rely on the positive impact of audience effects.

In looking at intelligence, our study contributes evidence of audience effects in an altogether new domain where choices signal ability. Broadly, our results support a wider role for audience effects than has been previously studied and our experiment paradigm – linking task payouts to an index of some trait like intelligence and then varying social observation of sorting – can be easily adapted for future studies of audience effects in other settings.

Appendix A.1 Tables and Figures

Table A.1. Comparison of Means for Ability Measures, Based on Baseline Sorting

	First Piece Rate			Second Piece Rate		
Ability Measure	**SAT**	**GRE**	**Diff.**	**SAT**	**GRE**	**Diff.**
ACT Scores	25.90	27.37	2.07**	26.85	29.09	2.24***
	(0.74)	(0.42)	(0.85)	(0.42)	(0.72)	(0.83)
GPA	3.174	3.370	0.196*	3.317	3.333	0.016
	(0.087)	(0.044)	(0.098)	(0.047)	(0.759)	(0.089)
Earnings	20.84	24.55	3.71***	23.44	24.17	0.73
	(0.93)	(0.52)	(1.00)	(0.52)	(1.00)	(1.13)

Notes: Standard errors are reported in parentheses. GRE-choosers defined by those who chose GRE for the piece rate in at least 2/4 rounds of the stage (analogous to signaling criteria). ACT scores and GPA were self-reported in questionnaire. Sample used in estimation excludes participant who did not correctly answer an SAT question in any round (incl. practice). In the spirit of the design, earnings in the experiment should be a viable proxy for academic ability. The results for the first piece rate show that earnings are significantly higher among those that sorted into the GRE. The difference is negligible across the two groups for the middle piece rate, however. One explanation is that for higher piece rates, fewer subjects optimally choose the GRE. Those "over-chasing" the GRE due to nonmonetary incentives or incorrect beliefs make up a relatively greater share of the group for higher piece rates and also earn less by behaving suboptimally, lowering the average and countering the predicted result.
* p<0.10, ** p<0.05, *** p<0.01

Table A.2. Strength of the GRE as a Signal of Ability, Based on Baseline Sorting

	Probability that GRE-chooser has higher ability than SAT-chooser	
Measure of Ability	**First Piece Rate**	**Second Piece Rate**
ACT Scores	0.64^{**}	0.66^{***}
GPA	0.61^{**}	0.51
Earnings	0.68^{***}	0.55

Notes: Significance levels are from a Wilcoxon rank-sum test of the hypothesis that the ability measure is different for GRE-choosers and SAT-choosers. Subjects are grouped by their majority choice in stage 1 for the respective piece rates (with the tie going to the GRE, like the decision rule for standing). ACT scores and GPA were self-reported in the questionnaire. Excludes participant who did not correctly answer an SAT question during the practice round or a decision round.
* p<0.10, ** p<0.05, *** p<0.01

Table A.3. Academic Ability: Self-Evaluated as Percentile Within Session

	Audience Treatment	Intrinsic Treatment + Control
Constant	60.407***	61.505***
	(5.295)	(5.390)
Percentile on Undergrad Entrance Exam	0.370***	0.132*
	(0.070)	(0.075)
Percent of Others Who Stood for GRE‡	-0.197**	0.010
	(0.085)	(0.010)
Observations	58	64
R-square	0.344	0.07

Notes: Standard errors are reported in parentheses. The second regressor is a subject's ACT score percentile within the distribution of experiment participants.

‡Subjects in intrinsic only and control were not asked to stand, but the share who qualified to stand is included as a regressor there as a point of comparison. A program crash in an intrinsic only session damaged the questionnaire data for those subjects. They are necessarily excluded from these regressions, as are subjects who declined to answer the relevant survey questions.

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

Table A.4. Success Rate Densities and FOSD Test Statistics

	Densities		Davidson-Duclos t-statistic:
Success Rate	**SAT**	**GRE**	$\boldsymbol{F_S(x) \leq F_G(x)\ \forall\ x}$
0.39	0.05	0.28	5.62
0.46	0.12	0.43	6.22
0.52	0.17	0.61	8.80
0.59	0.24	0.73	9.73
0.66	0.41	0.82	8.00
0.73	0.56	0.95	8.80
0.80	0.76	0.99	6.14
0.86	0.91	0.99	3.50
0.93	0.99	1.00	1.42

Table A.5. Probability of Selecting the GRE by Piece Rate: Pooled OLS Estimates

	Pair 1: $w_{SAT} = 0.6w_{GRE}$	Pair 2: $w_{SAT} = 0.8 \cdot w_{GRE}$	Pair 3: $w_{SAT} = w_{GRE}$
Stage 2	0.051	-0.066**	0.010
	(0.036)	(0.028)	(0.044)
Stage 2 x IO	0.036	0.168***	-0.014
	(0.060)	(0.039)	(0.055)
Stage 2 x AT	0.068	0.034	-0.058
	(0.055)	(0.056)	(0.050)
Stage 3	-0.001	-0.085*	-0.019
	(0.062)	(0.049)	(0.029)
Stage 3 x IO	0.108	0.173**	0.032
	(0.082)	(0.068)	(0.050)
Stage 3 x AT	0.079	0.053	-0.034
	(0.073)	(0.068)	(0.039)
Money-Maximizing	0.269***	0.118**	0.040
	(0.065)	(0.053)	(0.091)
Relative Success Rate in Previous Rounds	0.240***	0.331***	0.186***
	(0.065)	(0.067)	(0.043)
Observations	5262		
R-square	0.571		

Notes: Standard errors are reported in parentheses. Estimates are from a pooled OLS regression for all pair choices. Regression includes pair and subject fixed effects. Robust standard errors are clustered by both subject and by session-stage using variance estimator from Cameron et al. (2011).
* p<0.10, ** p<0.05, *** p<0.01

Table A.6. Difference-in-Difference Estimates for Middle Piece Rate: Changes in Probability of Selecting GRE Relative Stage 1 Baseline

Stage 2	Full Sample	Men	Women
Control	-0.066**	-0.019	-0.113***
	(0.028)	(0.032)	(0.041)
Intrinsic only	0.102***	0.164***	-0.020
	(0.022)	(0.050)	(0.060)
Audience Treatment	-0.033	-0.013	-0.048
	(0.048)	(0.060)	(0.062)
IO-C Diff.	*0.168****	*0.182***	*0.093*
	(0.039)	*(0.073)*	*(0.073)*
AT-C Diff.	*0.034*	*0.005*	*0.066*
	(0.056)	*(0.074)*	*(0.071)*
AT-IO Diff.	*-0.134***	*-0.177***	*-0.027*
	(0.055)	*(0.085)*	*(0.095)*
Stage 3	Full Sample	Men	Women
Control	-0.085*	-0.012	-0.159**
	(0.049)	(0.061)	(0.066)
Intrinsic only	0.089*	0.170	-0.051
	(0.048)	(0.111)	(0.041)
Audience Treatment	-0.032	-0.024	-0.032
	(0.053)	(0.075)	(0.070)
IO-C Diff.	*0.173***	*0.182*	*0.108*
	(0.068)	*(0.133)*	*(0.072)*
AT-C Diff.	*0.053*	*-0.013*	*0.127*
	(0.068)	*(0.097)*	*(0.082)*
AT-IO Diff.	*-0.120**	*-0.194*	*0.019*
	(0.070)	*(0.136)*	*(0.084)*
Observations	5262	2661	2193
R-square	0.571	0.618	0.555

Notes: Standard errors are reported in parentheses. Estimates are from pooled OLS regressions for all pair choices, first for the full sample and then for each gender subsample. Each regression includes pair and subject fixed effects. Robust standard errors are clustered by both subject and by session-stage using variance estimator from Cameron et al. (2011). Full sample includes subjects without gender information from questionnaire; excluding them does not significantly affect the results.
* p<0.10, ** p<0.05, *** p<0.01

Table A.7. Probability of Selecting the GRE as a Function of Others Seen Doing So

	Pair 1: $w_{SAT} = 0.6 \cdot w_{GRE}$	Pair 2: $w_{SAT} = 0.8 \cdot w_{GRE}$	Pair 3: $w_{SAT} = w_{GRE}$
Optimal	0.507***	-0.017	0.287***
	(0.187)	(0.065)	(0.0935)
Relative Success Rate in Previous Rounds	0.185	0.914***	0.0005
	(0.267)	(0.282)	(0.0421)
Percentage of Others Seen Choosing the GRE	0.0017	0.0029***	0.0012***
	(0.0011)	(0.0008)	(0.0004)
Percentile on Undergraduate Entrance Exam	-0.009	-0.114	-0.063**
	(0.197)	(0.089)	(0.030)
Observations	696		
R-square	0.643		

Notes: Standard errors are reported in parentheses. Estimates are from a pooled OLS regression for all pair choices. Regression includes pair fixed effects, a gender dummy, and school year dummies for each piece rate pair. Robust standard errors are clustered by both subject and by session using variance estimator from Cameron et al. (2011). * p<0.10, ** p<0.05, *** p<0.01

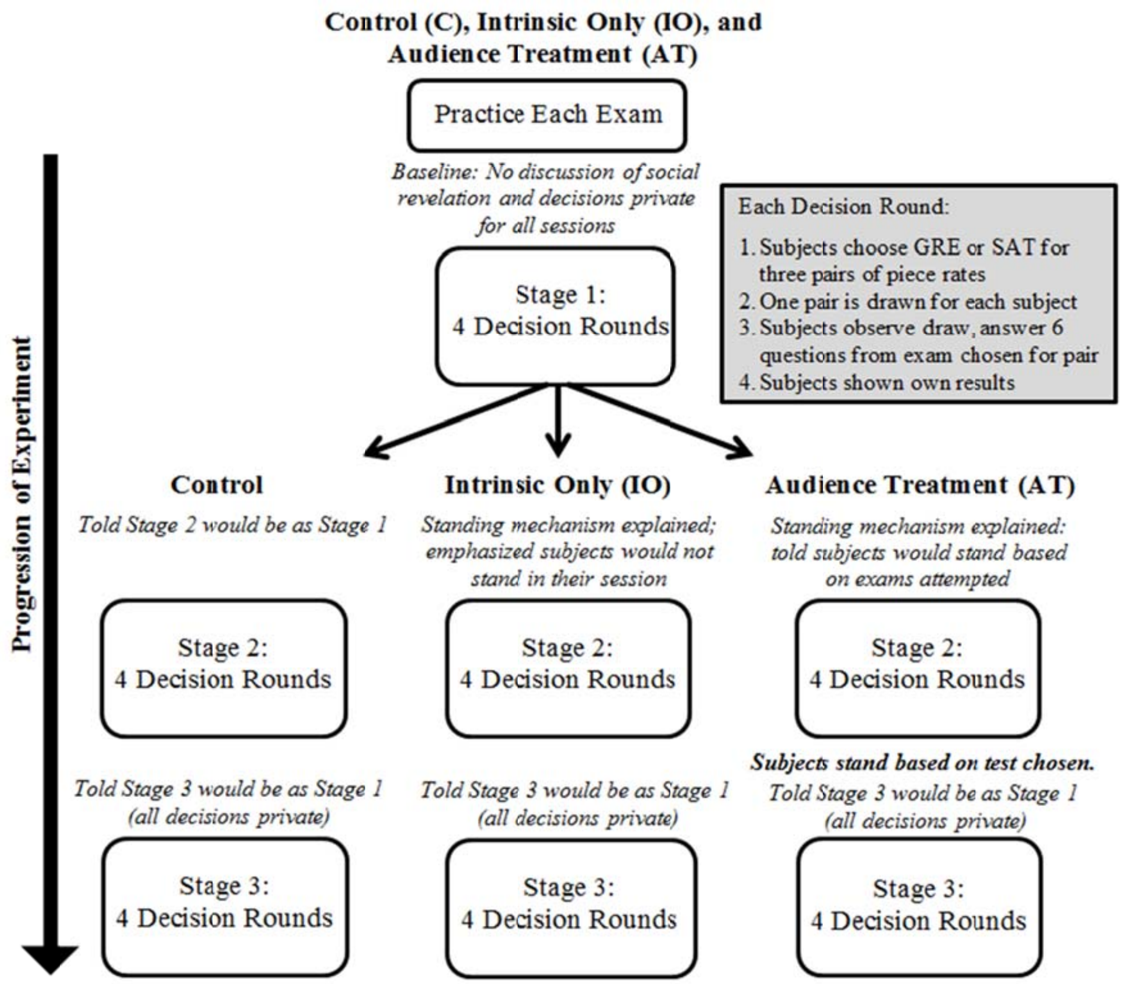


Figure A.1. Experiment Timeline

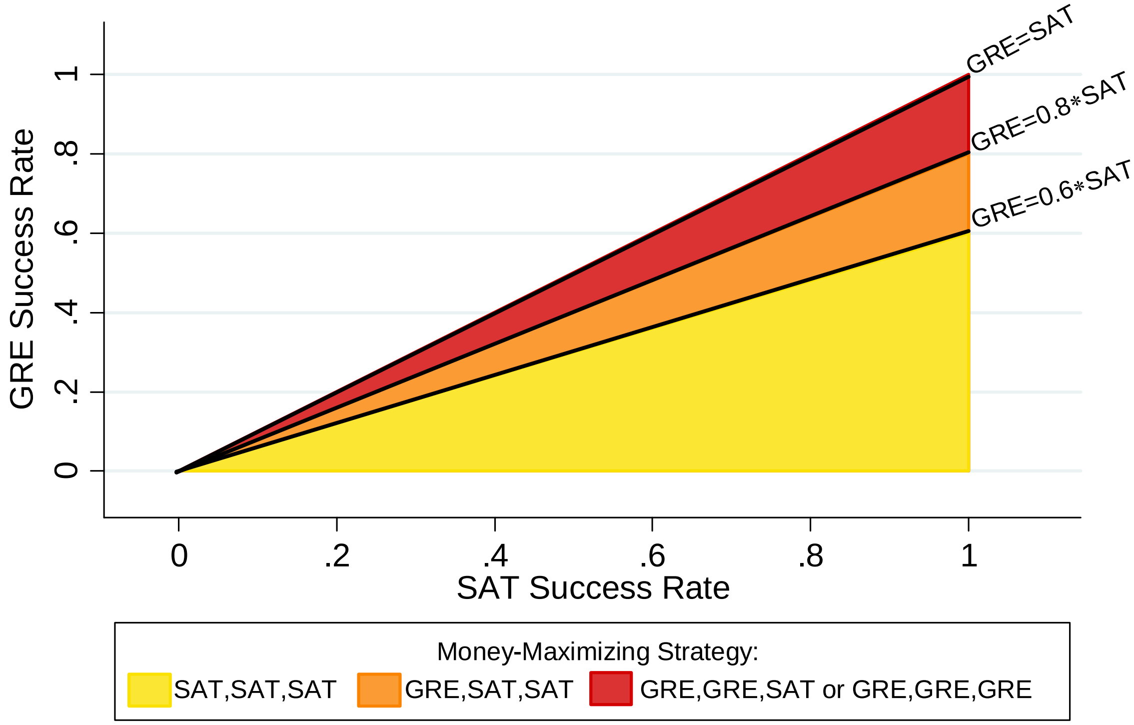


Figure A.2. Money-Maximizing Strategies by Success Rates

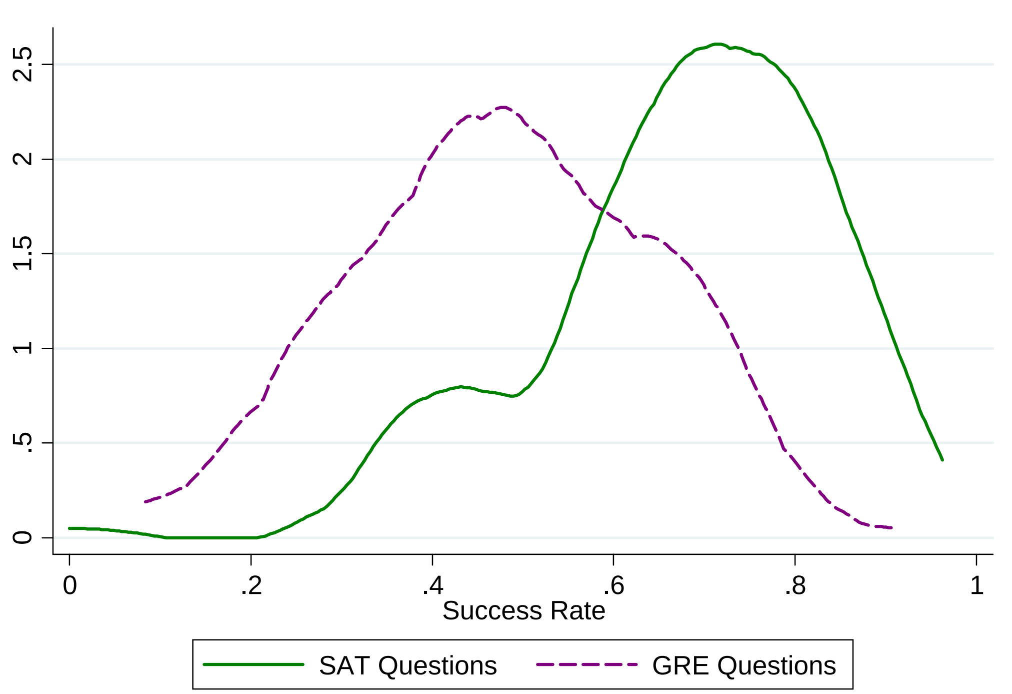


Figure A.3. Kernel density estimates of success rates for SAT and GRE questions.

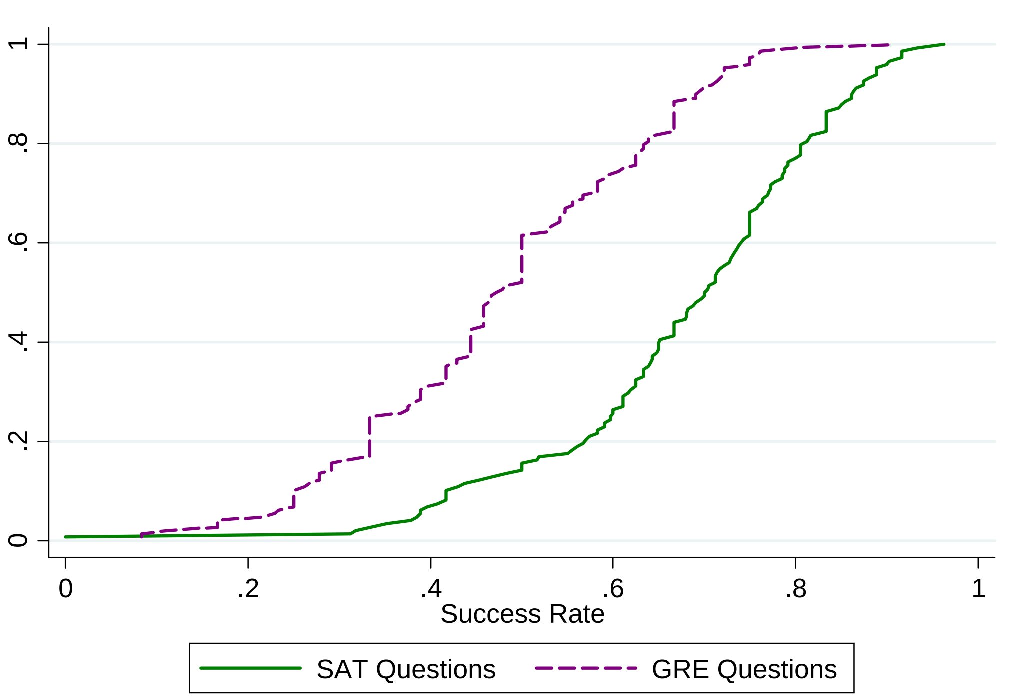


Figure A.4. Cumulative distribution of subject success rates for SAT and GRE questions.

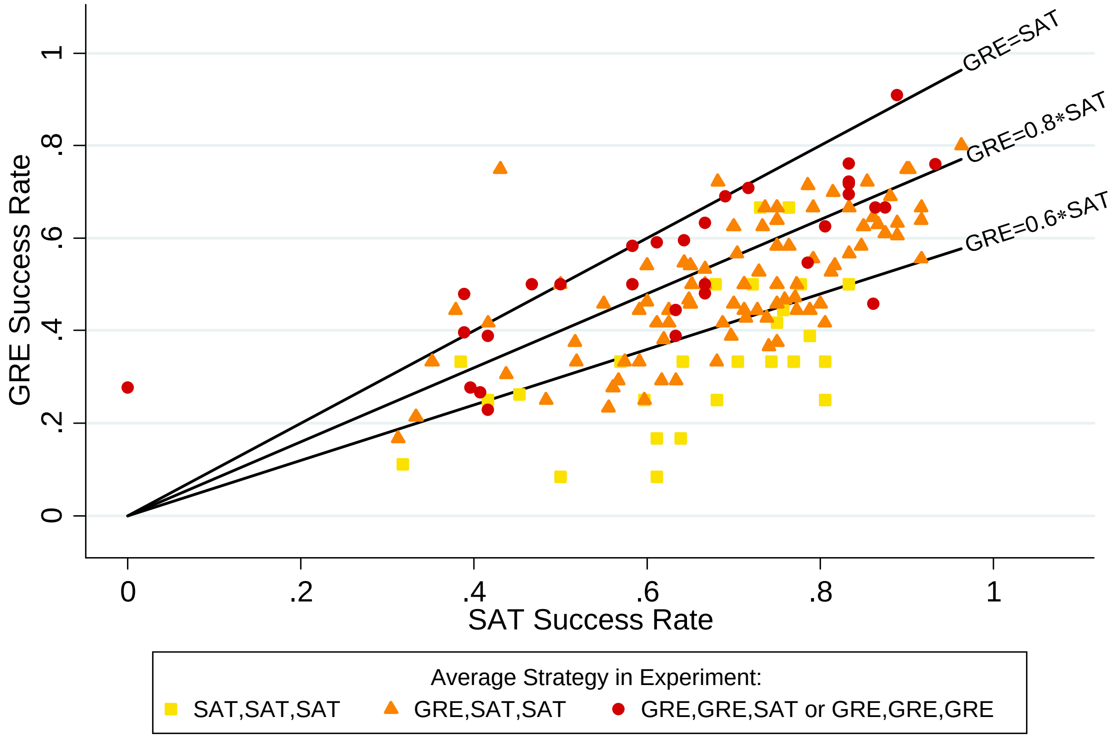


Figure A.5. Subject success rates in the experiment, colored by exam choice.

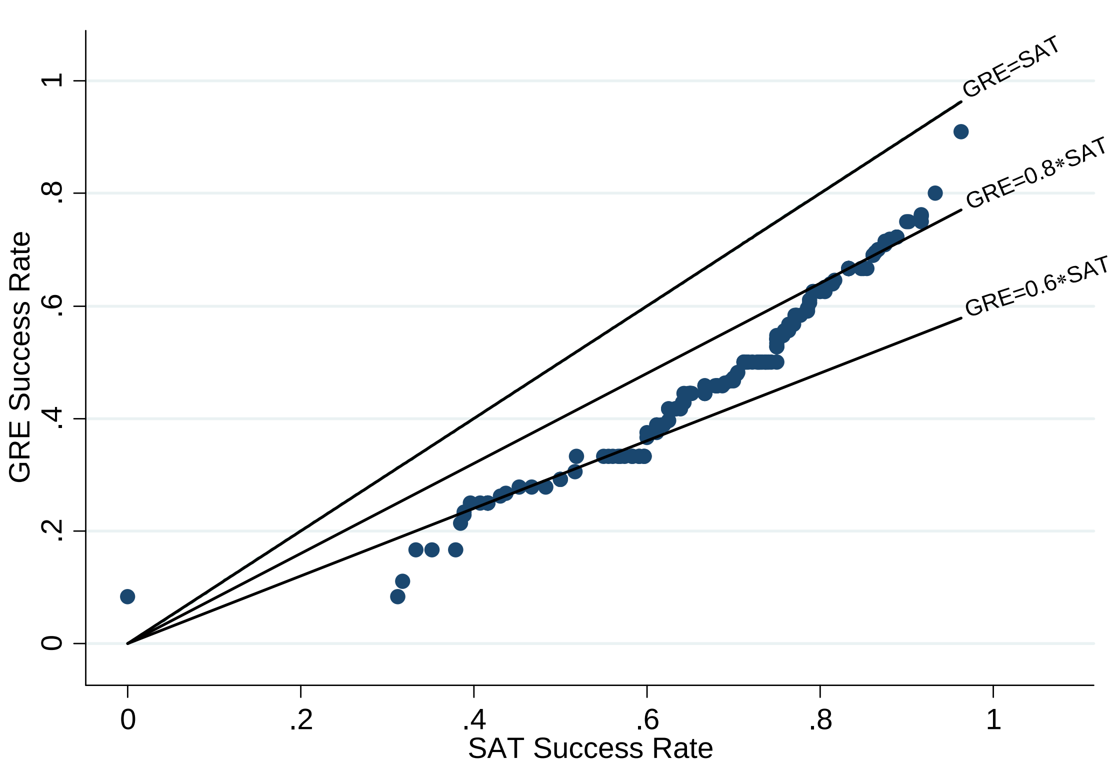


Figure A.6. Quintile-quintile plot of subject success rates on SAT and GRE questions.

Notes: Each point on the scatter plot represents a percentile and the coordinates reflect the associated success rate for questions from each exam. It seems reasonable that the rank order of subjects based on their true ability on SAT verbal analogy questions is quite similar to that for the GRE.

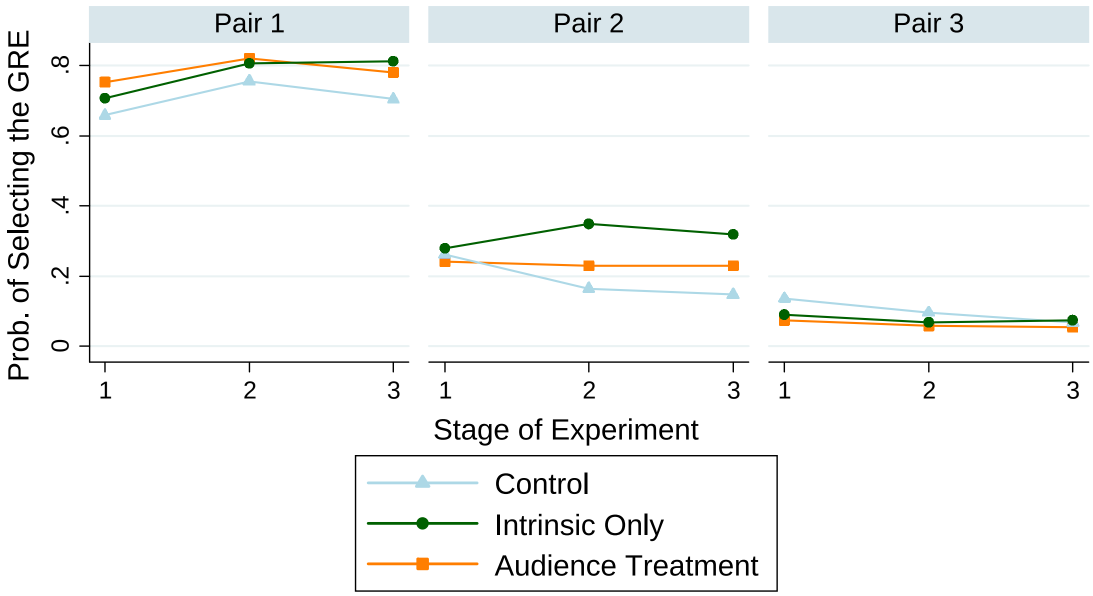


Figure A.7. Mean Likelihood of selecting the GRE by condition for each piece rate pair.

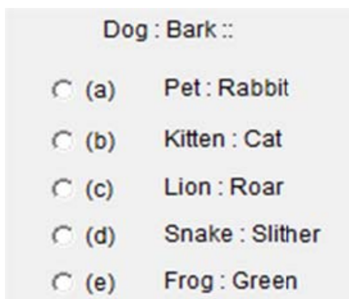
Appendix A.2 Experiment Instructions

We are conducting an experiment in the economics of decision making. A research foundation has provided funding for this research. You will receive at least $5 as a participation fee for showing up today. Your earnings beyond this will depend partly on your decisions and performance, and partly on chance. These decisions will be described below. Your earnings during the experiment will be measured in experimental tokens. At the end of the experiment, these tokens will be converted into dollars at the following rate and paid to you in cash:

10 Tokens = 1 Dollar

Tasks

This experiment consists of 12 independent rounds. In each round, you will have 4 minutes to answer 6 analogy questions and you will earn an amount of tokens for each correct answer. Analogies, common on many standardized tests, require you to analyze the relationship between two words and determine which answer choice contains the pair of words that share a relationship similar to the original pair. A very simple example is:

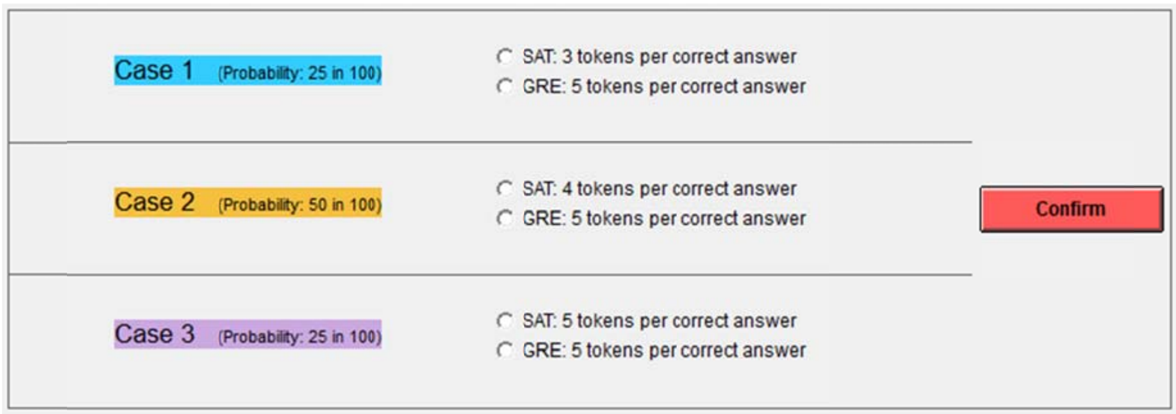


The correct answer would be (c): **Dog** is to **Bark** as **Lion** is to **Roar**. In both pairs of words, the second word is the sound made by the animal in the first word. While **Kitten** is similar to **Dog** in that both are pets, the relationship each shares with the other word in its pair are not the same. The other answers are incorrect for similar reasons.

Your Decisions

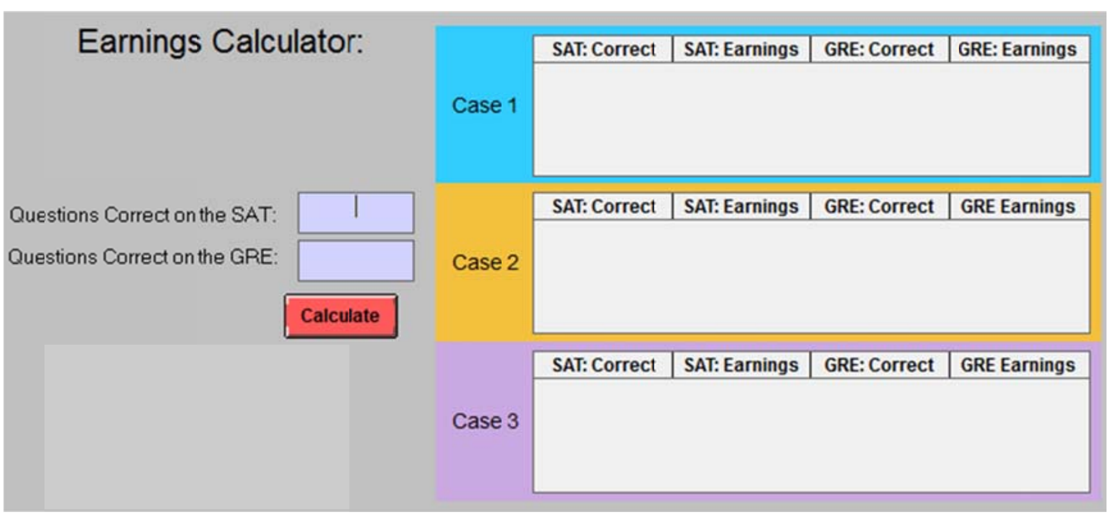
In each round, you can choose whether your analogy questions come from past SAT or GRE exams. The SAT is a college admissions test taken by high school juniors and seniors to assess their academic readiness for college. The GRE is a graduate admissions test taken by college seniors and college graduates applying to graduate programs to receive further education, such as a Master's degree or PhD. It is similar in this way to the LSAT, the exam taken to enter law school, or GMAT, the exam to enter a Master's in Business Administration program. Because SAT questions are intended for high school graduates and GRE questions for university graduates, questions from the SAT are generally easier than questions from the GRE.

The amount you earn per correct answer in each round will depend on which test you choose. You will be shown three pay schemes, each specifying the payment for a correct answer on the GRE and the payment for a correct answer on the SAT. In each round, you must select for each of the three cases whether you would like to answer questions from the SAT or questions from the GRE if that case is offered.



Note that the pay rate for GRE is always 5 tokens per correct answer, but the pay rate for SAT varies from 3 to 5 tokens. The probability or chance each case is chosen is as follows: Case 1, 25%; Case 2, 50%; Case 3, 25%. A case is randomly drawn for each participant in each round, so the case you are offered will differ from other participants. Because you do not know which case will be chosen, it is in your best interest to select for each pay scheme the test that you would like to answer from if that case is in fact selected.

On the computer screen, you will see an “Earnings Calculator” to help you select a test for each case:



In the light blue boxes, you can enter different values for the number of questions you might answer correctly out of the 6 questions given from each test and click "Calculate" to view your potential earnings for the two tests in each case. You may do this as many times as you like, and you will have the opportunity to do this before each round. Further, a list of the tests you have taken in each previous round and your performances is also provided to help you make decisions:

Your Previous Decisions and Performance:

Round	SAT Correct	GRE Correct	Pay Rate	Earnings
Practice				
1				
2				

Once you've made your decisions for each case and clicked "Confirm", a case will be drawn for you. You will be told which case was drawn and reminded of the test you chose for that case.

You will be given 4 minutes to answer 6 questions from that test. After you have submitted your answers, you will be told how many questions you answered correctly and how much you earned in that round. We will then begin a new round, and you will again choose a test for each pair of earnings and answer questions based on your choices and which pair is drawn. We will repeat this for a total of 12 rounds.

We ask that you do not talk with others during the experiment. If you have a question, please raise your hand and the experimenter will come to you. Please note that the computer screens have a protective cover that prevents others nearby from seeing your decisions and your earnings.

Practice Rounds

To help familiarize yourself with the interface and the tests, you will do one practice round of SAT questions and one practice round of GRE questions. Which test you practice first will be randomly determined. You'll be paid 5 tokens for each correct answer in the practice rounds.

Breaks

After rounds 4 and 8 we will take a short break and additional instructions will be read.

Your payments

You will be paid in a sealed envelope with the cash filled by a research assistant in another room. It will be identified with your subject ID. At the end of the experiment, the experimenter will leave the room and an unrelated third party will come in and pay each participant. This ensures that no other subject or the experimenter will know your earnings.

Summary

- You will answer test questions and earn tokens for correct answers; these are converted to cash at the end.
- Prior to each round, you choose to answer from either the SAT (generally easier) or the GRE for three different cases.
- The pay rates for correct answers are different for each of the three cases.
- It is in your best interest to choose the test you would like to take for each case if it is selected; there is an Earnings Calculator and your previous decisions are listed to help you decide.
- You have 4 minutes to answer 6 questions in each round.
- We'll repeat this for 12 rounds.

<u>Treatment-Specific Instructions Follow Below</u>

These instructions will be read aloud, and hard copies will not be distributed.

After Round 4

[AT] You have now completed 4 rounds. In the next block of 4 rounds we are introducing a new feature of the experiment. After the next 4 rounds are complete, participants will stand in two groups. Those participants who take GRE questions in 2 or more of the next 4 rounds will be asked to stand. They will remain standing for 8 seconds and then return to sitting. Those participants who did not take the GRE in 2 or more of the next 4 rounds will then be asked to stand for 8 seconds. They will return to sitting and the experiment will proceed.

[IO] You have now completed 4 rounds. In the next block of 4 rounds a new feature of the experiment is introduced in some sessions: subjects who take GRE questions in 2 or more rounds will be asked to stand. This will not occur in this session. All of your decisions will remain private throughout the experiment.

[Control] You have now completed 4 rounds. The next block of 4 rounds will proceed exactly as the first 4 rounds.

After Round 8

[In AT, subjects were asked to stand with their respective groups via a message on their computer screens. All subjects were shown the following: "At this point, participants who have answered questions from the GRE in 2 or more of the previous 4 rounds will stand". Below this, subjects were shown either "You have answered GRE questions in 2 or more of the previous 4 rounds. Please stand at this time" or "You have not answered questions from the GRE in 2 or more of the previous 4 rounds. Please remain seated at this time." After the appropriate subjects stood and returned to sitting, this was repeated for the complementary criterion of 3 or more SAT questions.]

[AT and IO] The final 4 rounds will be just like the first 4. Nobody will be asked to stand after these rounds no matter what their choices. After these 4 rounds, the experiment will conclude.

[Control] The final 4 rounds will be just like the previous rounds. After these 4 rounds, the experiment will conclude.

CHAPTER II
THE EFFECTS OF IMPORT COMPETITION ON WORKER HEALTH

This paper is collaborative work with Georg Schaur at the University of Tennessee.

Abstract

Occupational safety and health is an important determinant of workers' welfare. Our theory predicts that firms facing greater shut down risk reallocate resources to improve productivity at the expense of safety. Therefore, at firms facing the greatest shutdown risk due to import competition, safety conditions worsen following an import shock. We test this prediction with novel data on reported injuries at US manufacturers using growth in Chinese imports in the years 1996-2007 as a shock to competition. The data show that injury rates in the competing US industries increase over the short to medium run, particularly at smaller establishments. Back of the envelope calculations show that injury risk increases by 13 percent at the smallest establishments, costing workers the equivalent of a 1 to 2 percent reduction in annual wages.

2.1 Introduction

Health and injury outcomes are important to workers and firms. Estimates reveal that in 2007 US firms and workers saw as many as 9M occupational injuries and illnesses, 60,000 of which were fatal, that resulted in about 250B in costs to workers, firms and taxpayers (Leigh, 2011). Injury rates at US manufacturers are among the highest of any industry.[16] These same firms and workers also continue to see significant import competition from low cost markets, China in particular, which has important wage and employment effects. Labor standards, health and safety conditions are an important part of the employment contract, but have not been examined in the face of trade liberalizations or import competition. In this paper we ask if import competition from foreign markets affects injuries and worker health in US firms.

The link between import competition and worker injuries is supported by the intersection of evidence from the respective literatures. Import competition impacts firm survival (Pierce and Schott, 2013; Bloom et al., 2012, Bernard et al., 2006a, 2006b), labor markets (Autor et al., 2013), and firm investments in new technology (Ederington and McCalman, 2009). Literature on occupational safety and health (OSH) shows that injuries are determined by the relative priority the firm places on safety aside other goals like output (Zohar, 2000, 2002), technology upgrading and investments, and labor market conditions (Probst and Brubaker, 2001). Together, the bodies of literature suggest that foreign competition will impact occupational injuries and worker welfare by affecting the firms' incentives related to output and safety. Welfare evaluations based on wages alone miss this effect of trade on workers' welfare.

We combine plant-level panel data on injuries and illnesses at US manufacturers from the Occupational Safety and Health Administration (OSHA) with an industry and time varying measure of Chinese import competition, adapted from recent work by

[16] There were 4.6 injury and illness cases per 100 workers in manufacturing in 2012, compared to 3.8 in natural resources and mining and 3.7 in construction. (Bureau of Labor Statistics)

Autor, Dorn, and Hanson (2013). We apply differencing, fixed-effect and instrumental variable strategies to identify the causal effect of trade liberalization and supply-driven import shocks on injury and illness rates at competing domestic firms.

The estimates show that import competition has significant consequences for worker injuries. Import competition raises injury rates for all but the largest plants. The effect is greatest for the smallest plants. Looking at five-year log differences, the estimated elasticity of injury rates with respect to Chinese supply shocks is about 0.107 at the smallest decile of plants ($p<0.01$) and 0.085 at the median ($p<0.05$). Moving an industry from the 25th to the 75th percentile of total Chinese import growth increases injury rates by about 12% at the smallest decile of plants in the industry and 10% at the median.

Looking at worker welfare, estimates from the value of statistical life and injury literature show that the increases in injury risk resulting from Chinese import shocks are important in magnitude and are equivalent to wage decreases of approximately 0.4 – 1.6 percent.[17] For comparison, Arkolakis et al. (2012) discuss gains in real income due to trade liberalization of 1.4 percent. If variable trade costs are eliminated, Melitz and Redding (forthcoming) find a welfare effect of 17 percent. In total, we estimate that Chinese supply shocks in the US were responsible for between 62,000 and 90,000 injuries and illnesses annually during 2001-2007, about 7 percent of all cases in manufacturing, and a total loss to worker welfare between \$2.2 and \$9 billion each year. With this in mind, the effects of import competition on injuries are an important channel to consider for welfare overall and among workers at small plants in particular.

Differencing and fixed effect strategies mitigate the effect of unobserved plant, industry and geography specific characteristics. In addition, we tackle four identification problems. First, we consider long and short time differences to distinguish between short and long run effects of import competition and injuries. Second, underreporting is a recognized concern with self-reported injury data; workers hide or work through injuries rather than report them due to fear of being fired later on (Boone et al., 2011; Boone and van Ours, 2006). We discuss literature and incentives that mitigate the issue of misreporting and as a robustness check we estimate the model separately on the rates of injuries by severity and therefore susceptibility to misreporting. Third, it is difficult to identify exogenous trade shocks. In addition to instrumental variable techniques based on Autor et al., we identify import competition by adopting a liberalization in US trade policy towards China as a natural policy experiment according to Pierce and Schott (2013). Finally, an identification assumption we maintain throughout is that small plants, those with fewer employees, are less productive and therefore face a greater threat of insolvency from import competition. Findings in Holmes and Stevens (2014) provide an alternative explanation. Very small plants – and especially those located close to metropolitan areas – produce specialty goods and are therefore shielded from import competition. We perform robustness exercises that exclude plants from the sample that

[17] Estimates for the value of injuries vary across studies but generally range from 75% to 200% of yearly income. See Viscusi (1993) for a survey, and Hersch (1998) and Leeth and Ruser (2003) for later work.

are more likely to be specialized based on size, industry, and metropolitan location to show that specialized firms are not driving our results.

Our theory motivates why import competition raises injuries at small firms in the short run. At the firm, managers increase output by extracting greater effort from workers, but this increases the risk of injury. When making these decisions, firms consider the long term costs and the reputation effects of additional injuries. Because firms are judgment proof and future costs like higher insurance premiums, demand penalties, and productivity losses only matter if the firm is operational, the expected long run costs of injuries today are decreasing in the probability of the firm shutting down. Therefore, because import competition lowers the probability of firm survival, especially for less productive firms in the market, marginal firms respond by raising productivity at the cost of more injuries. Existing models are silent with respect to these short term firm-level adjustment effects. Comparing steady state equilibria before and after trade liberalizations, the least productive firms simply drop out.

Bloom et al. (2012) finds that the growth in Chinese imports in Europe significantly increased the average total factor productivity in competing domestic industries as a result of both low productivity-firm exit and improvements within surviving firms. Sources of observed intra-firm productivity gains include, in support of our theoretical approach, manager adjustments to improve factor productivity and increase firm survival (Bloom and van Reenen, 2007).[18] Also consistent with our intuition, Lazear et al. (2013) finds empirical evidence that firm's productivity gains during a recession are the result of getting more effort from their workers.[19] Our model shows that this happens because of a fall in the firm's survival prospects, and it generally explains how recessions, demand shocks, greater domestic competition, and other market shocks affect productivity and workplace safety.

Existing models' predictions regarding the long-run labor market effects of trade provide some useful insight for how import competition affects occupational health, an important component of the employment contract. Egger and Kreickemeier (2009) and Amiti and Davis (2011) consider fair wage mechanisms where firms share rents with their workers to keep them from shirking on the job. Helpman et al. (2011) considers a labor market with search frictions where firms and workers bargain over the rents of a successful match. Supposing firms share rents in the form of better health outcomes, trade liberalization will improve working conditions at large firms where profits rise and deteriorate conditions at smaller firms that are not exporting and face more competition on the domestic market. We take away the conclusion that, in the long-run, the effect of import competition on injuries is also heterogeneous across firms: safety standards

[18] See Syverson (2010) for a review of empirical studies on intra-firm productivity gains following trade liberalizations.

[19] Their model differs from ours in that the adjustment in their model is driven by workers who exert more effort to reduce the probability of job loss when unemployment increases as job loss is more costly. To the extent that import competition similarly deteriorates the value of being currently unemployed for a manufacturing worker – who might remain unemployed for longer or be forced to switch to a lower-wage occupation (as found in Ebenstein et al., 2014) – workers may increase their effort and be willing to accept worse safety outcomes following import competition.

deteriorate at low-productivity firms if rents fall as a result of import competition. Because we do not know how long it takes for long-run effects to materialize, our model complements these mechanisms to provide intuition for how imports affect injuries in the short run as firms adjust between long-run equilibria.

Our empirical exercise is most closely related to recent work by Autor et al. (2013), Pierce and Schott (2013) and Hummels et al. (2014). Our finding that greater import competition leads to higher injury rates in competing industries provides evidence of non-monetary welfare effects of import competition and supplements the results on adverse wage and employment effects.

Hummels et al. present theory that additional hours worked due to greater export opportunities leads to an increase in total injuries, consistent with literature on health and safety. They find empirical support for this scaling effect using Danish firm-level data. In contrast, our theory explains that firms facing import competition trade off safety for productivity and this affects injury rates. This implies that total injuries may increase even if number of workers and number of hours remain fixed. Therefore, we estimate the effect of import competition on injury rates and not the number of injuries.

We also contribute to the literature on the interrelation between international trade and labor standards (see Brown, Deardorff and Stern, 1996; Brown, 2007). Most studies focus on the developing world.[20] Our theory and evidence shows that occupational safety is an important determinant of welfare in developed economies in the face of import competition.

The remainder of the paper is organized as follows: in Section 2 we develop a theoretical model relating worker injury rates to firm survival; we describe our empirical strategy, data, and measurement in Section 3; Section 4 presents our primary empirical results; we discuss the robustness of our findings and alternative explanations in Section 5; Section 6 concludes. Tables and figures are provided in Appendix B.1.

2.2 Theoretical Model

With their backs against the wall, firms facing import competition and struggling for survival increase productivity and prolong their existence. Our theory shows that firms can improve productivity at the expense of their workers' health and safety. The partial equilibrium model provides intuition for what happens during the adjustment process at the most affected firms. These effects of import competition on injuries are absent in existing general equilibrium models that compare before and after liberalization steady state equilibria.

[20] In addition to wealth effects, long-run improvements in OSH can also be attributable to an increase in the relative price of labor and technological progress, both in production technology that alters workers exposure to risk and in safety technology. (Ruser and Butler, 2009) Trade exposure may accordingly affect standards also through technology diffusion/spillovers from exporters and changes in factor prices, but neither channel seems a strong explanation of the effects among US workers studied here.

2.2.1 Accidents and Injury Costs

Workers at the firm face a risk each period of being hurt on the job. Firms bear long-run costs for their workers' occupational injuries, including workers' compensation payments, medical expenses, fines, and associated legal fees. Firms shielded by workers' compensation insurance still face higher premiums if their plans are experience-rated. Firms further incur indirect costs that are more substantial than the direct costs and are not covered by insurance, including costly disruptions in production, damage to capital, hiring and training costs for replacement workers, demand penalties from bad publicity, and productivity losses due to lower employee morale[21] or employing a debilitated worker or inexperienced replacement. Therefore, an injury today results in costs today and the future. We assume that injury costs depreciate exponentially over time at a rate of d such that injury costs in period t are $c(t) = cd^t$.

Let δ be the exogenous shut-down probability in each period. The expected total cost of an injury today is then $\mathrm{E}[C] = \int_{t=0}^{\infty} e^{-\delta t} e^{-rt} c(t) dt$, where r is the firm's discount rate. It follows that, in expectation, the total cost over the firm's lifetime of an accident today is $\mathrm{E}[C] = \frac{c}{(\delta+r)d}$.[22]

The firm's decision maker is in effect "judgment proof". Future costs are only incurred if the firm is still operates.[23] An increase in the shutdown rate δ lowers the expected future cost of an injury today because the firm is less likely to be active and responsible for the cost.

2.2.2 Technology, Safety and Accident Rates

Plants employ labor L and elicit worker effort e to produce output $q = \varphi eL$, given exogenous productivity φ.

Taking wages w as given, the firm expends resources m to motivate worker effort, $e(w,m) = m^{\alpha} w^{\beta}$, where $0 < a, \beta < 1$. There is a tradeoff. The firm's endowed resources, such as management time, are limited and must be split between motivating effort and monitoring safety, s. To simplify notation let $m + s = 1$. The cost of increasing effort at the expense of safety then comes in the form of a higher injury probability $P(s) = (1-s)^{\gamma}$, where $s \in [0,1]$ (i.e. s and m are both nonnegative)

---

[21] In the fair-wage literature, worker effort depends on how they are treated by the firm in the form of the wages they are paid. Analogously, worker effort might depend on how they are treated by the firm with regards to the workplace environment and safety.

[22] This is derived in the following manner: Let C_t be the cost of an injury as of time t for all future time periods. For a short interval of τ periods, $C_t = \tau c + (1 - r\tau)(1 - \delta\tau)C_{t+1}$. As $C_{t+1} = dC_t$, we have $C_t = \frac{\tau c}{1-(1-r\tau)(1-\delta\tau)d}$. Taking the limit as $\tau \to 0$ yields $C_t = \frac{c}{(\delta+r)d}$. Alternatively, specifying discrete periods such that $E[C] = \sum_{t=0}^{\infty}(1-\delta)^t(1+r)^{-t}c(t)$ yields $C_t = \frac{c(1+r)}{1+r-d(1-\delta)}$ and the similar prediction $\frac{\partial E[C]}{\partial \delta} < 0$.

[23] In practice, this is implicitly true for many indirect costs – e.g. higher insurance premiums are inconsequential if the firm shuts down. For direct costs, incorporation laws limit many decision-makers from financial liability following a shutdown.

guarantees that the probability is bounded by 0 and 1. Let $\gamma \geq 1$ such that the marginal returns to safety are decreasing.

Now write $P(m) = (m)^{\gamma}$ for $m \in [0,1]$. Monitoring $m = 1 - s$ is the share of resources a firm allocates to motivating effort. Therefore a firm that motivates more effort improves the productivity per worker φe, but raises the expected number of worker injuries $m^{\gamma}L$.

2.2.3 Equilibrium Effort and Safety

The firm faces the inverse demand curve $p_M(q_M) = \theta q_M^{-\frac{1}{\sigma}}$, where $\theta > 0$ is a demand shifter (e.g. price index) and $\sigma = \frac{1}{1-\rho} > 1$ is the constant elasticity of demand. The firm maximizes total expected profits taking into account future expected costs of injuries. The firm chooses L and $m \in [0,1]$ to maximize profits taking into account total discounted injury costs, $\pi = \theta\varphi^{\rho}\left(m^{\alpha}w^{\beta}L\right)^{\rho} - wL - m^{\gamma}L\frac{c}{(\delta+r)d}$.

Normalizing wages to one[24 25], the firm's optimal share of resources to motivating effort is then $m^* = \left[\left(\frac{\alpha}{\gamma-\alpha}\right)\left(\frac{d(\delta+r)}{c}\right)\right]^{\frac{1}{\gamma}}$, where an interior solution requires $\frac{c}{d(\delta+r)} \geq \frac{\alpha}{\gamma-\alpha}$. The equilibrium worker injury rate is then

$$P(m^*) = \left(\frac{\alpha}{\gamma-\alpha}\right)\left(\frac{d(\delta+r)}{c}\right) \quad (1)$$

Notably, worker effort and injury rates under this specification are independent of the demand shifter θ and firm productivity φ; these come through as scaling effects in labor only.[26] Demand shocks, including changes in the price index, affect injuries only through their role in determining the shutdown rate δ. These solutions lead to the following proposition:[27]

---

[24] A few examples of considerable empirical support for wage stickiness are Blinder and Choi (1990), Campbell and Kamlani (1997), and recently Barattieri et al. (2014).

[25] We focus our attention on health and so we simplify the model by assuming that wages are unaffected by injury risk in the short run, either because workers have asymmetric information and do not observe changes in risk or because workers are paid their compensating variation in the event of an injury (this cost being subsumed in c) and they are therefore indifferent. The model's predictions are consistent under an alternative specification that allows for compensating wage differentials so long as other costs of injuries exist in the manner prescribed.

[26] The firm's optimal employment is $L^* = \left[\rho\theta\varphi^{\rho}\left(\frac{\alpha}{\gamma-\alpha}\right)^{\frac{\alpha\rho}{\gamma}}\left(\frac{\gamma-\alpha}{\gamma}\right)\left(\frac{d(\delta+r)}{c}\right)^{\frac{\alpha\rho}{\gamma}}\right]^{\frac{1}{1-\rho}}$.

[27] We use the functional specifications and closed-form solutions introduced above to present our results, but the predictions hold for general form functions such that e and P are both increasing and concave in m and C is decreasing in δ. These results are available on request.

Proposition 1: At an interior solution, an increase in the shutdown rate δ leads firms to reduce workplace safety in exchange for greater worker effort and accept an increase in the injury rate.

The intuition is straight forward. An increase in the shutdown rate lowers the expected future costs of an injury and the firm responds by pushing workers harder at the expense of higher injury rates.

Note that as the firm increases focus on effort as a result of a decrease in the survival probability, the labor productivity φe increases. This implies that estimated productivity parameters capture a tradeoff between safety and effort. This is a testable hypothesis given appropriate firm level data.

2.3 Empirical Strategy

In this section we describe our microdata on worker injuries and derive an empirical model based on our theory. We then combine our empirical model with existing measures of import competition and discuss the identification strategy.

2.3.1 Data

We employ plant-level data from the Occupational Safety and Health Administration on worker injuries and illnesses at US manufacturers during 1996-2007. Each year, OSHA and participating state regulators sample approximately 50,000 manufacturing establishments, or plants, from the universe of private sector plants with at least 40 employees per BLS records (GAO 2009).[28] The data are an unbalanced panel of 473,014 plant-year observations that covers 53% of all US manufacturing workers during this time.

Each plant self-reports the total number of job-related injury and illness cases recorded in the year, the number of cases that required time away or restricted work or transfer to other duties, its primary industry at the SIC4 level, the average number of all workers during the year, and a measure of equivalent full-time employment for the year imputed from total hours worked. We construct the injury rate P as the OSHA-defined Total Case Rate (TCR). For each establishment, TCR is computed as the number of reported injuries in the year divided by the imputed measure of employment, scaled up by 100. An increase in TCR therefore measures an increase in the rate of injuries normalized to 100 workers.[29] Potential reporting errors in both hours worked and recorded injuries contribute to mis-measurement of TCR.[30] We drop outliers with an

[28] Per the Quarterly Census of Employment and Wages, less than 0.1% of manufacturing plants in the US are government-owned and employ less than 0.5% of all manufacturing workers.

[29] It is calculated with the imputed employment measure so that P is a measure of injury risk per unit of time worked, and therefore it is not directly affected by lengthening or shortening work schedules.

[30] We also drop the eight observations with employment above 25000 as apparent reporting errors – none report employment above 10000 in any preceding or following year. Our results are robust to their inclusion.

imputed TCR greater than 60, twice the highest industry average and eight times the overall average. The correlation coefficient between average industry-year TCRs constructed with our sample and the population estimates reported by the BLS is 0.812 when these 0.5% of observations are excluded and 0.086 when they are not. The results are qualitatively similar when we include these outliers and most remain statistically significant.[31] Summary statistics for these data are given in Panel A of Table B.1. More detailed information on the data is provided in Appendix B.2.

To measure import competition we use trade data from two sources. First, we use data on yearly imports from China at the SIC4 industry level provided by Autor et al. (2013) and described in detail there.[32] Briefly, they concord product-level import data from the UN Comtrade Database to the industries that manufacture like products. We make use of their data on Chinese exports to the US and to an aggregate of eight other high-income nations.[33]

For our policy experiment based on Pierce and Schott (2013) we employ US tariff data from Feenstra et al. (2002).[34] These data have both the Normal Trade Relations (NTR) tariff rates and the higher Column 2 tariff rates for eight-digit Harmonized Tariff System products. Following Pierce and Schott, we define the difference in the two rates as the NTR gap and construct a SIC4 industry-level measure by averaging the tariff differentials for all goods produced by the industry.[35] The mean NTR gap for our sample is about 0.31 and the standard deviation is 0.16. Summary statistics for the trade data are provided in Panel B of Table B.1.

2.3.2 Specification

Let i denote plants, j denote industries, and t denote years. Our reduced form model, derived in Appendix B.3 from equation (1), relates changes in injury rates over s-years to the change in import competition during the period, base-year employment, and industry and year-specific effects:

$$(2)\quad \ln\left(P_{ij}\right)^{t:t+s} = \beta_0 + \beta_1 \ln\left(M_{uc,j}\right)^{t:t+s} + \beta_2 \ln\left(M_{uc,j}\right)^{t:t+s} \times \ln L_{ijt} + \beta_3 \ln\left(L_{ijt}\right)$$

[31] For example, we estimate the mean import shock over five years increased injury rates by 12 ppts at the smallest decile of plants (p<0.05) and by 9 ppts at the median (p<0.05) when using the full sample, as compared to 13 ppts (p<0.01) and 10 ppts (p<0.05) when using our main (trimmed) sample. These results are available upon request.

[32] This bilateral industry-level trade data is publicly provided on David Dorn's website, for which we are grateful.

[33] They are Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

[34] We are grateful to Feenstra et al. (2002) for making the US tariff data publicly available through John Romalis' website.

[35] We use the matching procedure in Pierce and Schott (2012) to concord HTS8 product tariffs to SIC4 industries. We are grateful to them for making their concordance available through Peter Schott's website. We first use the concordance from the Bureau of Economic Analysis and match 63% of products. Next, we concord an HTS8 product to a SIC4 if all other products in its seven-digit "family" match to the same SIC4. We repeat the process using six-digit families and so forth, and this matches an additional 29% of products. Last, in the ordered set of HTS8 codes, we assign a product to a SIC4 if the industry preceding the gap is the same as the industry following the gap, which matches an additional 3% of products.

$$+\tau_j + \theta_t + \varepsilon_{ijt},$$

where the change in import competition is identified as the log growth in US imports from China in industry j, $\ln(M_{uc,j})^{t:t+s}$. The coefficients of interest β_1 and β_2 capture the elasticities of the shutdown rate with respect to import competition and import competition interacted with plant employment, our proxy for productivity differences. Employment in the base year is assumed to be exogenous of the ensuing growth in import competition.

Existing literature and our theory explain the intuition for the model. Import competition raises shut-down rates (Bernard et al. 2006a; Pierce and Schott, 2013), especially at the least productive plants (Bernard et al. 2006b) with low levels of employment. Therefore, based on our theory, we expect $\beta_1 > 0$, $\beta_2 < 0$, and, $\beta_1 + \beta_2 \ln(L_{ijt}) > 0$. An increase in import competition raises shut-down rates and thus increases injury rates, especially at small firms that are most affected by import competition.

Measuring import competition using the growth in Chinese imports by industry provides rich identifying variation across industries and time, but this measure is correlated with demand shocks. The coefficient estimates from an OLS regression of equation (2) will be biased if demand shocks also affect injury rates, and there are reasons to believe they do. Our theory explains that a demand shock will decrease injury rates if it improves the firm's survival prospects, while Hummels et al. (2015) shows that a demand shock will increase total injuries by making output more valuable. Elsewhere, studies link worker injuries to the business cycle, with similarly mixed findings on the direction of the effects.[36] To control for this potential bias, we employ an instrumental variables strategy to exogenously identify import growth driven by supply shocks. Notably, the OLS estimates for $\beta_1 + \beta_2 \ln(L)$ in equation (2) are generally smaller than the IV estimates, suggesting the endogeneity would bias against our results.

Following Autor et al (2013), we instrument for Chinese import growth in the US with Chinese import growth in a set of other OECD countries and estimate equation (2) using two-stage least squares (2SLS). In the first stage, we regress the two endogenous variables $\Delta \ln(M_{uc,j})^{t:t+s}$ and $\Delta \ln(M_{uc,j})^{t:t+s} \times \ln(L_{ijt})$ on the instruments $\Delta \ln(M_{oc,j})^{t:t+s}$ and $\Delta \ln(M_{oc,j})^{t:t+s} \times \ln(L_{ijt})$, along with base-year employment and any fixed effects included in equation (2), using OLS.

2.3.3 Identification

The main identifying assumption for the IVs is that the correlated growth across the markets is driven by changes in trade costs and China's productivity. Under these

[36] For example, Ruhm (2000) finds evidence that worker injuries are pro-cyclical, while Brooker et al. (1997) and Boone et al. (2011) find the opposite.

assumptions, the estimates for β_1 and β_2 from 2SLS on equation (2) are consistently identified.

We estimate the model in differences and include a rigorous set of fixed effects to ensure that no systematic information is moved into the error term that is correlated with our instrumental variables or results in endogeneity concerns for the variables that we treat as exogenous in our model. Differencing accounts for unobserved heterogeneity in plant characteristics that potentially affect injury rates, such as injury costs and productivity. Industry and base-year fixed effects allow for industry-specific trends in injury rates across the full panel and for shocks that affect all establishments in a given period.

We estimate the model for $s = [1,6]$. We regard five-year differences as our preferred specification, consistent with related work (e.g. Bloom et al., 2012; Amiti and Davis, 2011; Bernard et al., 2006a, 2006b), but we compare the estimates across intervals to speak to how these effects evolve over time. For the shortest intervals, the instruments will be weakened to the degree supply shocks are asynchronous across high income countries. The instrument for longer differences is more forgiving of smaller delays in the shocks across markets, but we are left with fewer observations, due to both the longitudinal limits of the panel and the increasing likelihood of plants shutting down or switching primary industries over longer time horizons.

Autor et al. list three threats to the IV validity that are also relevant for our analysis. One concern is that product demand shocks may be correlated across the US and other high-income countries, and therefore the first-stage IV results are not capturing the supply-side effect only. To check this, they apply a gravity strategy that isolates the import growth attributable to changes in trade costs and China's productivity. They find similar results and conclude that correlated demand shocks are not driving their estimates for employment and wage effects using the IV. We proceed with the same assumption for our identification of injury effects. Second, the correlation may be driven by adverse productivity shocks to US producers that are being supplanted by Chinese imports in US and other OECD markets, but China's explosive productivity growth – about 8% over this period (Brandt et al., 2012), more than double that in the US – make it the more likely driver of its export growth across the markets. Lastly, Chinese imports may rise in both the US and other OECD countries following common technology shocks that adversely affect labor-intensive industries. However, Chinese export growth to the US greatly outpaced that of other low and middle-income nations which may similarly exploit an adverse shock, suggesting that Chinese productivity shocks and falling trade costs for Chinese imports instead drove the growth.

To alleviate remaining endogeneity concerns associated with the instrumental variable technique outlined above, we also employ an alternative identification strategy based on a natural policy experiment. The US permanently reduced import tariffs on Chinese goods in 2001 when it acceded to the WTO. Following Pierce and Schott (2013), we use the industry-specific tariff reductions to proxy for the changes in import competition associated with the policy shock, and we identify the effects on injuries using a difference-in-difference estimator.

The second identification concern we consider is potential measurement error in the dependent variable. If the measurement error is not random in our estimating equations, it will bias our coefficient estimates. For example, it may be the case that workers at dying plants are more like to report injuries. In that case, part of β_2 captures misreporting. It is not clear that this is true, however. Even at dying plants, workers want to preserve a good reputation as they likely have to hunt for new jobs in the future. Nevertheless we will perform a robustness check using the rate of fatalities and more serious nonfatal injuries and illnesses only – those that required time away from work, restricted work or transfer – under the premise that these incidences cannot be as easily hidden and are less susceptible to strategic misreporting.

Third, a key connection between our empirical specification and theory is the interaction of import competition with plant size. We assume plant size proxies for differences in productivity, which we do not observe in the data. This assumption is consistent with our theory and with new trade models such as Melitz (2003), where the least productive firms are the smallest. Notably, using labor as a proxy is advantageous for regulators and policymakers who are looking to identify the plants most affected by import competition: plant size is directly observed and more widely available than is data on plant-level total factor productivity. Findings in Holmes and Stevens (2014) suggest an alternative explanation. They show that the very smallest firms and firms closely located to metropolitan areas are more specialized and are therefore shielded from import competition. We will perform several robustness checks that show that this is not affecting our results.

Next, our identification is potentially biased because the yearly samples only target plants with 40 or more workers. Our sample of differences is therefore biased towards excluding plants that shrank during the period as some become ineligible for selection. Optimally, we would observe a balanced sample of all surviving plants, but that's not possible with our data source. This threatens our identification if the injury effects of import competition are different at growing plants such that our estimates are not representative of the full population. Further, this explains heterogeneity in the estimates as the selection bias is greatest for smaller plants already near the margin. Our theory suggests that this will bias our estimates towards zero if plants that grow tend to be those less affected by import competition, but we still address the issue with robustness checks that limit the sample to larger plants in the base year that are unlikely to be excluded with moderate falls in employment.

Finally, we consider the potential bias in our main specification arising from heterogeneity in establishment discount rates. This will bias our estimates towards zero. The intuition is that plants with high discount rates already place a low value on the future and changes in shut-down rates are less relevant. Notably, this bias would seem to work against our predictions that the effects are greatest at small plants given evidence in the literature that they are the more fiscally constrained and have higher discount rates (Beck et al., 2005). Nevertheless, we show in Appendix B.3 that specifying the model in levels allows us to separate discount rates from shutdown rates such that they difference out. Therefore as an alternative specification we estimate

$$(3) \qquad P_{ij}^{t:t+s} = \beta_0 + \beta_1 \ln\left(M_{uc,j}\right)^{t:t+s} + \beta_2 \ln\left(M_{uc,j}\right)^{t:t+s} \times \ln L_{ijt} + \beta_3 \ln\left(L_{ijt}\right)$$

$$+\theta_t + \varepsilon_{ijt}.$$

Estimating this model in levels has the advantage that we can include zero injury observations which we have to drop in the log-linear model. The trade-off is that, as we show in the appendix, in this case the estimates are affected by heterogeneity in plant-specific injury costs. Therefore, we prefer model (2) over model (3) because according to our theory we can account for plant specific injury costs.

2.4 Results

2.4.1 Baseline Specification

Table B.2 reports our main results for the effect of Chinese import shocks on injury rates at competing US manufacturers. The six columns give the second-stage estimates from the 2SLS regression of equation (2) for one- to six-year differences. The IVs fail a weak instrument test in the first-stage regressions for one-year differences, but are significantly correlated for all longer intervals.[37] For all intervals, we find that Chinese import shocks increased injury rates at the smallest US plants in the industry ($\beta_1 > 0$), and the effects were decreasing across plants in size ($\beta_2 < 0$). Under the identifying assumptions for our 2SLS strategy, these effects are the result of Chinese productivity gains and falling trade costs. The estimates for both coefficients of interest are statistically significant for all specifications save that for one-year differences.

We estimate the marginal effect $\beta_1 + \beta_2 \times \ln(L)$ at approximately the smallest decile (L =40), median (L =100), and largest decile (L =400) of plant employment size in the sample. These estimates are reported in Table B.2 for each difference specification. For five-year differences, we estimate that the elasticity of injury growth with respect to import shocks is about 0.107 at the smallest decile of plants (p<0.01) and 0.085 at the median (p<0.05). The effects are positive and significant at plants with less than 250 employees and weakly positive at all but the very largest plants in the sample.

These estimates imply that moving an industry from the first to third quartile of total Chinese import growth over five years increases injury rates by about 12 percentage points at the smallest decile of plants and 10 percentage points at the median. Alternatively, scaling the estimates by the interquartile range for the distribution of supply-driven import growth yields magnitudes half these sizes, based on estimates in Autor et al. (2013) that show supply shocks comprise 48% of total growth.

[37] For differences of two or more years, the robust F statistics for joint significance of the instruments considerably exceed 10, the benchmark for weak instruments suggested by Staiger and Stock (1997). That the F statistics for one-year differences fall below this value suggests the instruments in that specification are weak and the second-stage estimates may be inconsistent as a result. The test statistics for weak instruments are suppressed in Tables 5 and 8 as they are quite similar to those reported in Table 2 (the first stage equations are the same).

Looking across time intervals, we find that the estimated elasticities at the smallest decile of plants are small and statistically insignificant for two to three year intervals and become larger and significant after four to five years and appear to plateau there.[38] The estimates for the median plants follow a similar trend, while those for the very largest plants are not significantly different from zero for any specification. The results demonstrate that, at smaller plants in affected industries, the injury effects of import competition are persistent over four to six years.

2.4.2 Magnitudes

Our estimates show that Chinese import shocks worsen injury rates at smaller plants in competing US industries. We use a back-of-the-envelope exercise to derive estimates for the magnitude of these effects, recognizing that they are only part of the story. Other channels like trade-induced labor reallocations and industry and local labor market spillovers likely contribute to the general equilibrium effect but lie beyond our analysis. Nonetheless, our estimates give an idea of the scale of these non-wage effects and their relative importance for worker welfare, providing a point of comparison with existing studies on the wage effects of trade.

We derive these estimates using the 5-year difference model to predict log growth in plant injury rates under both the realized growth in Chinese imports and the counterfactual of no import growth over the period, holding the other parameters of the model constant. From the employment-weighted average of the estimates, we back out the predicted yearly injury rates for all US manufacturing workers under the two scenarios.[39] The difference in a given year, scaled by US manufacturing employment, gives our estimate for the number of injuries attributable to Chinese import competition through these within-plant effects.

These back-of-the-envelope estimates are presented in Table B.3. Our estimates attribute 7.4% of all US manufacturing worker injuries and illnesses to Chinese import shocks, which increased annual morbidity by 0.51 cases per 100 workers on average. Figure B.1 shows the predicted injury rates for 2007 across plant size bins and illustrates the heterogeneity in these effects. The increased morbidity at plants with 50 or fewer workers was 0.98 cases per 100 workers (13% increase), while at plants with 1000 or more workers it was only 0.12 (2% increase). In aggregate, import shocks were responsible for between 62,000 and 90,000 injuries, where the lower bound assumes injury rates were unchanged at the small plants that lie outside our sample and the upper bound applies estimates for the smallest plants in our sample to out-of-sample plants.

The value of statistical life literature estimates that nonfatal occupational injury and illness cost worker welfare between \$35,000 and \$100,000 in 2004 USD (Viscusi,

---

[38] It seems unlikely that the trends across time intervals are driven by changes in the composition of the estimating sample related to plants dropping out. The results look similar when limiting the estimating sample for each specification to only plants that survive at least five years past the base year (the criterion for inclusion in the five-year difference sample). These results are available upon request.

[39] In aggregating, we first create an industry average using establishment-level effects and weighting each by plant employment, reported in the ODI data, and then we aggregate across industries and weight each by the manufacturing employment shares reported by the BLS.

1993; Hersch, 1998; Leeth and Ruser. 2003).[40] Based on those values, the estimated within-plant injury effects of Chinese imports cost US manufacturing workers the equivalent of a 0.4%-1.6% reduction in wages and between \$2.2 billion and \$9 billion in aggregate each year.[41] This is in addition to direct and indirect costs borne by firms and the government.

2.4.3 Identifying Import Shocks Using Natural Policy Experiment

The main threat to identification is that we are not exogenously identifying supply shocks with our IV and our results are instead driven by unobserved factors related to both the import measure and the number of injuries, such as demand shocks. We address this concern by testing the model's predictions using a second strategy for identifying import competition based on the policy experiment applied by Pierce and Schott (2013). The measure is grounded in trade policy and relies on a different set of exogeneity assumptions.

The US granted China Permanent Normal Trade Relations (PNTR) in 2001 when it acceded to the WTO. Previously, the US had granted Chinese imports the lower NTR tariff rates annually based on a vote of Congress, and the votes were often heavily contested.[42] The policy change eliminated uncertainty over future trade costs and encouraged Chinese firms to incur the fixed cost to export to the US: Handley and Limão (2013) finds that the reduction in tariff uncertainty explains 22-30% of Chinese exports following its joining the WTO. The ensuing import growth is especially large for products where the difference in the NTR and non-NTR tariffs is large, and so the shock hits some industries harder than others based on the import tariffs protecting their products.

Following Pierce and Schott, we use the NTR gap for each industry – the average difference in the two tariff rates for its goods – to quantify the import shock that hit each manufacturing industry in 2001. We use the following difference-in-difference (DiD) estimator to compare within-plant changes in injury rates before and after PNTR in 2001 (first difference) and across industries with higher and lower NTR gaps (second difference):

$$\text{(4)} \qquad \ln\left(P_{ij}\right)^{t:t+s} = \beta_0 + \beta_1 \bar{\tau}_j \times 1_{t\geq 2001} + \beta_2 \bar{\tau}_j \times 1_{t\geq 2001} \times \ln\left(L_{ijt}\right) + \gamma \ln\left(L_{ijt}\right)$$

[40] These values do not account for the documented heterogeneity in VSL related to worker income, age, gender, and other characteristics (Evans and Smith, 2006, 2010; Knieser et al., 2010). We do not observe these in the data. The total welfare effects will differ if these injuries are disproportionately concentrated among some groups.

[41] These estimates exclude fatalities, which we do not identify. Supposing import competition explained 7% of the 392 worker fatalities in the US manufacturing industry (Census of Fatal Occupational Injuries) would imply an additional \$190 million in welfare loss based on a benchmark VSL estimates of \$7 million (Kniesner et al., 2010).

[42] The House of Representatives voted *against* temporary NTR status for China in 1990, 1991 and 1992, following the Tiananmen Square incident, but the Senate failed to act on those votes and the status was granted in each year.

$$+\theta_t + \theta_j + \varepsilon_{ijt}$$

where $\bar{\tau}_j$ is the (time invariant) NTR gap for industry j. We separately estimate the model for changes over $s = [1,5]$ years using for each the sample of observed changes on either side of PNTR, $t = \{2001 - s, 2001\}$. Industry fixed effects θ_j account for industry-specific trends in injury rates that span both pre- and post-PNTR periods, and base-year fixed effects θ_t control for the common trend across industries within each time period. The results for equation (4) are given in Table B.4. We find that permanently reducing Chinese import tariffs significantly affected injury rates at competing US manufacturers. The results support our main findings: the import shock increases injury rates at the smallest plants in the industry ($\beta_1 > 0$) and the effects are diminishing in plant size ($\beta_2 < 0$). The coefficient estimates are statistically significant for intervals of three or more years. We calculate the marginal effect of the NTR Gap on injury rate growth at the smallest decile, median, and largest decile of plant employment and report those estimates for each specification in Table B.4. The trend and the timing are consistent with outside evidence on labor adjustments following a trade shock. Pierce and Schott (2013) estimate that liberalizing US tariffs on Chinese goods in 2001 caused a cumulative decline in US manufacturing employment growth over the following $s = [1,6]$ years of (in percentage points): 3, 6, 11, 13, 15, 16.

For interpretation, we scale the marginal effects by the mean NTR gap for the sample, about 0.28. In the mean-affected industry, we estimate that the liberalization increases injury rates over the five years that follow by 5.8 percentage points at the smallest decile of plants ($p<0.01$). The effect was weakly positive at the median plant and weakly negative at the largest decile. These results provide direct evidence of changes in tariff policy affecting worker health at competing domestic plants, and are, to our knowledge, the first results to do so.

Several identification concerns related to these estimates are worth discussing. Our main identifying assumption here is that the NTR gap is exogenous to market conditions near PNTR in 2001 that may also correlate with injury rates. This assumption is grounded in the fact that the tariff rates were set long before this period and there were few changes in the decade leading up to PNTR and none to non-NTR rates, which explain 89% of the variation in the measure. In further support, Pierce and Schott show that their results are robust to using lagged tariff rates to construct the NTR gap. We follow their main specification and use 1999 tariff rates.

A related threat to identification arises because PNTR coincides with the business cycle peak in 2001. Pierce and Schott deal with the issue by comparing post-2001 changes to those that followed the previous peak in 1990, but we cannot do this because our data does not begin until 1996. It seems unlikely that our results are biased by correlation between the NTR gap and the injury effects of the business cycle across industries for a few reasons. First, industry NTR gaps are plausibly exogenous to changes in the macroeconomy during this period for reasons discussed above. Second, the literature finds that reported injury rates are procyclical, so any bias would seem to work against our finding an increase in injury rates following PNTR. Further, it is not evident that this bias explains intra-industry heterogeneity across plants.

As compared to our import growth measure, we have less intertemporal variation when identifying the model using the NTR gap. It only explains within-industry changes in injury growth before and after PNTR in 2001, and not changes within either period (e.g. 1996:1998 vs. 1998:2000) or in periods spanning the change (e.g. 1999:2003), which limits our estimating sample. Identification with the NTR gap holds appeal in that it directly links trade policy to worker health effects.

2.4.4 Robustness

We run robustness checks to address several remaining threats to identification in our main empirical results. Out of concern for potential bias in our primary dependent variable due to misreporting, we estimate equation (2) using the rate of fatal and serious nonfatal injuries and illnesses only. These injuries are more difficult to hide and less susceptible to strategic reporting by workers, and therefore the results are more robust to this potential bias in the left-hand variable of the equations. The results, shown in Table B.5, are consistent with those from our main specification. We conclude that import competition affects worker injuries and not reporting behavior only.

We next address our identifying assumption that size captures productivity differences across plants in an industry. This assumption connects our heterogeneous empirical specification to the theory and underpins our interpretation of the regression results. Holmes and Stevens (2014) provide an alternative explanation for why import competition affects small plants differently: producers of specialty and custom goods are likely smaller and better insulated from Chinese import competition. For our empirical results, this suggests that differential effects among these specialized plants might be driving the heterogeneity across plant size.

Holmes and Stevens discuss plants with less than 20 employees, which suggests we are largely shielded from this issue as our sample is intended to cover only plants with at least 40. Less than 4% of observations in our data report 20 or fewer employees. We also limit issues relating to heterogeneity within industries by defining them according to SIC codes. Holmes and Stevens point out that the newer North American Industrial Classification System (NAICS) redefines several manufacturing industries in a way that increases the presence of custom and specialty plants relative SIC-defined industries. For example, some small custom cabinet-makers are classified in the wood kitchen cabinet industry under NAICS, but these are classified as retail under SIC and excluded from our sample.

Still, we run several robustness checks that are presumed to limit the presence of specialty plants in the estimating sample based on Holmes and Stevens findings. These results are presented in Table B.6. In the first column, we exclude plants with less than 100 workers in the base year. In the second column, we exclude plants in industries with the highest shares of specialty plants based on estimates from Holmes and Stevens. Lastly, Holmes and Stevens find that specialty plants tend to be located in urban areas, and so we exclude plants in more densely populated five-digit ZIP codes using US Gazetteer data and present the results in the third column. For each check, the results are consistent with our main specification. Importantly, the heterogeneity across plant size persists within each of these subsamples, supporting our claim that it is driven by productivity differences and not specialization.[43]

The estimates in column 1 of Table B.6 also demonstrate that our results are not driven by bias due to the sample not covering plants where employment shrinks below the selection minimum. Less than 2% of plants with at least 100 employees in the base year report below 50 employees five years later, suggesting few among this group were operating but ineligible for selection because their employment was below 40. Our results using this estimating sample that is more robust to omitting shrinking plants are qualitatively similar to our main results.

We lastly estimate the model using the level-difference specification for injury rates given in equation (3). Presented in Table B.7, the results support our main finding that injury rates went up at small plants hit with import competition. The trends across plant size and interval durations generally resemble those in our main results, with the exception that the estimated effects at larger plants are positive and statistically significant for longer differences in the levels specification.

2.5 Conclusion

We find empirical evidence that the growth in Chinese imports in the US in the late 1990s and early 2000s significantly increased worker injury and illness rates in the competing industries in the short to medium run. We find that one significant contributor to these effects was the change in US trade policy in 2001, when import tariffs on Chinese goods were permanently reduced. The heterogeneous within-industry effects were greatest for small plants. Back-of-the-envelope estimates show that Chinese import shocks accounted for 7.4% of worker injuries and illnesses in US manufacturing during 2001-2007. Injury rates rose by 13% at the smallest plants, costing worker welfare the equivalent of a 1-2% reduction in annual wages.

Our theoretical model predicts that firms respond to greater shutdown risk by allocating resources towards productivity at the expense of safety. Based on the theory, we hypothesize that import competition will deteriorate worker health outcomes in the short run at marginal firms at risk of being pushed from the market during the transition to the new open economy equilibrium. This mechanism – substituting resources away from safety and into productivity – allows firms to prolong their life in the market and slow down the adjustment process following a trade liberalization, but does so at the cost of its workers' health.

Our theory and interpretation of the estimation results is consistent with recent evidence from Lazear et al. (2013) that finds increases in worker effort during a recession drive firm productivity gains. In our data we cannot identify changes in effort or other

[43] Notably, we find that the 5-year injury effects among plants with 40 or fewer workers are *lower* at the smallest plants: $\beta_1 = -1.400^{*}$ and $\beta_2 = 0.429^{**}$. This is consistent with the story that the very smallest plants are specialized and insulated from import competition, but the results should be taken with the caveat that we do not observe many of these plants in the data (N=3212) and their reported employment falls below the intended minimum for the sample.

factors that might affect injury rates, such as changes in technology or shifts towards employing more temporary or part-time workers. Future research matching more comprehensive firm level data with detailed worker level data will be able to provide further evidence related to the underlying mechanisms for these effects.

Appendix B.1 Tables and Figures

Table B.1. Summary Statistics of US Manufacturing Plant-Level Data, 1996-2007

PANEL A: Injury and Employment Data (Full Sample)			
	Obs.	**Mean**	**Std. Dev.**
Observations per Plant	172188	3.289	2.622
Years Spanned in Data	172188	4.034	4.088
Observations/Year Spanned	172188	0.793	0.246
Employment	526931	185.64	432.36
Injury Rate (TCR)	473014	10.38	9.50
Injury Rate (TCR) | Non-zero	420046	11.68	9.30
Log Injury Rate (TCR)	420046	2.125	0.889
1-year log-change	192851	-0.124	0.646
2-year log-change	152186	-0.180	0.715
3-year log-change	137015	-0.219	0.764
4-year log-change	117785	-0.281	0.796
5-year log-change	93509	-0.347	0.816
6-year log-change	72673	-0.410	0.836
Severe Injury Rate (DART)	474637	5.475	5.988
Non-Severe Injury Rate	475549	5.071	6.350

Notes: Employment imputed from total hours worked as equivalent full-time workers, and excludes observations with imputed plant employment of greater than 25000. Injury rate statistics exclude observations with imputed injury rates above 60 per 100 workers in a year. Observation intensity is an upper bound as time spanned in data is a lower bound of the spell in eligible pool for selection.

PANEL B: Import Competition Measures			
	Obs.	**Mean**	**Std. Dev.**
US Imports from China (000s)	560784	760.0	1868.7
1996 Levels (000s)	59568	235.3	512.3
2007 Levels (000s)	43965	1587.4	3151.9
Log US Imports from China	558746	11.06	3.02
1-year log-change	245123	0.264	0.750
2-year log-change	199097	0.527	0.990
3-year log-change	181189	0.803	1.111
4-year log-change	155399	1.062	1.196
5-year log-change	123452	1.263	1.297
6-year log-change	95514	1.558	1.406
OTH Imports from China (000s)	560558	471.0	978.3
NTR Gap	164464	0.313	0.161

Notes: OTH refers to the set of other OECD countries used to instrument for imports in the US, following Autor et al. (2013). The distribution of NTR gap is across plants only as it is time invariant, all others are across plant-years.

Table B.2. Injury Effects of Realized Chinese Import Growth (Second Stage IV Results)

Dependent Variable: Log Growth in Injury Rates at Plant $\ln(TCR)^{t:t+s}$
Import Competition Measure: Log Growth in Chinese Imports in Industry $\ln(M_{uc})^{t:t+s}$

Interval Duration (s)	**1 Year**	**2 Years**	**3 Years**	**4 Years**	**5 Years**	**6 Years**
Import Competition	0.357	0.192**	0.114*	0.267***	0.193***	0.164***
	(0.234)	(0.086)	(0.061)	(0.063)	(0.064)	(0.061)
× Employment	-0.031	-0.037**	-0.023**	-0.043***	-0.024**	-0.022**
	(0.024)	(0.014)	(0.010)	(0.009)	(0.009)	(0.010)
Employment	0.026***	0.017**	-0.004	0.018	-0.010	-0.015
	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.016)
Estimated Marginal Effects by Plant Employment Size						
40 Employees	0.245	0.057	0.031	0.107***	0.107***	0.083**
	(0.163)	(0.042)	(0.035)	(0.037)	(0.039)	(0.036)
100 Employees	0.217	0.024	0.010	0.067**	0.085**	0.063*
	(0.148)	(0.036)	(0.031)	(0.035)	(0.035)	(0.033)
400 Employees	0.174	-0.027	-0.021	0.007	0.052	0.033
	(0.130)	(0.035)	(0.031)	(0.031)	(0.033)	(0.033)
F-Statistics for Joint Significance of Instruments in First-Stage Regressions						
$\ln(M_{uc,j})^{t:t+s}$	3.88	42.62	70.91	110.55	58.64	58.94
$\ln(M_{uc,j})^{t:t+s} \times \ln L_{ijt}$	9.51	47.97	90.45	206.19	126.54	135.33
Base Years (t)	1996-2006	1996-2005	1996-2004	1996-2003	1996-2002	1996-2001
Number of Plant Clusters	62780	49981	46523	44224	38787	32204
Observations	185416	142312	126210	107639	84552	64939

Notes: All regressions include a constant and year and industry (SIC4) fixed effects. Robust standard errors are clustered by plant and reported in parentheses. Employment is the natural log of the number of employees in the base year. Regressions include all plants with observations that span the respective interval, including those in states that drop out of the sample in later years and in "catch-all" SIC4 industry categories, xxx9. Omitting either or both groups does not significantly affect the results. The employment levels 40, 100, and 400 are approximately the 10th, 50th, and 90th percentiles for the five-year difference estimating sample. The distributions for the samples in the other regressions are similar. F-statistics for weak instrument tests are adjusted for robust standard errors. * p<0.10, ** p<0.05, *** p<0.01

Table B.3. Predicted Injury Effects for Import Growth Specification Under Alternative Trade Environments

Year	Manufacturing Emp. (000s)	Sample Injury Rate	Predictions | Realized Chinese Import Growth		Predictions | No Chinese Import Growth		Difference in Predictions		Share Cases Attributable to Import Growth (ppts)
			Injury Rate	No. Injuries	Injury Rate	No. Injuries	Injury Rate	No. Injuries	
2001	16641	6.74	6.51	1083329	6.07	1010109	0.43	73220	6.6
2002	15259	7.31	7.36	1123062	6.87	1048293	0.49	74769	6.7
2003	14509	7.41	8.57	1243421	7.90	1146211	0.67	97210	7.8
2004	14315	6.21	8.62	1233953	7.96	1139474	0.66	94479	7.7
2005	14227	5.55	6.15	874961	5.75	818053	0.40	56908	6.5
2006	14155	5.23	5.47	774279	5.03	711997	0.45	62282	8.2
2007	13879	5.46	5.70	791103	5.22	724484	0.48	66619	8.4
Mean	**14712**	**6.31**	**6.92**	**1017730**	**6.41**	**942660**	**0.51**	**75070**	**7.4**

Notes: Injury rates are calculated as the Total Case Rate (TCR), the number of reported injuries and illnesses in a year per 100 equivalent fulltime workers. Estimates are from the specification of equation 2 for five-year intervals. Average rates across years are weighted by yearly manufacturing employment. Injury rates are estimated using a data sample restricted to firms with at least 40 employees, but the estimates here are applied to all manufacturing employment. About 17% of manufacturing workers over this period are employed at plants of less than 40 workers, which lie outside our sample. The estimates for the number of injuries should therefore be multiplied by 0.83 to obtain the estimated effects for workers at plants covered by our sample only and the number of injuries attributable to Chinese import growth under the assumption that injury rates at out-of-sample plants are unaffected.

Table B.4: Injury Effects of Permanent Import Tariff Reductions Granted China in 2001

Dependent Variable: Log Growth in Injury Rates at Plant $\ln(TCR)^{t:t+s}$
Import Competition Measure: Mean Tariff Reduction for Industry's Goods Permanently Granted China in 2001 $1(\text{Post-NTR})^t \times$ NTR-Gap

	Interval Duration (s)				
	1 Year	**2 Years**	**3 Years**	**4 Years**	**5 Years**
Import Competition	-0.094	0.148	0.414^{**}	0.634^{***}	0.705^{***}
	(0.141)	(0.165)	(0.173)	(0.179)	(0.190)
× Employment	0.022	-0.017	-0.067^{**}	-0.116^{***}	-0.122^{***}
	(0.026)	(0.031)	(0.032)	(0.034)	(0.036)
Employment	0.000	-0.007	-0.005	0.001	-0.023^{***}
	(0.005)	(0.007)	(0.007)	(0.007)	(0.008)
Estimated Marginal Effects by Plant Employment Size					
40 Employees	-0.015	0.084	0.167^{**}	0.205^{**}	0.255^{***}
	(0.066)	(0.077)	(0.079)	(0.082)	(0.084)
100 Employees	0.005	0.068	0.106	0.098	0.143
	(0.057)	(0.068)	(0.069)	(0.072)	(0.073)
400 Employees	0.035	0.044	0.013	-0.063	-0.027
	(0.062)	(0.076)	(0.076)	(0.080)	(0.082)
Base Years (t)	2000,2001	1999,2001	1998,2001	1997,2001	1996,2001
Number of Plant Clusters	21468	17953	20601	20672	21552
Observations	31781	25205	27194	26114	27496
R^2	0.010	0.017	0.017	0.019	0.018

Notes: All regressions include a constant and year and industry (SIC4) fixed effects. Robust standard errors are clustered by plant and reported in parentheses. Employment is the natural log of the number of employees in the base year. Regressions include all plants with observations that span the respective interval, including those in states that drop out of the sample in later years and in "catch-all" SIC4 industry categories, xxx9. Omitting either or both groups does not significantly affect the results. The employment levels 40, 100, and 400 are approximately the 10^{th}, 50^{th}, and 90^{th} percentiles for the five-year difference estimating sample. The distributions for the samples in the other regressions are similar. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table B.5: Robustness to Misreporting: Severe Injury Effects of Chinese Import Growth (Second Stage IV Results)

Dependent Variable: Log Growth in Severe Injury Rates at Plant $\ln(DART)^{t:t+s}$
Import Competition Measure: Log Growth in Chinese Imports in Industry $\ln(M_{uc})^{t:t+s}$

	Interval Duration (s)					
	1 Year	**2 Years**	**3 Years**	**4 Years**	**5 Years**	**6 Years**
Import Competition	0.249	0.104	0.047	0.170**	0.153**	0.166**
	(0.283)	(0.097)	(0.073)	(0.070)	(0.068)	(0.071)
× Employment	-0.008	-0.016	-0.005	-0.027***	-0.024**	-0.022*
	(0.030)	(0.017)	(0.012)	(0.010)	(0.010)	(0.011)
Employment	0.033***	0.004	-0.028***	-0.016	-0.029**	-0.039**
	(0.008)	(0.009)	(0.010)	(0.012)	(0.014)	(0.016)
Estimated Marginal Effects by Plant Employment Size						
40 Employees	0.220	0.044	0.028	0.071*	0.065	0.084**
	(0.191)	(0.048)	(0.042)	(0.043)	(0.041)	(0.043)
100 Employees	0.212	0.029	0.023	0.047	0.043	0.064
	(0.172)	(0.042)	(0.039)	(0.039)	(0.037)	(0.039)
400 Employees	0.201	0.006	0.016	0.009	0.009	0.033
	(0.147)	(0.041)	(0.038)	(0.037)	(0.036)	(0.039)
Base Years (t)	1996-2006	1996-2005	1996-2004	1996-2003	1996-2002	1996-2001
Number of Plant Clusters	57339	45526	41872	39733	34575	28680
Observations	170419	130346	113785	96890	75895	58266

Notes: All regressions include a constant and year and industry (SIC4) fixed effects. Robust standard errors are clustered by plant and reported in parentheses. Employment is the natural log of the number of employees in the base year. Regressions include all plants with observations that span the respective interval, including those in states that drop out of the sample in later years and in "catch-all" SIC4 industry categories, xxx9. The employment levels 40, 100, and 400 are approximately the 10^{th}, 50^{th}, and 90^{th} percentiles for the five-year difference estimating sample. The distributions for the samples in the other regressions are similar. * p<0.10, ** p<0.05, *** p<0.01

Table B.6: Specialization Robustness Checks

Dependent Variable: 5-Year Log Growth in Injury Rates at Plant $\ln(TCR)^{t:t+5}$
Import Competition Measure: 5-Year Log Growth in Chinese Imports in Industry $\ln(M_{uc})^{t:t+5}$

Dimension of Selection (Base-Year Value)	Plant Employment ≥100	Population Density (ZIP5) ≤ 75th %ile	Share Specialized Plants in Industry ≤ 75th %ile
Import Competition	0.248***	0.207***	0.164***
	(0.088)	(0.076)	(0.058)
× Employment	-0.031**	-0.023**	-0.021**
	(0.013)	(0.011)	(0.009)
Employment	0.002	-0.008	-0.017
	(0.017)	(0.014)	(0.013)
Estimated Marginal Effects by Plant Employment Size			
40 Employees		0.121***	0.088***
		(0.047)	(0.033)
100 Employees	0.106**	0.100**	0.069**
	(0.045)	(0.042)	(0.029)
400 Employees	0.064	0.067*	0.041
	(0.040)	(0.039)	(0.029)
Base Years (t)	1996-2002	1996-2002	1996-2002
Number of Plant Clusters	22659	29546	28531
Observations	48975	65684	63063

Notes: All regressions include a constant and year and industry (SIC4) fixed effects. Robust standard errors are clustered by plant and reported in parentheses. Employment is the natural log of the number of employees in the base year. Results shown are for differences over five years using the supply growth measure of import competition. The results for other intervals and for the tariff gap measure resemble the trends for the main specifications and are available on request. The sample for the third column excludes plants in ZIP codes with population densities above 2,530 people/sq. mile. The sample for the fourth column excludes plants in industries (SIC4) with shares of specialized firms above 78%, based on estimates from Holmes and Stevens (2014). Sensitivity checks demonstrate the results are robust to the respective cutoffs used, and are available on request as well. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table B.7: Levels Specification for Injury Effects of Realized Chinese Import Growth (Second Stage IV Results)

Dependent Variable: Level Growth in Injury Rates at Plant $(\boldsymbol{P})^{t:t+s}$
Import Competition Measure: Log Growth in Chinese Imports in Industry $(\boldsymbol{M_{uc}})^{t:t+s}$

	Interval Duration (s)					
	1 Year	**2 Years**	**3 Years**	**4 Years**	**5 Years**	**6 Years**
Import Competition	4.950^{**}	2.293^{**}	1.177	3.182^{***}	3.283^{***}	1.968^{**}
	(2.223)	(1.053)	(0.725)	(0.736)	(0.786)	(0.765)
× Employment	-0.606^{**}	-0.383^{**}	-0.159	-0.399^{***}	-0.332^{***}	-0.169
	(0.236)	(0.175)	(0.124)	(0.114)	(0.113)	(0.124)
Employment	0.393^{***}	0.224^{**}	-0.008	0.260^{**}	0.195	-0.083
	(0.065)	(0.096)	(0.105)	(0.126)	(0.149)	(0.197)
Estimated Marginal Effects by Plant Employment Size						
40 Employees	2.713^{*}	0.881	0.590	1.712^{***}	2.057^{***}	1.344^{***}
	(1.633)	(0.544)	(0.407)	(0.430)	(0.486)	(0.453)
100 Employees	2.157	0.530	0.444	1.347^{***}	1.752^{***}	1.190^{***}
	(1.528)	(0.472)	(0.376)	(0.389)	(0.442)	(0.420)
400 Employees	1.317	-0.001	0.223	0.794^{**}	1.292^{***}	0.956^{**}
	(1.418)	(0.459)	(0.392)	(0.375)	(0.419)	(0.425)
Base Years (t)	1996-2006	1996-2005	1996-2004	1996-2003	1996-2002	1996-2001
Number of Plant Clusters	70740	56197	52910	49941	43737	36414
Observations	202096	155358	140036	119008	93507	71854

Notes: All regressions include a constant and year and industry (SIC4) fixed effects. Robust standard errors are clustered by plant and reported in parentheses. Employment is the natural log of the number of employees in the base year. Regressions include all plants with observations that span the respective interval, including those in states that drop out of the sample in later years and in "catch-all" SIC4 industry categories, xxx9. Omitting either or both groups does not significantly affect the results. The employment levels 40, 100, and 400 are approximately the 10^{th}, 50^{th}, and 90^{th} percentiles for the five-year difference estimating sample. The distributions for the samples in the other regressions are similar. * p<0.10, ** p<0.05, *** p<0.01

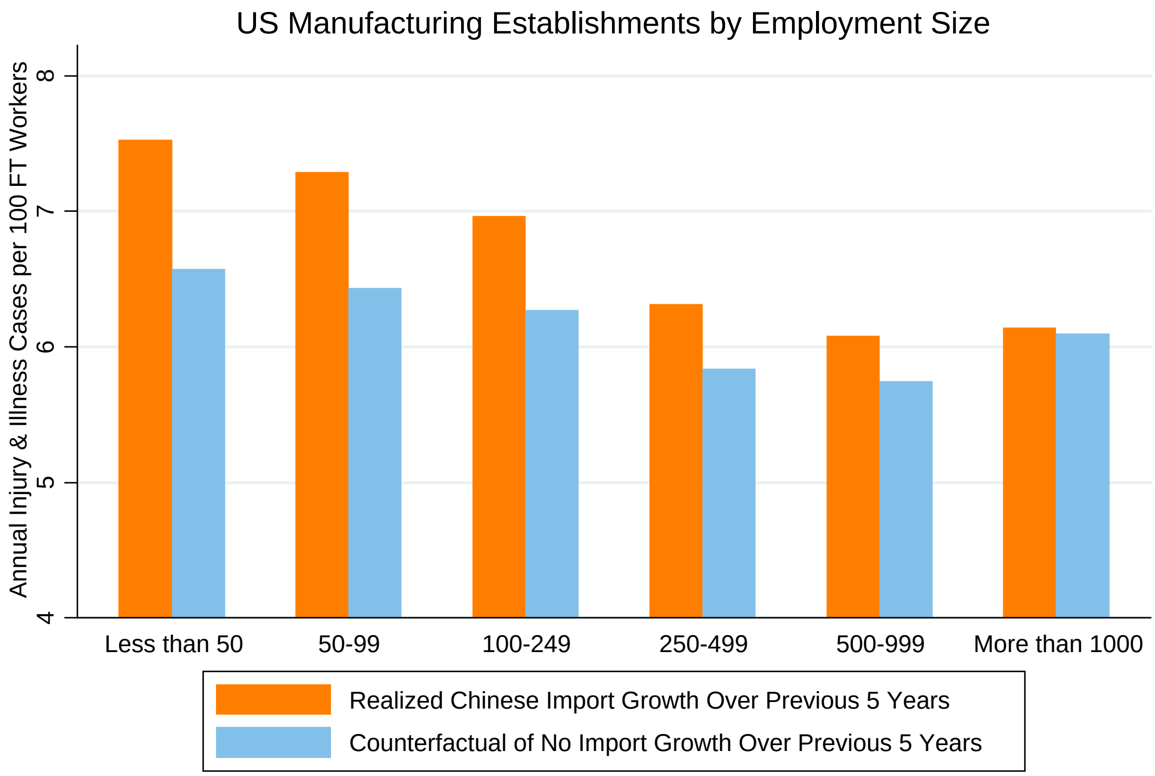


Figure B.1: Predicted Injury Rates in 2007 under Alternative Trade Environments

Appendix B.2 Online Data Appendix

In this appendix, we provide further detail on our injury data and measurement. As we note in the main text, we take the Total Case Rate (TCR) as our main measure of worker injury rates. It is calculated as $\text{TCR} = \frac{\#\text{Injuries and Illnesses} \times 200{,}000}{\text{Total Employee Hours Worked}}$.

We evaluate the representativeness of our sample with regards to both SIC4 industry coverage and the injury rates at plants in our sample. First, we find that yearly industry employment in our sample is highly correlated with population estimates reported in the NBER-CES Manufacturing Industry Database, with a correlation coefficient of 0.823. We next compare yearly industry averages of TCR within our sample (weighted by employment) to estimates for the full industry population reported by the BLS and find that the correlation coefficient to be 0.812. We might expect it to be less than unity because our sample is not representative of the full population of workers – it does not cover those at plants with less than 40 employees.

While the data are highly detailed and allow for the identification of the effect of import competition on injuries, they do not come without problems. The size minimum for sampled plants was originally 60 employees in 1996 and 50 in 1997, and changed to 40 thereafter. Further, selection is based on a BLS record of employment that may not equal the amount reported in our data as a result of changes in employment over the year and potential reporting errors.[44] Lastly, all states participated in the ODI program at first but several self-regulated states discontinued participation during our panel: OR, WA, and WY in 1997, SC in 1999, and AK and AZ in 2006. We estimate our main specification using all eligible observations in the sample, but we run several robustness checks with respect to sample selection issues. First, we estimate the model on a sample that excludes plants with imputed employment below the minimum for the year to account for potential misreporting. We then further restrict the sample to exclude observations in all years with an employment measure of less than 60 FTE to address dynamic selection concerns associated with the changes in eligibility. The first cut eliminates 14% of plant-year observations – representing 2% of employment – from the sample. The second cut eliminates an additional 16% of plant year observations and 4% of employment. Next, we drop all states that discontinued participation in the ODI program at any point during our data window to account for potential endogeneity due to non-random attrition at the state level, which account for about 2% each of total plant-year observations and employment in the full sample. Our overall conclusions remain the same for each robustness check.

Lastly, plants that change SIC4-level industries across an interval of time are excluded from the sample for our difference specifications because the switch creates a spurious change in the industry-level measure of import competition and also likely

[44] About 15% of observed plants have an imputed employment measure less than the year's sampling minimum. Most are just below the margin (likely symptomatic of a true disparity between recorded employment and imputed employment as fulltime equivalent of hours worked), but about 4% are less than half of the minimum (potentially due to inaccurate self-reporting).

affects injury rates via changes in production processes and technology. For our estimating samples of changes over $s = [1,6]$ years, this excludes 2.5%, 5.1%, 6.5%, 7.3%, 8.3% and 9.3% of observations, respectively.

Appendix B.3 Online Theory Appendix

In this appendix, we explain how we derive our reduced form estimating equations from the injury rate equation (1) in our theory.

For our main specification, we generalize equation (1) to derive a log linear specification for injury rates at plant i in industry j during year t. We consider heterogeneous injury costs $c_{ijt} = \lambda_{ij}\exp(t)^{\theta_j} exp(\theta_t)$, that vary in levels across industries and plants, trend at different rates within industries over time and shocks that affect all industries and plants within a year. Take the log of both sides and linearize $\ln(\delta_{ijt} + r)$ around δ^* to get

(B.3.1) $$\ln P_{ijt} = \ln(\delta^* + r) + \left(\frac{1}{\delta^*+r}\right)\left(\delta_{ijt} - \delta^*\right) + \ln\left(A_{ij}\right) + \theta_{\mathrm{j}} t - \theta_t + \eta_{ijt}$$

where $A_{ij} = \frac{\alpha_{ij}}{\gamma_{ij}-\alpha_{ij}} \frac{d_{ij}}{\lambda_{ij}}$ captures plant-specific effects and η_{ijt} captures approximation error. Take differences over s years to obtain

(B.3.2) $$\ln\left(P_{ij}\right)^{t:t+s} = \left(\frac{1}{\delta^*+r}\right)\delta_{ij}^{t:t+s} + \theta_{\mathrm{j}} s - \theta_t^{t:t+s} + \varepsilon_{ijt}.$$

We do not observe survival rates, instead we specify shut-down rates as a function of firm size and a measure of import competition such that

(B.3.3) $$\delta_{ij}^{t:t+s} = \alpha_0 + \alpha_1 \ln\left(M_{uc,j}\right)^{t:t+s} + \alpha_2\left(\ln\left(M_{uc,j}\right)^{t:t+s} \times \ln L_{ijt}\right) + \alpha_3 \ln L_{ijt},$$

where $\ln\left(M_{uc,j}\right)^{t:t+s}$ is the log growth in US imports from China in industry j during the years t to $t+s$[45] and L_{ijt} is a measure of employment. Based on the literature we expect $\alpha_1 > 0$ and $\alpha_2 < 0$. Import competition raises shut-down rates (Bernard et al. 2006a; Pierce and Schott, 2013), especially at the least productive plants (Bernard et al. 2006b) with low levels of employment.

We cannot estimate equation (B.3.3) structurally with our data because plant exit is indistinguishable from non-sampling in our unbalanced panel. We instead substitute it into equation (B.3.2) and estimate the reduced form equation:

(2) $$\ln\left(P_{ij}\right)^{t:t+s} = \beta_0 + \beta_1 \ln\left(M_{uc,j}\right)^{t:t+s} + \beta_2 \ln\left(M_{uc,j}\right)^{t:t+s} \times \ln L_{ijt} + \beta_3 \ln\left(L_{ijt}\right)$$
$$+\tau_j + \theta_t + \varepsilon_{ijt}$$

[45] Our linear probability specification directly relates to existing empirical work which regresses an indicator for plant exit (survival) on the import competition measure.

where $\beta_k = \left(\frac{1}{\delta^* + r}\right) \alpha_k$.

As an alternative, we also consider the level-difference specification of equation (1) which, approximating around the points $c_t = c^*$ and $\delta_{ijt} = \delta^*$, is given by

(B.3.4) $$P_{ij}^{t:t+s} = \frac{A_{ij}}{c^*} \delta_{ij}^{t:t+s} + A_{ij}(c_{t+s}^{-1} - c_t^{-1})(\delta^* + r) + \eta_{ijt}.$$

As before, we substitute equation (A3) for $\delta_{ij}^{t:t+s}$ and derive the estimating equation

(3) $$P_{ij}^{t:t+s} = \beta_0 + \beta_1 \ln\left(M_{uc,j}\right)^{t:t+s} + \beta_2 \ln\left(M_{uc,j}\right)^{t:t+s} \times \ln L_{ijt} + \beta_3 \ln\left(L_{ijt}\right)$$

$$+\theta_t + \varepsilon_{ijt}.$$

where $\beta_i = \frac{A_{ij}}{c^*} \alpha_i$.

CHAPTER III
THE EFFECTS OF IMPORT COMPETITION ON HEALTH IN THE LOCAL ECONOMY

This paper is collaborative work with Georg Schaur at the University of Tennessee.

Abstract

We study the effects of Chinese import exposure in the US on self-reported health measures. We find that average mental, physical, and general health worsens in local labor markets exposed to greater import competition between 2000 and 2007. The effects are greatest for mental health. Moving a region from the 25th to 75th percentile of import exposure corresponds to a 5.5% increase in the time individuals report suffering from poor mental health, adding about 0.18 days per month. The effects are greatest for the employed, consistent with theory from the health literature pertaining to the documented effects of import competition on wages, employment and job security. These estimates provide direct evidence that trade affects welfare through changes in overall mental and physical well-being.

3.1 Introduction

The trade literature documents consequences of globalization for employment and labor market conditions that are separately identified in the health literature as important determinants of mental and physical health among affected workers and their families. The intersection of the two literatures suggests that trade affects welfare through changes in health, which are overlooked by traditional gains-from-trade estimates based on real wages and employment. Looking at one channel, recent work by McManus and Schaur (2015) and Hummels, Munch, and Xiang (2015) finds that globalization affects workplace injuries and illnesses. Studies in the health literature imply more far-reaching implications, however. Further, recent work by Autor, Dorn, and Hanson (2013) finds that import exposure affects labor market outcomes throughout the local economy and not just in import-competing industries. In this paper, we study how the surge in Chinese import exposure in US labor markets, driven by Chinese productivity gains and falling trade costs, has affected overall mental and physical health outcomes. Our empirical results close the gap between these recent findings and provide evidence of broader health effects of trade.

We combine Autor et al.'s strategy for identifying import exposure at the local labor market-level with health microdata from the Behavioral Risk Factor Surveillance System (BRFSS) to study the causal relationship. To generate a regional measure of import exposure per-worker, Chinese imports in each manufacturing industry are mapped to Census commuting zones (CZ) based on the share of the local labor force employed in the industry.[46] An instrumental variables strategy exogenously identifies the change in imports explained by Chinese supply shocks, as opposed to demand shocks that may otherwise be correlated with health outcomes.

[46] Differences in employment composition across regions provide the identifying variation in the model. Anecdotally, the intuition is that a surge in automobile imports will disproportionately affect Detroit given the relative concentration of its workforce in the industry.

Our results show that Chinese import competition in local labor markets tends to worsen self-reported mental, physical, and overall health. Moving a CZ from the 25th percentile to the 75th percentile of import exposure in 2000-2007 explains a 1.06 percentage point increase in the share of the population rating their general health as fair or poor, a 6.7% increase over the mean. The effects are greatest for mental health, which encompasses stress, anxiety, and depression. An increase in import exposure equal to the interquartile range adds 0.18 additional days of poor mental health to the average resident's month, 5.5% of the sample mean. The incidence of poor physical health days, encompassing injury and illness, among the total population is weakly increasing in import competition.

The adverse health effects are greatest among the employed and especially those working for wages (as opposed to self-employed). Among that group, import competition significantly increases the average number of poor mental health days ($p<0.01$) and the number of poor physical health days ($p<0.10$), as well as the population share reporting some incidence of poor mental health ($p<0.10$) and some incidence of poor physical health ($p<0.05$). Given that import exposure is shown to increase firm shutdown rates in competing industries (Pierce and Schott, 2013; Bloom et al., 2012; Bernard et al., 2006a, 2006b), these results are consistent with evidence in the health literature that job insecurity hurts worker mental and physical well-being (Cheng and Chan, 2008; Sverke et al., 2002).

Our results provide direct evidence that changes in the trade environment have consequences for overall mental and physical health, and they complement recent findings for occupational injury and illness. In particular, the deterioration in physical health among the employed that we identify here corroborates findings in McManus and Schaur (2015) that Chinese import shocks increase worker injury and illness rates at US manufacturers in competing industries. Other related work by Hummels et al. (2015) meanwhile looks at export shocks and finds that exogenous export growth at Danish firms increases occupational injuries and illnesses and increases incidence of severe depression. We contribute to these studies in the following ways: First, we identify a broader range of health effects by using measures of overall mental, physical, and general health rather than specific outcomes like occupational injury or visits to a psychiatrist only. Second, our analysis at the regional level captures spillover effects within local economies of the sort identified in Autor et al. (2013), as well as interpersonal spillovers such as the mental health costs of spousal job insecurity (Wilson, 1993) and job loss (Marcus, 2013; Bubonya et al., 2014), and therefore captures a greater share of the total health effects. We also capture indirect effects that arise from workers being displaced by import exposure into unemployment or lower-paying jobs (as found in Ebenstein et al., 2014). To our knowledge, this is the first paper to look at the effects of trade on overall health in the economy, and we present the first evidence that directly links import competition to mental health.

We also make use of data on access to medical care to provide new evidence on the consequences of import competition with relevance for healthcare policy. We find that increasing 2000-2007 import exposure in a region by the interquartile range of $1000 per worker increases the share of the employed who are unable to afford necessary medical care by 2.4 percentage points over that time, a 24% increase over the sample mean. These findings suggest that the welfare costs of the health effects identified in this paper and in McManus and Schaur (2015) may be compounded by import competition hurting people's ability to pay for the needed medical care. The health costs of import shocks may be lower under socialized healthcare systems or mandated employer health plans that shield workers from an endogenous fall in their health insurance coverage. We find a positive but insignificant relationship between import exposure and the share of people without health coverage.

This paper also contributes to existing evidence on the local labor market effects of globalization, including McLaren and Hakobyan's (2012) study of the effects of NAFTA and Autor et al.'s (2013) study of Chinese import exposure.[47] We employ the methodology from Autor et al. such that our results for health outcomes directly compare to theirs for the fall in average wages, employment, and labor market participation. Their estimates imply that an additional $1000 in Chinese import exposure per worker results in a 0.8 percentage point reduction in log weekly earnings and a 4.5% fall in manufacturing employment. In comparison, our estimates imply a 6.8% increase in the share of the population with fair or poor general health and a 5.7% increase in days of poor mental health for the equivalent shock.

Lastly, this paper contributes to the literature on the macro- and microeconomic determinants of health, much of which focusses on the effects of business cycle fluctuations and unemployment. Employment status and job security are robustly linked to mental health and subjective well-being.[48] Blanchflower and Oswald (2004) estimate it would take $108,000/year[49] to compensate a man for the disutility associated with being unemployed. Luechinger et al. (2010) and Clark et al. (2010) find that high unemployment rates hurt the employed also via greater economic distress and fear of job loss. Studies in occupational health and organizational psychology document the consequences of greater job insecurity for various mental and physical health outcomes, including stress, anxiety, depression, hostility, somatization, high blood pressure, heart disease, and stroke.[50] Our results show that the trade environment, like the business cycle, is an important determinant of mental and physical health in the economy.

The remainder of the paper is organized as follows: Section 2 discusses the related literature; Section 3 presents the data and measurement; Section 4 explains the empirical specification and discusses identification; Section 5 provides the results; Section 6 presents robustness checks, and Section 7 concludes. Tables and figures are provided in Appendix C.1.

---

[47] See also Kovak (2013) for a specific-factors model of wage and price changes within a regional economy and empirical support from Brazil, along with Chiquiar (2008) and Topalova (2010) for empirical evidence of local labor market effects in Mexico and India, respectively.

[48] For example, see Karsten and Moser (2009) and McKee-Ryan et al. (2005) for studies on unemployment, and see Cheng and Chan (2008), Sverke et al. (2002), and De Witte (1999) for the effects of job insecurity on worker health.

[49] In 2015USD, converted from their estimate of $60,000 in 1990USD.

[50] For a review of that literature, see meta-analyses by Bohle et al. (2001) and Sverke et al. (2002).

3.2 Related Literature

Chinese exports to the US and the rest of the world surged in the 2000's, driven by Chinese productivity gains related to its transition to a market-oriented economy (Chen, Jin, and Yue, 2010) and falling trade costs associated with its accession to the WTO in 2001 (Pierce and Schott, 2013). The intersection of evidence in the trade and health literatures explains several channels for these import shocks to affect mental and physical health in the US both positively and negatively.

Import shocks predominately reduce employment in import-competing industries, as Pierce and Schott (2013) show for US manufacturers following the trade liberalization towards China. Displaced workers switch to other occupations and industries where wages are lower (Ebenstein et al., 2014) or enter unemployment, where mental and physical health suffers (Karsten and Moser, 2009; McKee-Ryan et al., 2005). Job loss carries additional health costs beyond those of being unemployed only (Kassenboehmer and Haisken-DeNew, 2009) and is directly linked to greater mortality rates (Sullivan and von Wachter, 2009). Workers who remain employed in import-competing industries face more tenuous employment conditions as firm shutdown rates rise (Pierce and Schott, 2013; Bloom et al., 2012, Bernard et al., 2006a, 2006b). The health literature explains this will hurt mental and physical health, as job insecurity is in part associated with greater stress and anxiety and increased risk of heart disease and stroke. (Cheng and Chan, 2008; Sverke et al., 2002; De Witte, 1999).

Studies on health and the business cycle explain mechanisms through which import competition can improve health as well. Ruhm (2000) finds that physical health outcomes improve during a recession, an effect that subsequent work attributes to a fall in working hours (Ruhm, 2005) and decreased economic activity such as driving (Miller et al., 2009).

The evidence on the effect of income on health is mixed, due in part to challenges in dealing with reverse causality.[51] Ettner (1996) finds that an exogenous increase in income improves mental and physical health, and Lindahl (2005) reports sizeable positive effects for overall health and mortality risk as well. Frijters et al. (2005) finds a positive but very small effect on self-reported general health, meanwhile Meer et al. (2003) finds that the effects are statistically indistinguishable from zero and Ruhm (2005) does not find evidence that income changes explain the health effects of the business cycle. These collective results suggest that wage losses from import competition will tend to weakly worsen health outcomes, or else have no effect at all.

Work-related injuries and illnesses are another important contributor to overall health, and the main channel thus far explored in the recent literature on trade shocks and health. McManus and Schaur (2015) finds that worker injuries and illnesses go up in US manufacturing industries hit with Chinese import competition, especially at the smallest and least productive establishments. The theory in that paper explains that marginal firms sacrifice safety for productivity when import shocks lower the firm's survival rate

[51] Notably, reverse causality issues would not seem to threaten our analysis here as country-level trade flows are presumably exogenous of individual health outcomes.

and the likelihood it incurs long-run costs for its workers' injuries. The empirical analysis focusses on within-firm effects only, however, and does not capture labor market reallocations. Occupational health will improve for workers displaced into safer jobs, though the associated wage losses (per Ebenstein et al.) potentially dampen the effects on overall health. Our analysis in this paper captures these general equilibrium effects by looking at overall health outcomes across the labor market. Hummels et al. (2015) also document the link between trade and occupational health: they find that exogenous export shocks at Danish firms increase worker injuries and sick days, a result they attribute to greater work intensity as firms respond to positive demand shocks.[52] Looking to mental health, they find that export shocks increase workers' likelihood of visiting a psychiatrist and taking anti-depressants. Our study complements their results in looking at mental and physical health outcomes in the US and as a consequence of import competition.

We follow Autor et al. (2013) and estimate the aggregate effects of import competition at the regional (subnational) level. In comparison to McManus and Schaur (2015) and Hummels et al. (2015), this spatial identification strategy captures both health effects among workers who are displaced by import competition into other industries or occupations (per Ebenstein et al., 2014) and indirect effects that arise due to spillovers within affected regions. Autor et al. document the importance of the latter channel: they find that Chinese import exposure significantly lowered wages at non-manufacturing jobs in regions with a high concentration of import-competing firms. Further, there is also reason to believe the consequences for workers adversely affect their spouse's mental health, as Wilson (1993) shows for job insecurity and Marcus (2013) and Bubonya et al. (2014) show for job loss and unemployment, and our analysis at the regional level captures these indirect effects also. One cost of this strategy is that we are left with considerably fewer observations than firm-level analyses; however, there is still considerable variation in both trade shocks and health outcomes across regions, as discussed below.

3.3 Data and Measurement

The trade data comes directly from Autor et al. (2013) and is discussed in more detail there. Briefly, they construct a per-capita measure of import competition within a local labor market in the following manner: first, they map UN Comtrade data on US imports from China at the HS6 product level to the SIC4 industries that manufacture each product. Next, they apportion each industry's imports to local labor markets– defined as the 722 Census commuting zones (CZ) covering the mainland US – based on their share of national industry employment. The result is the following per-capita measure of the change in Chinese import exposure in local labor market i following base year t:

---

[52] Different channels explain why both import and export shocks can adversely affect worker health. This is consistent with evidence in Sokejima and Kagamimori (1998) of a U-shaped relationship between working hours and myocardial infarctions and findings in the occupational safety and health literature that show both shorter and longer work schedules can increase injury rates.

$$(5) \qquad \Delta IPW_{it}^{US} = \sum_j \frac{L_{ijt}}{L_{it}} \frac{\Delta M_{jt}^{US}}{L_{jt}}$$

where $\frac{L_{ijt}}{L_{it}}$ is the share of region i's labor force employed in industry j in the base year and $\Delta M_{jt}^{US}/L_{jt}$ is the per-worker growth in US imports from China in the industry (in \$1000s). The variation in this measure is driven by heterogeneity in the employment structure across CZs. In particular, Autor et al. note that 75% of the variation is driven by differences in manufacturing industry concentration, with the remainder explained by variation in manufacturing versus non-manufacturing employment. For use as an instrument, they also construct a similar measure that uses lagged employment and growth in Chinese imports in an aggregate of eight other high income countries[53]: $\Delta IPW_{it}^{OTH} = \sum_j \frac{L_{ijt-1}}{L_{it-1}} \frac{\Delta M_{jt}^{OTH}}{L_{jt-1}}$. Along with ΔIPW_{it}^{US} and ΔIPW_{it}^{OTH} for the period 2000-2007, we use their data for CZ data on population, employment and demographics that are used as control variables.

The health data comes from the Center of Disease Control's Behavioral Risk Factor Surveillance System (BRFSS).[54] BRFSS is a national telephone survey that monitors a broad range of health-related behaviors and outcomes. As part of the core set of questions, the survey asks respondents the number of days in the last month that their mental health was not good and the number of days their physical health was not good.[55] Respondents are also asked to rate their general health on a five-point scale from poor to excellent.

We use individual responses in 2000 and 2007 to create variables on the change in average health outcomes within commuting zones over this period. These variables are: population share reporting fair or poor general health, share reporting some days of poor mental (physical) health in the past month, and the average number of poor mental (physical) health days reported.[56] Table C.1 provides summary statistics for these health measures for both CZs and for the individual microdata used to construct the CZ-level outcomes, along with statistics on the within-CZ coverage of the data. Figure C.1 shows the distributions of CZ health outcomes for the pooled 2000 and 2007 samples. We also use BRFSS data on employment status to calculate the respective health measures within the subgroups of: employed, self-employed, employed for wages, unemployed (out of

---

[53] These are: Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

[54] Centers for Disease Control and Prevention (CDC). *Behavioral Risk Factor Surveillance System Survey Data*. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2000, 2007.

[55] Specifically, they were asked: "Thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 was your mental health not good?", and "Thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?".

[56] We map the counties reported in BRFSS to CZs using a crosswalk from Autor et al. (2013). We thank them for making this publicly available on David Dorn's website.

work), and homemakers. More detailed job information, such as occupation and industry of employment, are not available in the data.

BRFSS is well-suited to this application for three reasons: First, it is administered to a representative sample across the entire US and all respondents are asked the same core questions, providing a comparable measure across regions. Second, the survey has asked these same core questions about health annually since 1993, such that we have an identical measure of health in 2000 and 2007 to compare. Third, BRFSS is the largest health survey in the world and provides large sample sizes even within regions defined as narrowly as CZs. On average, there are 681 individuals used to estimate each CZ-year measure in our estimating sample; the median is 303 and none are comprised of fewer than 44 individuals.

One issue that arises in using BRFSS data for our analysis is that small counties are censored in the data. As a result, we are only able to derive health measures for 340 CZs, as the rest are comprised entirely of censored counties. Further, the coverage within the CZs that we do observe will be biased towards urban residents if some counties in the CZ are censored. We discuss the threats to identification that stem from this data limitation below.

3.4 Identification

In this section, we present the empirical model that we take to the data on regional import exposure and health outcomes. We then discuss potential identification concerns.

Let i denote commuting zones and t denote years. The reduced form model for the change in a CZ health outcome O_{it} is given by:

$$\Delta O_{it} = \alpha + \beta \Delta IPW_{it}^{US} + \gamma X_{it} + \Delta\varepsilon_{it} \tag{6}$$

where ΔIPW_{it}^{US} is the measure of Chinese import exposure and X_{it} is a vector of CZ characteristics in the base year that are independent of the ensuing growth in imports. Any time invariant CZ characteristics related to health outcomes are differenced out. The common change in health outcomes across all CZs is captured by the constant. In most regressions, we also include Census division dummies (as in Autor et al., 2013) that allow for different trends across the nine major regions of the US.[57] We do not log transform O_{it} because the distribution in levels is relatively normal (see Figure C.1) and there a considerable number of zeros in some of the CZ-level data for smaller subgroups (e.g. no homemakers in the CZ sample report poor or fair general health).

We set out to identify the health effects of import exposure driven by China's productivity gains and falling trade costs. Import growth is correlated with demand shocks, however, which will bias OLS estimates for β if demand shocks are also

[57] The Census divisions are: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.

correlated with health outcomes as the literature suggests.[58] We follow Autor et al. and address the concern by using Chinese import growth in other OECD countries to instrument for the growth in the US. The instrument ΔIPW_{it}^{OTH} is also constructed using ten-year lagged employment values to account for any potential anticipation effects that may render base-year employment endogenous to the ensuing import growth. We estimate equation (6) using two-stage least squares (2SLS), where the first stage regresses the endogenous ΔIPW_{it}^{US} on the instrument ΔIPW_{it}^{OTH}, along with the controls in the second-stage equation. The identifying assumption here is that the common growth in imports across the markets is driven by Chinese supply shocks and not demand-side factors that are potentially correlated with health outcomes. Under the additional assumption that $\Delta\widehat{IPW}_{it}^{US}$ is uncorrelated with $\Delta\varepsilon_{it}$, the OLS estimate of β in equation (6) using $\Delta\widehat{IPW}_{it}^{US}$ consistently identifies the net effect of import exposure on regional health averages.[59]

Autor et al. identify several threats to the validity of the IV in their study on wages and employment that are relevant here also. First, product demand shocks in the US and other high-income countries may be correlated such that the IV estimate does not exogenously capture supply-side effects only. They check this by employing a gravity identification strategy that isolates import growth attributable to changes in China's productivity and trade costs only and find similar results. They conclude their IV results are not driven by correlated demand shocks, and so we proceed with the same assumption for these health effects. Next, it may be the case that US producers who suffer adverse productivity shocks are supplanted by Chinese imports in US and other OECD markets, but it is more likely that China's export growth is driven by the huge gains in its own productivity, which grew by about 8% during this time (Brandt et al., 2012), more than double that in the US. Finally, the correlated growth in Chinese imports across the markets may be driven by common technology shocks that hurt labor-intensive industries at home. That Chinese export growth dwarfed that of other low and middle-income nations, however, suggests that the growth was instead driven by falling trade costs and productivity shocks unique to China. Notably, the OLS estimates for β are generally lower than the IV estimates, suggesting that any remaining correlation between the IV estimates for ΔIPW_{it}^{US} and demand shocks will bias against finding adverse health effects in the 2SLS results.

The next concern for identification is that import exposure in a region drives changes in the composition of the population that are related to health outcomes. This would change average health outcomes within a CZ over time independent of any changes in individual health. This seems unlikely, however. Autor et al. (2013) find no significant relationship between CZ import exposure and population size or composition

[58] For example, Ruhm (2000) and others relate income and employment shocks to health. Under the IV assumption, our analysis will only capture the health effects of these shocks to the extent they are correlated with import exposure.

[59] We opt to use OLS instead of a fractional logit regression for the population shares reporting poor mental/physical/general health outcomes because we cannot difference out CZ fixed effects in the nonlinear logit specification and a FE specification with T=2 suffers from an incidental parameters problem.

across age or education level.[60] Still, we regress import competition on CZ age, gender, and racial demographics from our sample as a robustness check to test that import exposure is not shifting the makeup of the population along these potential covariates of health outcomes.

Another concern for our identification is that cross-sectional differences in commuting zones might systematically explain both import exposure and trends in health. In this case our estimates would not identify a causal link between import exposure and health, but rather a spurious relationship driven by an omitted variable bias. In a robustness check, we regress the 1993-1999 change in poor health outcomes on *future* import exposure in 2000-2007 to test for a relationship between the shock in a CZ and long run health trends.

Last, we note that the estimating sample is not representative of the entire US population. County identifiers in the public BRFSS data are censored for low-population counties and so we cannot assign these individuals to commuting zones when calculating the average health measures.[61] We can only generate CZ averages for 340 of the 722 commuting zones covering the US. The out-of-sample regions are the most sparsely populated, such that the estimating sample of CZ-level changes that we use is representative of about 71% of the national population.[62] Manufacturing employment shares and import exposure per worker are not significantly different across the CZs covered in our sample and those not. To the extent that other factors determine the effects of import competition on health, however, our results will not be representative of these out-of-sample areas.

3.5 Results

We first estimate equation (2) on the primary health measures for the overall population: share of the CZ reporting fair or poor general health, average days of poor mental health per month, and average days of poor physical health per month. The 2SLS results are shown in Table C.2. In all regressions CZ observations are weighted by base-year population and standard errors are clustered at the state level to allow for unobserved correlation across CZs in a state (e.g. due to a change in a state health program during the period).[63] The first-stage regression results for each specification are presented in the

[60] Autor et al. cite findings in Blanchard and Katz (1992) and Glaeser and Gyourko (2005) that show people are slow to move following labor demand shocks, most of all workers without a college education (Notowidigdo, 2013), a prominent share of manufacturing labor.

[61] Counties are censored if there are less than 50 individuals in the BRFSS sample in a year. The censoring is more prevalent in earlier years, which had smaller sample sizes – e.g. the sample size grew from about 180,000 to 430,000 between 2000 and 2007, amounting to greater coverage within counties and fewer being censored.

[62] About 32% of individual observations are in counties that are either censored in 2000 or 2007, but this represents only 29% of the national population (adjusting for sampling weights).

[63] The results are qualitatively similar when observations are weighted instead by number of employed workers and number of unemployed workers in the CZ in the base-year. Keeping the weights constant across specifications ensures that they are not responsible for differences in the estimates for the full population and the subgroups broken out in Table 3.

first columns and thereafter suppressed. The coefficient estimates and robust F-statistics from a weak instruments test reveal that the instrument is significantly correlated with US import exposure in each specification.

In Table C.2, the first column for each health measure is a univariate regression of its change on the contemporaneous change in import exposure. The results show that the markers for poor general health and poor mental health are significantly and positively correlated with import exposure, while the effect on poor physical health is positive but not statistically significant. The second specification adds dummies for the nine Census divisions and the third adds the full set of controls from Autor, Dorn, and Hanson's (2013) regressions for employment and wage changes. These additional covariates are: the share of the population employed in manufacturing, the share with a college education, the percentage that are foreign-born, the percentage of employment among women, the share of workers employed in routine occupations, and the average offshorability index of the occupations. We consider that these potential correlates of trends in other labor market outcomes can also be correlated with changes in health, along with import exposure, and so they belong in the estimating equation. The additional controls increase the magnitude of the estimated effect of import competition for share of fair/poor general health and average poor mental days, which both remain statistically significant, while the effect on poor physical days remains positive but not significantly different from zero. We regard the specification with the full set of controls as our preferred model.

The coefficients give the estimated effect of a \$1000 increase in competing Chinese imports per worker in the CZ over a seven-year period. As this is the approximate interquartile range for import exposure in our sample, the coefficients translate to the implied effects of moving a CZ from the 25th percentile of import exposure to the 75th percentile. We estimate that this shock adds about 0.18 days/month of poor mental health for the average adult in the CZ, a 5.5% increase over the sample mean. It increases the population share with fair or poor general health by 1.06 percentage points *ceteris paribus*, 6.7% of the mean. These results are comparable in magnitude to estimates in McManus and Schaur (2015) that imply increasing import exposure in an industry by the interquartile range increases worker injury rates by 10% over five years at the median plant size.

It is a priori unclear what drives the decline in overall health and mental health in particular. One candidate explanation from the literature is that workers face greater job insecurity and stress at work as a result of Chinese import competition. These effects would be concentrated among the employed. Another explanation is that import competition exerts greater stress and anxiety on the unemployed, who face worse prospects for reemployment. We make use of employment status in the data to identify the relative contributions of each to the overall effects. We construct CZ measures of each health outcome using the subsample of observations in the BRFSS data that are employed and that are unemployed and estimate the model on the difference in outcomes within each group in 2000 and in 2007. Table C.3 gives the estimates of equation (6) for each health measure across each of the employment subgroups. In this table we also estimate the model for the share of the population reporting any incidence of poor mental health and any incidence of poor physical health. The results in panel B show that health is worsening especially among the employed, who are more likely to report both poor mental health days ($p<0.10$) and poor physical health days ($p<0.05$) as a result of greater import competition. The estimates imply that employed adults in a CZ with import exposure at the 75th percentile will all else equal have an additional 0.31 days of poor health per month than their counterparts in a CZ at the 25th percentile, a 9.3% increase over the sample mean. Shown in panel C, the effects are even greater for the subset of the employed working for wages (as opposed to self-employed). These results are consistent with the documented effects of job insecurity on mental and physical health, though we cannot directly test the mechanism here.[64] The estimates in panel D show that for the unemployed import exposure weakly improves mental health and has the opposite effect on physical and general health, however none of the results are statistically significant.[65]

Another empirical question is to what extent these health effects spillover throughout the economy. The health psychology literature argues that health outcomes are interpersonally determined: Wilson et al. (1993) finds that job insecurity affects spousal stress and mental well-being, and Marcus (2013) and Bubonya et al. (2014) provide similar evidence for unemployment. To test for this with our data, we estimate the model for the CZ-level health measures among individuals who self-identify as homemakers. The results, shown in Table C.3, do not provide any evidence of adverse spousal health effects. Instead, there is weak evidence that mental and physical health improves among this subgroup.[66]

We lastly consider how import competition affects health through changes in access to healthcare, a prominent concern for policymakers. For one, health outcomes might deteriorate because individuals struggle to afford medical care as a result of import exposure. Individuals in BRFSS report whether there was a time in the past year when they needed to see a doctor but could not because of the cost. We regress the change in this measure on import exposure to test for the relationship. The results for the total population, the employed, and the employed for wages are shown in the first three columns of Table C.4. The estimated coefficient of interest is positive and statistically significant for all three groups, indicating that import exposure increases the share of people who find themselves unable to afford necessary healthcare. The effects are greatest for the employed and even more so for those employed for wages. For that group, the coefficient estimate implies that increasing import exposure by the interquartile range of \$1000 per worker increases the share of wage-earners unable to afford medical care by 2.4 percentage points, almost one-quarter of the average of 10%.

---

[64] Another candidate explanation is that well-being deteriorates for workers displaced into other industries or occupations as a result of import competition, a consequence documented by Ebenstein et al. (2014).

[65] For one, the gains in mental health are consistent with the finding in Clark (2003) that the unemployed benefit from a “social norm effect” as more people become unemployed.

[66] For one, this is consistent with homemakers from lower-income households leaving that group for employment and leaving a wealthier and healthier population of homemakers as a result. This can also be explained by trade liberalization increasing wages for the most skilled workers, to whom homemakers are more likely married.

In addition to imposing welfare costs in its own right, reduced access to health care would seem to exacerbate the other health consequences of import competition, including the occupational injury and illness effects identified in McManus and Schaur (2015). In conjunction, these results imply that import competition both increases injuries and illnesses and reduces workers' ability to pay for related medical care.

That workers are more likely to be priced out of medical care because of import exposure can in part be attributed to the documented wage losses in affected CZs (Autor et al., 2013), but another explanation is that firms respond to import competition by downgrading or canceling worker health plans such that workers cannot afford the higher co-pays or uncovered medical expenses. Universal healthcare in Denmark helps shut down this channel in Hummels et al.'s (2015) study on worker injury and illness, but it is a potentially important margin of adjustment in economies with private healthcare systems – like the US during this period – and one of particular importance to policy makers. To look at endogenous changes in insurance coverage using the BRFSS data, we regress the change in the share of people without any health coverage on import exposure. The results are shown in columns 4-6 of Table C.4. The estimated coefficient of interest is positive for all groups and greatest for wage employees, but it is not statistically different from zero in the specification with the full set of controls and so we cannot reject the null hypothesis that import competition does not affect health coverage.

3.6 Robustness Checks

We run several robustness checks related to our identifying assumptions and alternative specifications. First, we address our assumption that changes in the population which are correlated with health are not driven by import competition. We regress equation (6) on changes in demographic variables in our data to test if these potential covariates of health outcomes are in fact endogenous to our measure of import exposure, which would bias our main results. The results are shown in Table C.5. Shown in the first row of estimates, we do not find any evidence that import exposure within a CZ significantly affects the composition of the overall population across gender, age, or race during the seven-year interval. These results with our sample data are consistent with evidence on age and education presented in Autor et al. (2013).[67] We also look at potential endogenous changes in the composition of the population in the employment status groups we look at above. For example, we may be concerned that import exposure disproportionately displaces older workers – who tend to have better mental health outcomes – and this change in the composition of the employed contributes to the positive relationship between import exposure and poor mental health that we identify. This does not appear to be the case, however. The estimates in the lower rows of Table

[67] Time-varying individual outcomes like educational attainment are potentially endogenous – Greenland and Lopresti (2013), for example, shows high school dropout rates fall in US labor markets exposed to Chinese import competition.

C.5 reveal that import competition does not generally explain changes in population characteristics with any degree of statistical significance.[68]

We next consider the threat that some unobserved CZ characteristics explain both the change in health and the level of import exposure such that our identification suffers from omitted variable bias. We look at this by regressing the change in health outcomes in 1993-1999 on the future import growth in 2000-2007 to test for pre-trends. A positive relationship would suggest that CZ characteristics which are correlated with long run deteriorations in average health are also correlated with greater import exposure. The results for each health outcome by subgroup are shown in Table C.6. We instead find a weakly negative relationship in most cases, especially for mental health: areas with high import exposure were seeing relative improvements in health preceding the 2000-2007 shock. Assuming health trends in 1993-1999 and 2000-2007 are positively related, these results indicate that the potential bias – if there is any – works against our main findings. We conclude that our results are not driven by omitted CZ characteristics that explain both trends in health and import exposure.

3.7 Conclusion

We find empirical evidence that Chinese import exposure in the US worsened average health outcomes in affected local labor markets. The estimates imply that the mean import shock in a region increased the share of the population reporting fair or poor general health by 6.7%. The effects are greatest for self-reported mental health. The mean shock increases the incidence of poor mental health among all adults by 0.18 days per month, about 5.5% of the average. Self-reported mental and physical health especially worsens among the employed. The results are consistent with the intersection of evidence that import competition reduces employment, wages, and job security, and findings in the health literature that link these channels to worse mental and physical health.

Looking to related policy implications, we find that import exposure significantly increases the share of people who are at times unable to afford necessary medical care. This suggests that the injury and illness effects of import competition in the workplace and beyond may be exacerbated by reduced access to healthcare. The health costs of trade shocks will likely be less under a system of social healthcare or mandatory employer-provided health plans that cannot be endogenously downgraded or eliminated. Further research with richer data can provide additional insight on how trade affects health insurance coverage and quality, workers' ability to afford healthcare, and other non-wage determinants of mental and physical health outcomes.

[68] One exception among wage employees is that import competition appears to weakly displace white workers from that group and replace them with Hispanic workers. Regressing changes in self-reported health on changes in racial makeup across all adults – which we find to be exogenous to import exposure in the first row of Table C.5 – reveals that Hispanics tend to report fewer days of poor mental and physical health than white non-Hispanics. This shift in the racial makeup of wage-earning workers would therefore seem to bias against our results.

Appendix C.1 Tables and Figures

Table C.1. Summary Statistics of BRFSS Data

	N	Mean	Std Dev	Min	Max
Individual-Year Data					
Days/Month of Poor Mental Health	457905	3.36	7.61	0	30
Indicator: > 0 Poor Mental Days	457905	0.32	0.47	0	1
Days/Month of Poor Physical Health	456684	4.00	8.51	0	30
Indicator: > 0 Poor Physical Days	456684	0.36	0.48	0	1
Indicator: General Health Fair or Poor	465040	0.17	0.38	0	1
Sample Sizes Within Commuting Zone-Year Aggregates					
Full Sample	340	681.5	1184.3	44	14180
Employed	340	388.8	675.3	24	7814
Self-Employed	339	57.2	101.0	1	1181
Employed for Wages	340	331.7	577.6	21	6742
Unemployed	314	25.9	49.8	1	672
Homemakers	335	53.5	89.7	1	1140
Commuting Zone-Year Data					
Average Days of Poor Mental Health	680	3.27	0.86	0.40	6.39
2000-2007 Change	340	0.15	1.02	-2.97	3.85
Share Reporting Poor Mental Days (ppts)	680	32.4	6.3	10.3	52.2
2000-2007 Change	340	0.01	7.5	-25.3	26.9
Average Days of Poor Physical Health	680	3.55	0.97	0.58	7.93
2000-2007 Change	340	0.42	1.01	-2.38	5.11
Share Reporting Poor Physical Days (ppts)	680	34.3	5.4	13.7	56.8
2000-2007 Change	340	2.2	6.9	-15.8	31.4
Share Reporting Fair/Poor Gen. Health (ppts)	680	15.8	5.5	0.3	36.7
2000-2007 Change	340	1.8	4.7	-13.8	16.2

Table C.2. Chinese Import Shocks and Health Outcomes in Local Labor Markets: 2SLS Estimates

	General Health Share Reporting Fair or Poor (ppts)			Mental Health Average Poor Days/Month			Physical Health Average Poor Days/Month		
	(G.1)	(G.2)	(G.3)	(M.1)	(M.2)	(M.3)	(P.1)	(P.2)	(P.3)
ΔIPW_i^{US}	0.575**	0.706**	1.063**	0.107***	0.120***	0.183***	0.070	0.081	0.075
	(0.283)	(0.319)	(0.427)	(0.373)	(0.036)	(0.091)	(0.048)	(0.052)	(0.064)
Initial Mfg. Emp. Share			-0.090			-0.018			-0.003
			(0.076)			(0.017)			(0.013)
Initial Share College-Educ.			0.072			-0.015			0.003
			(0.056)			(0.014)			(0.010)
Initial Share Foreign-Born			-0.026			-0.019**			-0.008
			(0.030)			(0.010)			(0.007)
Initial Female Emp. Rate			-0.090			-0.028			-0.009
			(0.080)			(0.023)			(0.014)
Initial Emp. Share in Routine Occ.			-0.164			0.015			-0.059**
			(0.263)			(0.030)			(0.026)
Avg. Offshorability Index of Initial Occ.			-0.636			0.135			0.026
			(1.848)			(0.227)			(0.204)
Census Division FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
	First Stage Results for ΔIPW_i^{US}								
ΔIPW_i^{OTH}	0.756***	0.737***	0.526***						
	(0.104)	(0.117)	(0.121)						
Wk Instrument F-Stat	50.76	35.95	18.77						
R^2	0.425	0.445	0.502						

Notes: N=340 for all regressions. Observations are at the commuting zone- level and weighted by base-year population. Robust standard errors are clustered by state and reported in parentheses. First stage regressions and results are identical for the corresponding columns of the mental and physical health regressions and are suppressed, as are coefficient estimates for the other covariates in the first stage equation. * p<0.10, ** p<0.05, *** p<0.01

Table C.3. Local Health Effects by Employment Status

	Avg Poor Mental Days/Month	Share w/Poor Mental Days (ppts)	Avg Poor Physical Days/Month	Share w/Poor Physical Days (ppts)	Share w/ Fair or Poor General Health
			Panel A: Full Population		
ΔIPW_i^{US}	0.183**	1.162	0.075	1.267*	1.063**
	(0.091)	(0.748)	(0.064)	(0.699)	(0.427)
			Panel B: Employed		
ΔIPW_i^{US}	0.305**	1.597*	0.144	1.891**	0.814
	(0.121)	(0.873)	(0.096)	(0.866)	(0.523)
			Panel C: Employed Wage Earners		
ΔIPW_i^{US}	0.322***	1.696*	0.149*	2.383**	0.932
	(0.124)	(0.925)	(0.091)	(0.998)	(0.586)
			Panel D: Unemployed		
ΔIPW_i^{US}	-0.642	-4.216	0.347	4.049	4.020*
	(0.840)	(4.468)	(0.524)	(3.903)	(2.331)
			Panel E: Homemakers		
ΔIPW_i^{US}	-0.449	-2.902*	-0.196	-4.968	-0.261
	(0.312)	(1.713)	(0.167)	(3.055)	(0.729)

Notes: N=340 for the regressions for full population, employed, and employed wage earners, N=312 for unemployed, and N=335 for homemakers. Observations are at the commuting zone level and are weighted by base-year share of national population. All regressions use the preferred specification with Census division dummies and the full set of controls from Autor et al. (2013), listed below Table C.2. Robust standard errors are clustered by state and reported in parentheses. First stage results are available upon request. Employed wage earners include all employed not classified as self-employed. Unemployed persons are those indicated to be out of work (for either less than a year or more than a year), and exclude self-identified retirees, students, and homemakers. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table C.4. Import Exposure and Health Care Access

	Share Couldn't Afford MD Visit (ppts)			Share Without Health Plan (ppts)		
	Full Pop.	Employed	Employed for Wages	Full Pop.	Employed	Employed for Wages
ΔIPW_i^{US}	0.939* (0.565)	2.099** (1.030)	2.403** (1.097)	0.787 (0.522)	0.497 (0.540)	0.930 (0.627)
Initial Mfg. Emp. Share	-0.018 (0.069)	-0.023* (0.131)	-0.251* (0.150)	0.008 (0.075)	0.074 (0.085)	0.058 (0.098)
Initial Share College-Educ.	0.161*** (0.053)	0.117 (0.073)	0.146* (0.075)	0.029 (0.066)	0.039 (0.064)	0.043 (0.063)
Initial Share Foreign-Born	-0.027 (0.034)	-0.007 (0.053)	0.002 (0.051)	-0.081*** (0.030)	-0.062 (0.041)	-0.027 (0.041)
Initial Female Emp. Rate	-0.184*** (0.071)	-0.013 (0.074)	-0.048 (0.097)	-0.122 (0.084)	-0.098 (0.103)	0.043 (0.105)
Initial Emp. Share in Routine Occ.	0.214 (0.147)	0.481** (0.228)	0.716*** (0.256)	0.059 (0.135)	0.150 (0.224)	0.235 (0.229)
Avg. Offshorability Index of Initial Occ.	-1.562 (0.955)	-3.826** (1.689)	-5.01*** (1.769)	0.295 (1.083)	-0.827 (1.370)	-2.171 (1.421)

Notes: N=340 for all regressions. All regressions use the preferred specification with Census division dummies and the full set of controls from Autor et al. (2013), listed below Table C.2. Observations are at the commuting zone- level and weighted by base-year population. Robust standard errors are clustered by state and reported in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table C.5. Robustness Check: Testing for Endogenous Changes in Population Composition

	Male Share	Age	White Share	Black Share	Hispanic Share	Other Race Share
			Panel A: Full Population			
ΔIPW_i^{US}	-0.743	-0.106	-0.945	0.039	0.846	0.060
	(0.958)	(0.226)	(0.688)	(0.405)	(0.556)	(0.361)
			Panel B: Employed			
ΔIPW_i^{US}	-0.818	0.204	-1.432	0.132	0.957	0.343
	(0.831)	(0.367)	(0.918)	(0.372)	(0.745)	(0.301)
			Panel C: Employed Wage Earners			
ΔIPW_i^{US}	-1.040	0.243	-1.690*	0.090	1.388*	0.212
	(0.873)	(0.325)	(0.984)	(0.379)	(0.834)	(0.350)
			Panel D: Unemployed			
ΔIPW_i^{US}	0.289	1.748*	1.504	-0.485	4.098	-5.117
	(3.183)	(1.056)	(1.850)	(2.034)	(2.901)	(3.668)
			Panel E: Homemakers			
ΔIPW_i^{US}	0.118	0.686	1.765	0.628	-2.451	0.058
	(0.396)	(0.599)	(2.447)	(0.622)	(2.474)	(0.612)

Notes: N=340 for the regressions for full population, employed, and employed wage earners, N=314 for unemployed, and N=336 for homemakers. Observations are at the commuting zone level and are weighted by base-year share of national population. All outcome variables except for age are shares of the respective employment group self-identifying by the demographic trait and measured in percentage points. All regressions use the preferred specification with Census division dummies and the full set of controls from Autor et al. (2013), shown in Table 2. Robust standard errors are clustered by state and reported in parentheses. First stage results are available upon request. Employed wage earners include all employed not classified as self-employed. Unemployed persons are those indicated to be out of work (for either less than a year or more than a year), and exclude self-identified retirees, students, and homemakers. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table C.6. Robustness Check: Testing for Pre-trends in Health Correlated with Import Exposure

	Avg Poor Mental Days/Month	Share w/Poor Mental Days (ppts)	Avg Poor Physical Days/Month	Share w/Poor Physical Days (ppts)	Share w/ Fair or Poor General Health
			Panel A: Full Population		
ΔIPW_i^{US}	-0.395**	-1.583	-0.204	-0.591	-0.501
	(0.179)	(1.049)	(0.122)	(1.156)	(0.589)
			Panel B: Employed		
ΔIPW_i^{US}	-0.338*	-2.966*	0.148	-0.090	0.682
	(0.199)	(1.565)	(0.111)	(0.979)	(0.692)
			Panel C: Employed Wage Earners		
ΔIPW_i^{US}	0.330*	-3.106**	0.192	0.025	0.544
	(0.186)	(1.460)	(0.130)	(1.158)	(0.703)
			Panel D: Unemployed		
ΔIPW_i^{US}	0.378	2.314	-0.422	4.951	-2.759
	(0.813)	(3.560)	(0.591)	(4.556)	(2.970)
			Panel E: Homemakers		
ΔIPW_i^{US}	-0.661*	1.470	0.069	0.488	-0.373
	(0.344)	(2.430)	(0.431)	(2.462)	(1.276)

Notes: N=186 for the regressions for the unemployed and N=200 for all others. Observations are at the commuting zone level and are weighted by base-year share of national population. All regressions use the preferred specification with Census division dummies and the full set of controls from Autor et al. (2013), listed below Table 2. Robust standard errors are clustered by state and reported in parentheses. First stage results are available upon request. Employed wage earners include all employed not classified as self-employed. Unemployed persons are those indicated to be out of work (for either less than a year or more than a year), and exclude self-identified retirees, students, and homemakers. * p<0.10, ** p<0.05, *** p<0.01

Table C.7. Robustness Check: Pre-period Health Outcomes as a Covariate

	Avg Poor Mental Days/Month	Share w/Poor Mental Days (ppts)	Avg Poor Physical Days/Month	Share w/Poor Physical Days (ppts)	Share w/ Fair or Poor General Health
			Panel A: Full Population		
ΔIPW_i^{US}	0.134	0.408	0.067	0.787	0.746*
	(0.096)	(0.491)	(0.077)	(0.543)	(0.402)
Pre-Period Health Outcome	-0.237***	-0.331***	-0.131*	-0.237***	-0.264***
	(0.066)	(0.063)	(0.078)	(0.075)	(0.078)
			Panel B: Employed		
ΔIPW_i^{US}	0.256**	0.985	0.058	1.084	0.148
	(0.124)	(0.783)	(0.103)	(0.681)	(0.289)
Pre-Period Health Outcome	-0.094	-0.305***	-0.090	-0.238***	-0.239
	(0.080)	(0.061)	(0.103)	(0.082)	(0.147)
			Panel C: Employed Wage Earners		
ΔIPW_i^{US}	0.277**	1.155	0.056	1.489**	0.190
	(0.122)	(0.792)	(0.082)	(0.709)	(0.301)
Pre-Period Health Outcome	-0.055	-0.301***	-0.109	-0.247***	-0.240*
	(0.084)	(0.066)	(0.089)	(0.078)	(0.143)

Notes: N=311 for all regressions for full population, employed, and employed wage earners. Pre-period health outcomes are the average values for 1998 and 1999. Observations are at the commuting zone level and are weighted by base-year share of national population. All regressions use the preferred specification with Census division dummies and the full set of controls from Autor et al. (2013), listed below Table 2. Robust standard errors are clustered by state and reported in parentheses. First stage results are available upon request. Employed wage earners include all employed not classified as self-employed. * p<0.10, ** p<0.05, *** p<0.01

Figure C.1. Distribution of Health Measures Across Commuting Zones

CONCLUSION

This dissertation is comprised of three studies in economics. The first chapter studies the effects of social scrutiny on behaviors that signal academic ability using a new laboratory experiment. Subjects in the experiment are paid for correctly answering verbal analogy questions taken from either a harder or easier set. Payments are carefully calibrated so that high-ability participants tend to make more money attempting the hard questions while low-types make more attempting the easy questions. This makes choosing the hard questions a credible signal of academic ability, and the results of the experiment support this. In treatment, subjects reveal to others in the experiment which questions they attempted. Contrary to predictions from the social psychology literature, we find that social observation pushes subjects away from choosing the signal of intelligence, despite evidence that it was privately valuable. These results suggest that, in some settings, social scrutiny may push people away from behaviors that signal their intelligence and ability.

The second chapter of this dissertation is an empirical study of the effects of trade shocks on worker safety and health. Our theoretical model explains that firms shift resources away from worker safety and into investments to increase output when they face greater competition and risk of shutdown. The model therefore predicts that import shocks, robustly shown to decrease firm survival, increase worker injury rates, especially at the smallest and least productive firms which are at the greatest risk of being pushed out of the market. To test this prediction empirically, we identify the effects of Chinese import shocks in the US in 1996-2007 on worker injury rates at US manufacturers in competing industries. The results demonstrate that injury rates rise in industries hit with import competition, particularly at smaller establishments, and the effects persist over at least five years. The magnitude of the effects and the implied welfare implications are significant: per our back of the envelope exercise, injury rates increase by 13 percent on average at the smallest establishments in the sample, and this costs worker welfare the equivalent of a 1 to 2 percent fall in their wage.

The final chapter studies the effects of import competition on broader health measures throughout affected regions. We find that Chinese import exposure in US local labor markets significantly worsens average mental, physical, and general health. The effects are greatest for self-reported mental health and among the employed in affected regions, consistent with evidence from the health literature on the consequences of job insecurity and wage losses. Further, we find that import exposure significantly increases the share of a region's population which is unable to afford medical care. Together, the latter two chapters of this dissertation provide some of the first evidence that trade shocks affect welfare through changes in physical and mental health, as well as access to healthcare. These findings have implications for trade, health, and regulatory policy, and they motivate further studies on the non-pecuniary effects of trade shocks.

REFERENCES

Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman. 2008. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. Journal of Public Economics 92(5): 1047-60.
Amiti, Mary, and Donald R. Davis. 2012. Trade, firms, and wages: Theory and evidence. The Review of Economic Studies 79(1): 1-36.
Andreoni, James and B. Douglas Bernheim. 2009. Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. Econometrica 77(5): 1607-36.
Andreoni, James and Justin M. Rao. 2011. The power of asking: How communication affects selfishness, empathy, and altruism. Journal of Public Economics 95(7): 513-20.
Ariely, Dan, Anat Bracha and Stephan Meier. 2009. Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. American Economic Review 99(1): 544-55.
Ariely, Dan, Uri Gneezy, George Loewenstein and Nina Mazar. 2009. Large Stakes and Big Mistakes. Review of Economic Studies 76(2): 451-69.
Arkolakis, C., Costinot A. and Rodriguez-Clare A. 2012. New Trade Models, Same Old Gains? American Economic Review, 2012, 102(1), 94-130
Ashraf, Nava, Oriana Bandiera, and Scott Lee. 2014. Awards Unbundled: Evidence from a Natural Field Experiment. Journal of Economic Behavior and Organization 100: 44-63.
Autor, David H., David Dorn, and Gordon H. Hanson. 2013. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. American Economic Review 103(6): 2121-68.
Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song. 2014. Trade Adjustment: Worker Level Evidence. NBER Working Paper 19226.
Baker, Dean B. 1985. The Study of Stress at Work. Annual Review of Public Health 6: 367-81.
Bandiera, Oriana, Iwan Barnkay and Imran Rasul. 2005. Social Preferences and the Response to Incentives: Evidence from Personnel Data. Quarterly Journal of Economics 120(3): 917-62.
---. 2010. Social Incentives in the Workplace. Review of Economic Studies 77(2): 417-58.
---. 2013. Team Incentives: Evidence From a Firm Level Experiment. Journal of the European Economic Association 11(5): 1079-1114.
Barattieri, Alessandro, Susanto Basu, and Peter Gottschalk. 2014. Some Evidence on the Importance of Sticky Wages. American Economic Journal: Macroeconomics 6: 70-101.
Baumeister, Roy F. 1998. The Self. In The Handbook of Social Psychology, 4th Ed, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, 680-740. New York: McGraw-Hill.
Beck, Thorsten, Asli Demirguc-Kunt and Vojislav Maksimovic. 2005. Financial and Legal Constraints to Growth: Does Firm Size Matter? Journal of Finance 60(1): 137-77.

Bem, Daryl J. 1972. Self-perception Theory. Advances in Experimental Social Psychology 6: 1-62.
Bénabou, Roland and Jean Tirole. 2003. Intrinsic and Extrinsic Motivation. Review of Economic Studies 70(3): 489–520.
---. 2004. Willpower and Personal Rules. Journal of Political Economy 112(4): 848-87.
---. 2006. Incentives and Prosocial Behavior. American Economic Review 96(5): 1652–78.
---. 2002. Self-Confidence and Personal Motivation. Quarterly Journal of Economics 117(3): 871-915.
Bernard, Andrew B, J Bradford Jensen and Peter K. Schott. 2006a. Survival of the Best Fit: Exposure to Low-Wage Countries and the (Uneven) Growth of US Manufacturing Plants. Journal of International Economics 68(1): 219-237.
---. 2006b. Trade Costs, Firms and Productivity. Journal of Monetary Economics 53(5): 917-37.
Bernheim, Doublas B. 1994. A Theory of Conformity. Journal of Political Economy 102(5): 841-77.
Blanchard, Olivier Jean and Lawrence F. Katz. 1992. Regional Evolutions. Brookings Papers on Economic Activity 1:1-75.
Blanchflower, David G. and Andrew J. Oswald. 2004. Well-Being Over Time in Britain and the USA. Journal of Public Economics 88(7): 1359-86.
---. 2008. Hypertension and Happiness Across Nations. Journal of Health Economics 27(2): 218-33.
Bloom, Nicholas and John can Reenen. 2007. Measuring and Explaining Management Practices Across Firms. Quarterly Journal of Economics 122(4):1351-1408.
Bloom, Nicholas, Mirko Draca, and John van Reenen. 2012. Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. NBER Working Paper No. 16717
Bodner, Ronit and Drazen Prelec. 2003. Self-Signaling and Diagnostic Utility in Everyday Decision Making. In The Psychology of Economic Decisions. Volume I: Rationality and Well-Being, edited by Isabella Brocas and Juan D. Carrillo, 105-26. Oxford: Oxford University Press.
Bohle, Phillip, Michael Quinlan, and Claire Mayhew. 2001. The Health and Safety Effects of Job Insecurity: An Evaluation of the Evidence. The Economic and Labor Relations Review 12(1): 32-60.
Boone, Jan and Jan C. van Ours. 2006. Are Recessions Good for Workplace Safety? Journal of Health Economics 25(1): 1069-1093.
Boone, Jan, Jan C. van Ours, Jean-Philippe Wuellrich, and Josef Zweimuller. 2011. Recessions are Bad for Workplace Safety. Journal of Health Economics 30(4): 764-73.
Brekke, Kjell Arne, Snorre Kverndokk, and Karine Nyborg. 2003. An economic model of moral motivation. Journal of Public Economics 87(9): 1967-83.
Brooker, Ann-Sylvia, John W. Frank, and Valerie S. Tarasuk. 1997. Back pain claim rates and the business cycle. Social Science & Medicine 45(3): 429-439.

Brown, Drusilla K. 2007. Globalization and Employment Conditions Study. World Bank. Social Protections Discussion Paper 708.
Brown, Drusilla K., Alan V. Deardorff, and Robert M. Stern. 1996. International Labor Standards and Trade: A Theoretical Analysis, in Jagdish Bhagwati, and Robert Hudec, eds., Fair trade and harmonization: Prerequisites for free trade? Cambridge, MA: MIT Press. 227-80
Bubonya, Melisa, Deborah A. Cobb-Clark, and Mark Wooden. 2014. A Family Affair: Job Loss and the Mental Health of Spouses and Adolescents. University of Melbourne Working Paper.
Cameron, Colin A., Jonah B. Gelbach and Douglas L. Miller. 2011. Robust Inference with Multiway Clustering. Journal of Business and Economics Statistics 29(2): 238-49.
Campbell III, Carl M. and Kunai S. Kamlani. 1997. The Reasons for Wage Rigidty: Evidence from a Survey of Firms. Quarterly Journal of Economics 112(3): 759-89.
Centers for Disease Control and Prevention (CDC). *Behavioral Risk Factor Surveillance System Survey Data*. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2000, 2007.
Chen, Yuyu, Ginger Zhe Jin, and Yang Yue. 2010. Peer Migration in China. NBER Working Paper 15671.
Cheng, Grand H-L. and Darius K-S. Chan. 2008. Who suffers more from job insecurity? A meta-analytic review. Applied Psychology 52(7): 272-303.
Cherry, Todd L., Peter Frykblom and Jason F. Shogren. 2002. Hardnose the Dictator. American Economic Review 92(4): 1218-21.
Chiquiar, Daniel. 2008. Globalization, Regional Wage Differentials, and the Stolper-Samuelson Theorem: Evidence from Mexico. Journal of International Economics 74(1): 70-93.
Choi, Don and Alan S. Blinder. 1990. A Shred of Evidence on Theories of Wage Stickiness. Quarterly Journal of Economics 105(4): 1003-1015.
Clark, Andrew. 2003. Unemployment as a Social Norm: Psychological Evidence from Panel Data. Journal of Labor Economics 21: 323-51.
Clark, Andrew, Andreas Knabe, and Steffen Ratzel. 2010. Boon or Bane? Others' Unemployment, Well-Being and Job Insecurity. Labour Economics 17(1): 52-61.
Dana, Jason, Roberto A. Weber and Jason Xi Kuang. 2007. Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. Economic Theory 33(1): 67-80.
Davidson, Russell and Jean-Yves Duclos. 2000. Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. Econometrica 68(6): 1435-64.
Davis, Donald R., and James Harrigan. 2011. Good jobs, bad jobs, and trade liberalization. Journal of International Economics 84(1): 26-36.
De Witte, Hans. 1999. Job insecurity and psychological well-being: Review of the literature and exploration of some unresolved issues. European Journal of Work and Organizational Psychology 8(2): 155-77.

DellaVigna, Stefano, John A. List and Ulrike Malmendier. 2012. Testing for Altruism and Social Pressure in Charitable Giving. Quarterly Journal of Economics 127(1): 1-56.
Demidova, Svetlana. 2008. Productivity Improvements and Falling Trade Costs: Boone or Bane? International Economic Review 49(4): 1437-62.
Demidova, Svetlana and Andrés Rodriguez-Clarez. 2013. The Simple Analytics of the Melitz Model in a Small Open Economy. Journal of International Economics 90(2): 266-72.
Ebenstein, Avraham, Ann Harrison, Margaret McMillan, and Shannon Phillips. 2014. Estimating the Impact of Trade and Offshoring on American Workers Using the Current Population Surveys. Review of Economics and Statistics 96(4): 581-95.
Egger, Hartmut and Udo Kreickemeier. 2009. Firm Heterogeneity and the Labor Market Effects of Trade Liberalization. International Economic Review 50(1): 187-215.
Ellingsen, Tore and Magnus Johannesson. 2008. Pride and Prejudice: The Human Side of Incentive Theory. American Economic Review 98(3): 990-1008.
Ettner, Susan L. 1996. New evidence on the relationship between income and health. Journal of Health Economics 15(1): 67-85.
Evans, Mary F. and V. Kerry Smith. 2006. Do we really understand the age-VSL relationship? Resource and Energy Economics 28: 242-61.
---. 2010. Measuring how risk tradeoffs adjust with income. Journal of Risk and Uncertainty 40: 33-55.
Feenstra, Robert C., John Romalis, and Peter K. Schott. 2002. US Imports, Exports, and Tariff Data, 1989-2001. NBER Working Paper No. w9387.
Fischbacher, Urs. 2007. Z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10(2): 171-78.
Frey, Bruno S. and Felix Oberholzer-Gee. 1997. The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding Out. American Economic Review 87(4): 746-55.
Frijters, Paul, John P. Haisken-DeNew, and Michael A. Shields. 2005. The causal effect of income on health: Evidence from German reunification. Journal of Health Economics 24(5): 997-1017.
Funk, Patricia. 2010. Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System. Journal of the European Economic Association 8(5): 1077-103.
GAO. 2009. Workplace Safety and Health: Enhancing OSHA's Records Audit Process Could Improve the Accuracy of Worker Injury and Illness Data. United States Government Accountability Office. Report to Congressional Requesters: October 2009.
Geishecker, Ingo, Maximilliam Riedl, and Paul Frijters. 2012. Offshoring and Job Loss Fears: An Econometric Analysis of Individual Perceptions. Labour Economics 19(5): 739-47.
Glaeser, Edward L. and Joseph Gyourko. 2005. Urban Decline and Durable Housing. Journal of Political Economy 113(2): 345-75.
Glazer, Amihai and Kai A. Konrad. 1996. A Signaling Explanation for Charity. American Economic Review 86(4): 1019-28.

Gneezy, Uri and Aldo Rustichini. 2000a. A Fine is a Price. Journal of Legal Studies 29(1): 1-17.
---. 2000b. Pay Enough or Don't Pay at All. Quarterly Journal of Economics 115(3): 791-810.
Gottfredson, Linda S. 1997. Why g Matters: The Complexity of Everyday Life. Intelligence 24(1): 79-132
Greenland, Andrew and John Lopresti. 2013. The Good News About Disappearing Jobs: U.S. High School Dropout Rates and Import Exposure. FREIT Working Paper 585.
Greiner, Ben. 2004. The Online Recruitment System ORSEE 2.0- A Guide for the Organization of Experiments in Economics. Working Paper, University of Cologne.
Grossman, Zack. 2012. Self-Signaling versus Social-Signaling in Giving. Working Paper, University of California Santa Barbara.
Hamilton, Lee V., Clifford L. Broman, William S. Hoffman, and Deborah S. Renner. 1990. Hard Times and Vulnerable People: Initial Effects of Plant Closing on Autoworkers' Mental Health. Journal of Health and Social Behavior 31(2): 123-140.
Handley, Kyle and Nuno Limão. 2013. Policy Uncertainty, Trade and Welfare: Theory and Evidence for China and the US. NBER Working Paper No. 19376.
Harbaugh, William T. 1998. The Prestige Motive for Making Charitable Transfers. American Economic Review 88(2): 277-82.
Helpman, Elhanan, Oleg Itskhoki and Stephen Redding. 2010. Inequality and Unemployment in a Global Economy. Econometrica 78(4): 1239-83.
---. 2011. Trade and Labor Market Outcomes. NBER Working Paper No. 16662.
Holmås, Tor Helge, Egil Kjerstad, Hilde Lurås and Odd Rune Straume. 2010. Does monetary punishment crowd out pro-social motivation? A natural experiment on hospital length of stay. Journal of Economic Behavior and Organization 75(2): 261-67.
Holmes, Thomas and John Stevens. 2014. An Alternative Theory of the Plant Size Distribution, with Geography and Intra- and International Trade. Journal of Political Economy 122(2): 369-421.
Hummels, David, Jakob Munch and Chong Xiang. 2014. No Pain, No Gain: The Effects of Globalization on Job Injury and Sickness. Working Paper.
Karlan, Dean and Margaret A. McConnell. 2013. Hey Look at Me: The Effect of Giving Circles on Giving. Working Paper.
Karsten, Paul and Klaus Moser. 2009. Unemployment impairs mental health: Meta-analyses. Journal of Vocational Behavior 74(3): 264-82.
Kassenboehmer, Sonja C. and John P. Haisken-DeNew. 2009. You're Fired! The Causal Negative Effect of Entry Unemployment on Life Satisfaction. Economic Journal 119: 448-462.
Knieser, Thomas J., W. Kip Viscusi, and James P. Zilliak. 2010. Policy relevant heterogeneity in the value of statistical life: New evidence from panel data quantile regressions. Journal of Risk and Uncertainty 40: 15-31.

Koch, Alexander K. and Hans-Theo Normann. 2008. Giving in Dictator Games: Regard for Others or Regard by Others? Southern Economic Journal 75(1): 223-31.
Köszegi, Botond. 2006. Ego Utility, Overconfidence, and Task Choice. Journal of the European Economic Association 4(4): 673-707.
Kovak, Brian K. 2013. Regional Effects of Trade Reform: What is the Correct Measure of Liberalization? American Economic Review 103(5): 1960-1976.
Kreps, David M. 1997. Intrinsic Motivation and Extrinsic Incentives. American Economic Review 87(2): 359-64.
Lacetera, Nicola and Mario Macis. 2010. Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. Journal of Economic Behavior and Organization 76(2): 225-37
Lazear, Edward P., Kathryn L. Shaw, and Christopher Stanton. 2013. Making Do With Less: Working Harder During Recessions. NBER Working Paper No. 19328.
Lazear, Edward, Ulrike Malmendier and Roberto Weber. 2012. Sorting, Prices, and Social Preferences. American Economic Journal: Applied Economics 4(1): 136-63.
Leigh, J. Paul. 2011. Economic Burden of Occupational Injury and Illness in the United States. Millbank Quarterly 89(4): 728-772
Li, Norman P., J. Michael Bailey, Douglas T. Kenrick, and Joan A. W. Linsenmeier. 2002. The necessities and luxuries of mate preference: Testing the tradeoffs. Journal of Personality and Social Psychology 82(6): 947-55.
Liem, Ramsay and Joan Huser Liam. 1988. Psychological Effects of Unemployment on Workers and Their Families. Journal of Social Issues 44(4): 87-105.
Lindahl, Mikael. 2005. Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income. Journal of Human Resources 40(1): 144-68.
Lindbeck, Assar. 1997. Incentives and Social Norms in Household Behavior. American Economic Review 87(2): 370-77.
Lindbeck, Assar, Sten Nyberg and Jörgen W. Weibull. 1999. Social Norms and Economic Incentives in the Welfare State. Quarterly Journal of Economics 114(1): 1-35.
List, John A. and Todd L. Cherry. 2000. Learning to Accept in Ultimatum Games: Evidence from an Experimental Design that Generates Low Offers. Experimental Economics 3: 11-29.
Luechinger, Simon, Stephan Meier, and Alois Stutzer. 2010. Why Does Unemployment Hurt the Employed? Evidence from the Life Satisfaction Gap Between the Public and Private Sector. Journal of Human Resources 45(4): 998-1045.
Marcus, Jan. 2013. The effect of unemployment on the mental health of spouses – Evidence from plant closures in Germany. Journal of Health Economics 32(3): 546-558.
Mas, Alexandre and Enrico Moretti. 2009. Peers at Work. American Economic Review 99(1): 112-45.

McKee-Ryan, Frances, Zhaoli Song, Connie R. Wanberg, and Angelo J. Kinicki. 2005. Psychological and Physical Well-Being During Unemployment: A Meta-Analytic Study. Journal of Applied Psychology 90(1): 53-76.
McLaren, John and Shushanik Hakobyan. 2012. Looking for Local Labor-Market Effects of the NAFTA. NBER Working Paper 16535.
McManus, T. Clay and Georg Schaur. 2015. The Effects of Import Competition on Worker Health. University of Tennessee Working Paper.
Meer, Jonathon, Douglas L. Miller, and Harvey S Rose. 2003. Exploring the health-wealth nexus. Journal of Health Economics 22(5): 713-30.
Meier, Stephan. 2007. Do Subsidies Increase Charitable Giving in the Long Run? Matching Donations in a Field Experiment. Journal of the European Economic Association 5(6): 1203-22.
Melitz, Marc J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Productivity. Econometrica 71(6):1695-1725.
Melitz, Marc J. and Stephen Redding. 2014. New Trade Models, New Welfare Implications. American Economic Review, forthcoming
Mellström, Carl and Magnus Johannesson. 2008. Crowding Out in Blood Donation: Was Titmuss Right? Journal of the European Economic Association 6(4): 845-63.
Miller, Douglas L., Marianne E. Page, Ann Huff Stevens and Mateusz Filipski. 2009. Why Are Recessions Good for Your Health? American Economic Review: Papers & Proceedings 99(2): 122-7.
Murnighan, J. Keith, John M. Oesch, and Madan Pillutia. 2001. Player Types and Self-Impression Management in Dictatorship Games: Two Experiments. Games and Economic Behavior 37(2): 388-414.
Murphy, Nora A. 2007. Appearing Smart: The Impression Management of Intelligence, Person Perception Accuracy, and Behavior in Social Interaction. Personality and Social Psychology Bulletin 33: 325-39.
Notowidigdo, Matthew J. 2013. The Incidence of Local Labor Demand Shocks. Northwestern University Working Paper.
Oxoby, Robert J. and John Spraggon. 2008. Mine and yours: Property rights in dictator games. Journal of Economic Behavior and Organization 65: 703-13.
Park, Bernadette and Sue Kraus. 1992. Consensus in initial impressions as a function of verbal information. Personality and Social Psychology Bulletin 18: 439-46.
Pierce, Justin R. and Peter K. Schott. 2012. A Concordance Between Ten-Digit US Harmonized System Codes and SIC/NAICS Product Classes and Industries. Journal of Economic and Social Measurement 37(1): 61-96.
---. 2013. The Surprisingly Swift Decline of US Manufacturing Employment. CESifo Working Paper No. 4563.
Probst, Tahira M. and Ty L. Brubaker. 2001. The Effects of Job Insecurity on Employee Safety Outcomes: Cross-Sectional and Longitudinal Explorations. Journal of Occupational Health Psychology 6(2): 139-59.
Prokosch, Mark D, Richard G. Coss, Joanna E. Scheib, and Shelley A. Blozis. 2009. Intelligence and mate choice: intelligent men are always appealing. Evolution and Human Behavior 30(1): 11-20.

Ruhm, Christopher J. 2000. Are Recessions Good for Your Health? Quarterly Journal of Economics 115(2): 617-50.
---. 2003. Good times make you sick. Journal of Health Economics 24(4): 637-58.
---. 2005. Healthy living in hard times. Journal of Health Economics 24(2): 341-63.
Salm, Martin. 2009. Does job loss cause ill health? Health Economics 18(9): 1075-1089.
Shapiro, Carl and Joseph E. Stiglitz. 1984. Equilibrium Unemployment as a Worker Discipline Device. American Economic Review 74(3): 433-44.
Soetevent, Adriaan R. 2005. Anonymity in giving in a natural context- a field experiment in 30 churches. Journal of Public Economics 89(11-12): 2301-23.
Sokejima, Shigeru and Sadanobu Kagamimori. 1998. Working Hours as a Risk Factor for Acute Myocardial Infarction in Japan. Case-Control Study. British Medical Journal 317(7161): 775-80.
Solomon, Susan and Leonard Saxe. 1977. What Is Intelligent, As Well As Attractive, Is Good. Personal and Social Psychology Bulletin 3: 670-73.
Staiger, Douglas and James H. Stock. 1997. Instrumental Variables Regression with Weak Instruments. Econometrica 65(3): 557-86.
Sullivan, Daniel and Till von Wachter. 2009. Job Displacement and Mortality: An Analysis Using Administrative Data. Quarterly Journal of Economics 124(3): 1265-1306.
Svenson, Ola. 1981. Are we all less risky and more skillful than our fellow drivers? Acta Psychologica 47: 143-48.
Sverke, Magnus, Johnny Hellgren, and Katharina Naswall. 2002. No Security: A Meta-Analysis and Review of Job Insecurity and Its Consequences. Journal of Occupational Health Psychology 7(3): 242-64.
Syverson, Chad. 2010. What determines productivity? NBER Working Paper No. 15712.
Tonin, Mirco and Michael Vlassopoulos. 2013. Experimental evidence of self-image concerns as motivation for giving. Journal of Economic Behavior and Organization 90: 19-27
Topalova, Petia. 2010. Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. American Economic Journal: Applied Economics 2(4): 1-41.
Viscusi, Kip. 1983. Risk by Choice: Regulating Health and Safety in the Workplace. Cambridge: Harvard University Press.
Wilson, Stephan M., Jeffry H. Larson, and Katherine L. Stone. 1993. Stress Among Job Insecure Workers and Their Spouses. Family Relations 42(1): 74-90.

VITA

Thomas "Clay" McManus was born to Robert and Holly in Atlanta, Georgia on April 4, 1988. He was raised in Knoxville, Tennessee and graduated from West High School in 2006. He then moved to Charleston, South Carolina where he attended College of Charleston until graduating cum laude in May 2010 with a Bachelor of Science in Economics with honors from the Honors College and the Honors Program in Business and Economics. He moved back to Knoxville, Tennessee later that year to begin his graduate studies at the University of Tennessee. He received his Master of Arts degree in Economics in December 2011 and his Doctor of Philosophy degree in Economics in August 2015. His work in experimental and behavioral economics has been supported by external funding from the Russell Sage Foundation, and the first paper, the first chapter of this dissertation, was recently published in the *Journal of Economic Behavior and Organization*. His other work focusses on international trade, and specifically the effects of import competition on physical and mental health. He has presented his research at conferences across the United States, and recently at the Empirical Investigations in Trade and Investment conference in Bali, Indonesia. He has taught undergraduate and graduate students at the University of Tennessee and the University of Iowa, and received teaching awards from the Department of Economics and the Haslam College of Business at the University of Tennessee. This fall, he and his new wife, Ashley, will move to Cincinnati, Ohio, where he has accepted a position as Assistant Professor in Economics at Xavier University.