

NBER WORKING PAPER SERIES

PATH DEPENDENCE IN THE LABOR MARKET:
THE LONG-RUN EFFECTS OF EARLY CAREER OCCUPATIONAL EXPERIENCE

Jesse Bruhn
Jacob Fabian
Luke Gallagher
Matthew Gudgeon
Adam Isen
Aaron R. Phipps

Working Paper 34463
<http://www.nber.org/papers/w34463>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2025

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect the views or the official positions of the U.S. Department of the Treasury or of the National Bureau of Economic Research. Any taxpayer data used in this research was kept in a secured Treasury or IRS data repository, and all results have been reviewed to ensure that no confidential information is disclosed. Further, all views and findings expressed in this manuscript are those of the authors and do not represent the views or official positions of the United States Military Academy, Department of Defense, or the United States Army. This paper has also not been funded or endorsed by ISO New England. Research reported in this publication was supported by the Population Studies and Training Center at Brown University through the generosity of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P2C HD041020) and the Russell Sage Foundation. We thank seminar participants at the Federal Reserve Board, Northeastern University, Opportunity Insights, Columbia University Teachers College, Clark University, the Northeast Econ of Ed Workshop, the San Francisco Federal Reserve Micro/Macro Labor Workshop, Northwestern University, University of Chicago, Purdue University, the Cowles Summer Workshop, the SOLE annual meeting, Boston University, the Atlanta Federal Reserve Employment Conference, and the NBER Summer Institute Labor Studies session for insightful comments and feedback. All errors remain ours. Jesse Bruhn: Department of Economics, Brown University. Jacob Fabian, Market Development, ISO New England. Luke Gallagher: Department of Social Sciences, West Point. Matthew Gudgeon: Department of Economics, Tufts University. Adam Isen: Department of Economics, Johns Hopkins University. Aaron Phipps. Department of Social Sciences, West Point.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Jesse Bruhn, Jacob Fabian, Luke Gallagher, Matthew Gudgeon, Adam Isen, and Aaron R. Phipps. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Path Dependence in the Labor Market: The Long-run Effects of Early Career Occupational Experience

Jesse Bruhn, Jacob Fabian, Luke Gallagher, Matthew Gudgeon, Adam Isen, and Aaron R. Phipps

NBER Working Paper No. 34463

November 2025

JEL No. J20, J24, J3, J38, J4, J45

ABSTRACT

We study the causal effect of different early career occupational experiences on labor market outcomes. To do so, we pair over two decades of administrative tax data with internal personnel records from one of the largest employers of young adults in the United States: the US Army. Enlistees work in a diverse and varied set of occupations, including non-combat roles like mechanics, legal services, financial specialists, cooks, dental hygienists, police officers, and network/computer specialists. Occupational eligibility is determined by test score cutoffs which we leverage in a series of 35 regression discontinuity designs. We find that a typical early career occupational experience generates a substantial amount of path dependence, with point estimates that suggest a 19p.p. increase in the likelihood of being observed in an identical or closely related occupation as much as 20 years later. The corresponding impact of different occupations on earnings are highly heterogeneous, yet predictable: long-run changes in the average earnings of the occupations applicants are pushed into, and pulled out-of, can explain over 60% of the causal variation across cutoffs, with point estimates that suggest improvements in occupational earnings premia translate dollar-for-dollar into economic success. Taken together, our results highlight the importance of early career occupational experience as a key channel for promoting long-run well-being among young adults who are not college bound.

Jesse Bruhn
Brown University
Department of Economics
and NBER
jesse_bruhn@brown.edu

Jacob Fabian
ISO New England
jfabian@iso-ne.com

Luke Gallagher
United States Military Academy
at West Point
luke.gallagher@westpoint.edu

Matthew Gudgeon
Tufts University
Department of Economics
and NBER
matthew.gudgeon@tufts.edu

Adam Isen
Johns Hopkins University
adam.isen@gmail.com

Aaron R. Phipps
United States Military Academy
at West Point
aaron.phipps@westpoint.edu

1 Introduction

Early life occupational experiences are strongly related to long-run economic success. Every year, roughly three-million young Americans will enter the labor market for the first time.¹ For a typical worker, the following period is critical: over 66% of lifetime wage growth will occur during the first 10 years of their career (Topel and Ward, 1992; Rubinstein and Weiss, 2006; Wachter, 2020). This suggests an important connection between early career occupational experiences and sustained economic well-being, with potentially wide ranging implications for the design of the education system, the sources of inter-generational earnings transmission, and the role that high-quality job opportunities have in reducing economic inequality. In particular, understanding which career features lead to prosperity, and whether it is possible to cultivate durable connections to occupations with those features, is critical for the effective design of active labor market policy (e.g. Katz et al., 2022).

In this paper, we study the causal effect of different early career occupations on long-run occupational choice and earnings. There is little quasi-experimental evidence on this question. This is because it is challenging to isolate the causal influence of occupation separately from other worker and firm characteristics. Simple observational comparisons exhibit clear scope for selection: workers with higher potential earnings likely choose higher paying careers. Existing work typically relies on survey data and leverages job transitions among observably similar workers for identification (e.g. Deming and Noray, 2020; Kambourov and Manovskii, 2009a,b; Topel and Ward, 1992). Natural experiments with transparent, quasi-random variation are hard to find, let alone variation that spans a broad variety of careers. The scarcity of quasi-experimental estimates is also due, in part, to the fact that it is rare to have high-quality administrative data that contains any reliable measure of occupation *at all*.²

To overcome these challenges, we pair over two decades of administrative tax data with internal personnel records from one of the largest employers of young adults in the country: the United States Army. The US military is a massive employer, with a combined labor force larger than the population of Chicago, IL (2.86 million workers in 2022). Managing an organization of this magnitude requires a diverse and varied set of roles. Some of these roles—such as infantry and field artillery—are explicitly connected to the war-fighting mission. However, the majority of Army soldiers serve in non-combat occupations such as legal services, financial specialists, cooks, dental hygienists, police officers, network/computer specialists, mechanics, and many others.³ Thus, military occupational experiences frequently have direct civilian analogues.

Military careers also represent the rare example of a widely available pathway to opportunity for young adults who are not college bound (Greenberg et al., 2022). At this moment in time, there are currently 7 million working age male veterans, which equates to roughly 1 in 10 full time male workers in the US labor force (U.S. Bureau of Labor Statistics, 2025). The vast majority of these veterans entered the military without a four-year degree (Greenberg et al., 2022). Thus, the military plays a dominant role in the early career

¹We estimate this by multiplying the number of children of minimum working age (≈ 4 million aged 14 in 2022) by the age 25-54 labor force participation rate (83.3% in 2022), which will be a lower bound in steady state.

²This literature frequently relies on publicly available surveys like the National Longitudinal Survey of Youth (as in Deming and Noray, 2020) or the Current Population Survey (as in Wachter, 2020).

³About 60% of soldiers in our data have non-combat roles. See appendix table F.1 for a break-down of military occupations across several broad categories known as “Career Management Fields” among first term soldiers.

development of skills and credentials for a huge swathe of the non-college labor market. Examining early career occupational experiences in the military is therefore critical for understanding the well-being of the non-college educated, particularly young men, a population that is currently struggling both economically and socially (Case and Deaton, 2022; Reeves, 2022; Autor et al., 2020).

The Army determines occupational eligibility via a unique natural experiment. In order to enlist, applicants to the military must take a test known as the Armed Services Vocational Aptitude Battery (ASVAB). Sub-components of this test are then reformulated into 10 test scores known as “line-scores,” which contain cutoffs that determine eligibility for specific occupations. For causal identification, we leverage these line-score thresholds in a series of 35 regression discontinuity designs. Applicants induced into an Army occupation receive training in the relevant skills, followed by anywhere from 2-5 years of practical, on-the-job experience. Thus, by comparing applicants that barely cleared an eligibility threshold for a given occupation to those that did not, we can isolate variation in occupational training and experience that is orthogonal to other observable and unobservable confounders.

Before we turn to results, there are two aspects of the research design that are worth highlighting. First, we note that while the discontinuities yield precise estimates for many outcomes, such as the long-run impact on career trajectory, they do not *individually* yield precise estimates for certain other outcomes, such as earnings. Thus, our preferred method of summarizing the variation will be to increase power by reporting *average* relationships across cutoffs. Second, we note that crossing individual cutoffs often “unlocks” more than one military occupation; moreover, we also find that the counterfactual careers applicants are “pulled” out-of vary across cutoffs. Taking into account these differences in the long-run careers applicants are “pushed” into and “pulled” out-of will be important, especially for understanding how early career experiences shape earnings. Our preferred approach will consist of building natural benchmarks that allow us to answer interpretable questions like, “Do early career experiences that cause applicants to move from lower-earning occupations and into higher-earning occupations also improve long-run economic well-being?”

We begin the empirical analysis by studying the effects of early career occupational experiences on long-run career trajectories. We document substantial path dependence. For example, we find that applicants who enter a typical Army occupation as a result of crossing a threshold are, on average, 19p.p. more likely to be in an identical or closely related occupation as much as 20 years later, with little evidence of heterogeneity across cutoffs. Some of this effect is due to applicants who choose to make a career out of the Army. However, it is important to distinguish between path dependence that arises as the result of a durable connection with an employer, in this case the Army, versus path dependence that arises as a result of skills, credentials, or information acquired on the job. To examine this second set of mechanisms, we also produce estimates that explicitly exclude the “within-employer” retention channel by looking only at the impact of Army occupational experiences on career paths at non-Army employers. We find that the average early career experience causes a 12p.p. increase in the likelihood of continuing on that career trajectory at a non-Army employer up to 20 years later. This suggests that at least two-thirds of our baseline effect is due to the skills, credentialing, and/or information channels. We also find that the counterfactual careers applicants are “pulled” out-of in the long-run are quite heterogeneous across cutoffs. Interestingly, these long-run “pull” effects correspond closely to the military careers applicants are “pulled” out-of in the short-run. Overall, the pattern of findings

on career paths highlights how crossing a line-score cutoff creates changes in early career experience that have persistent and sizable impacts on long-run labor market trajectories.

Next, we show that different early career occupational experiences have consequential, yet highly heterogeneous, effects on long-run earnings. For example, crossing the EL93 cutoff, which is linked to Army occupations related to signal support (Computers/Electronics), leads to early career experiences that result in long-run earnings increases of \$11,560 (se = 7638) dollars per year. On the other hand, crossing the CL95 cutoff, which is linked to careers in the quartermaster (i.e. warehousing) and human resources fields, leads to early career experiences that result in long-run earnings *reductions* of \$13,600 (se = 5336) per year. More formally, we can reject the joint null hypothesis that the impact of these early career experiences on earnings are identical across cutoffs ($P = 0.0004$). Thus, to better characterize the heterogeneity in terms of the consequential changes in long-run career path, we benchmark the earnings effects to the causal difference in average earnings between the occupations that applicants were pulled out-of and pushed in-to in the long-run.⁴

We find that causal effects on earnings align closely with the corresponding causal effects on occupational average wages. Thus, Army occupations which, 11-20 years after enlistment, have shifted individuals out of careers with low average earnings and into careers with high average earnings also tend to cause large increases in realized earnings. Moreover, we find that the differences in these implied earnings premia can explain *over* 60% of the variation across cutoffs in the causal effects on actual earnings, with a slope coefficient that implies each dollar of improvement in occupational wage premia translates essentially 1-for-1 into increases in earnings. This is consistent with the realized average impact on earnings primarily reflecting average occupational wage differences rather than general human capital or match quality. This finding also suggests that the simple and commonly used heuristic of comparing average earnings differences across occupations does, indeed, accurately predict the average impact of a career move.

We find that many of the other observable changes caused by early career experience are less predictive of long-run well-being. Cutoff-induced changes in average industry earnings, likelihood of working in a government job, continuing a career in the military, occupational skill content, credentials, length/intensity of training, and college completion *do* explain some of the variation in the causal effect on earnings across cutoffs. However, occupational wage differences continue to be tightly linked to long-run improvements even conditional on the cutoff-induced changes in these other important outcomes. Other candidate outcomes, like deployment to a combat zone, receiving Veteran's Affairs Disability Compensation (VADC), and joining the Army at-all, have virtually no explanatory power and, consequently, conditioning on them does not change the rate at which occupational wage premia translate into average effects on earnings.

We conclude the empirical analysis by exploring in more detail the specific types of occupations that tend to put young adults on the pathway to success. To this point, we have focused on reporting causal effects on long-run outcomes averaged across cutoffs. However, this approach does not directly speak to the *specific types*

⁴To measure the implied changes in occupational wage premia, we assign each applicant the leave-out average earnings of their reported occupation and estimate the long-run causal effect of crossing each threshold on occupation-specific average earnings. Thus, if an applicant was shifted out of a career that typically paid \$30,000 per year, and into a career that typically paid \$40,000 per year, our measure of the earnings premia would be \$10,000 (see section 6.2 for more detail). We observe earnings and not wages; throughout the paper we use wages and earnings interchangeably.

of early career military occupations that improve economic well-being. Thus, we present two supplementary analyses aimed at shedding light as to which specific early career experiences in the military are most economically beneficial. Inspired by our finding that long-run occupational wage premia are closely connected to economic well-being, our first analysis explores the predictive power of changes in the *ex-ante* expected earnings implied by different military career paths. To measure ex-ante expected earnings, we calculate a simple average of the expected long-run earnings (i.e. 11-20 years later) of the civilian careers linked to each military occupation. This allows us to measure which cutoffs cause young adults to move across occupations with a higher or lower *expected* earnings trajectory. We find that early career experiences linked to occupations with a high-earnings career trajectory produce strong returns: on average, each dollar of expected earnings translates into 77 cents of actual increased earnings.

Second, we develop a partially-pooled, multiple endogenous variable, two-stage least squares model of early career occupational experience. Under strong, but testable, assumptions about the nature of treatment effect heterogeneity, this approach allows us to isolate the impact of groups of similar Army occupations relative to a common civilian counterfactual and with a reasonable degree of precision. The model reveals that Army careers in military occupations related to computers/electronics, combat/construction, and health/police/intelligence, all substantially improve long-run earnings relative to not joining the Army. However, Army careers for enlistees in “logistics” that often involve extensive work in warehouses, and that tend to be linked to below average paying civilian careers, lead to relatively lower earnings changes and do not improve long-run earnings relative to not joining the Army.

2 Related Literature

This paper contributes to the literature on occupational mobility and career dynamics by quasi-experimentally identifying the causal effect of different early career occupational experiences on long-run outcomes like career choice and earnings. Previous research has examined the importance of job mobility in wage growth and career development (Topel and Ward, 1992; Neal, 1998, 1999; Pavan, 2011; Abraham et al., 2024) and highlighted the role of occupation-specific human capital (e.g. Kambourov and Manovskii, 2009a,b; Roys and Taber, 2019; Sanders and Taber, 2012; Traiberman, 2019), intergenerational employment (Staiger, 2021; Haeck and Laliberté, 2025; Corak and Piraino, 2011), and occupational matching (e.g. Cullen et al., 2025).⁵ Much, but not all, of this literature relies on survey data to elicit and control for worker preferences over occupations.^{6,7} Other strands of related work abstract from specific occupations and, instead, focus on the importance of underlying tasks and skills (e.g. Acemoglu and Autor, 2011; Autor et al., 2003; Autor and Handel, 2013; Gathmann and Schönberg, 2010; Deming, 2017; Acemoglu and Restrepo, 2019, 2020). Our central contribution is to use large-scale, administrative data to estimate the causal impact of early career

⁵Cullen et al. (2025) is closely related: the authors find that ability mismatch between workers and jobs is an important contributor to early career outcomes within the Air Force. Interestingly, since our identifying variation comes from soldiers at the test score eligibility margin for their occupation, our results can be interpreted as applying to ‘underqualified’ soldiers by typical ability mismatch definitions. Even so, we find that these soldiers experience long-run earnings gains when shifted into higher paying occupations.

⁶An important exception to the reliance on survey data is Abraham et al. (2024), who use the LEHD to document how earnings varies descriptively with the average wage of the initial occupation. Consistent with our results, the authors find that initial occupations with higher average earnings are highly correlated with a higher-wage career trajectory.

⁷More recent work has used hypothetical choice and other modern survey instruments to isolate beliefs about careers and their connection to occupational decisions (e.g. Arcidiacono et al., 2020; Conlon and Patel, 2025).

occupational experiences.

We also add to a closely related literature documenting the impact of macroeconomic events on early career trajectories by separately isolating the impact of career choice from the broader influence of a recession (e.g. Oreopoulos et al., 2012; Kahn, 2010).⁸ Most closely connected are Huckfeldt (2022) and Altonji et al. (2016a), who argue that much of the “scarring” effect of recessions stems from workers’ initial placement in lower-skill or lower-wage occupations, which affects long-run human capital development. Other explanations for the scarring effects of recessions may be changes in skill demand (Hershbein and Kahn, 2018), job placement assistance (Crépon et al., 2013), or poor initial employer matches (Fadlon et al., 2020). Our research design permits us to identify the impact of early career experience and its attendant training, rather than the impact of macroeconomic conditions at the start of a career.

This paper also contributes to a rich literature on active labor market policy by exploring the long-run impact of occupational experience and associated training across a wide array of “blue-collar” careers. The potential for occupation and job training to improve labor market outcomes has a long history of study in labor economics; however, randomized evaluations have often yielded disappointing results (Heckman and Smith, 2000; Crépon et al., 2013). One notable exception is the small number of evaluations examining sectoral job training – a model that, much like the military, involves an initial period of occupational training paired with subsequent on-the-job experience (Katz et al., 2022). Recent work evaluating active labor market policy in Denmark finds that re-skilling can be effective for individuals who switch occupations (Humlum et al., 2023). Our results support the idea that shifting early career experiences can generate long-run economic returns.

Finally, our paper contributes to an emerging literature focused on the economic and social struggles of the non-college educated by highlighting the potential for different military careers to serve as pathways to opportunity. Recent work has documented declines in the relative standing of non-college graduates along both economic and social dimensions, particularly for young men (Autor, 2014; Case and Deaton, 2022; Reeves, 2022).⁹ Military service is the rare example of a widely available career path for young adults with little prior exposure to college (Greenberg et al., 2022). Much of the existing literature focuses on compulsory service, typically for older cohorts, and finds negative impacts on measures of long-run well being (Angrist, 1990; Imbens and van der Klaauw, 1995; Angrist et al., 2010, 2011; Bingley et al., 2020). An exception is Card and Cardoso (2012), who find beneficial effects of mandatory service among young men who would have otherwise held low-wage jobs. Research exploring voluntary military service in the US during more recent periods finds improvements in earnings for Black Americans (Angrist, 1998; Greenberg et al., 2022). Closely related to our work is Borgschulte and Martorell (2018), who link macroeconomic trends to the decision to remain in service. There is also a broader, related literature that examines specific aspects of military service, to include combat deployments, educational benefits, gender integration, military families, and military-specific

⁸See Wachter (2020) for a complete review.

⁹While it is clear that educational quality and field of study are important for long-run earnings among populations that are college bound (Altonji et al., 2016b; Kirkeboen et al., 2016; Zimmerman, 2019; Autor et al., 2020; Mountjoy, 2022; Andrews et al., 2025), there is less consensus on how to improve long-run outcomes among those who struggle in a traditional classroom environment. For example, there is mixed evidence on the returns to higher-education among academically marginal students (Zimmerman, 2014; Heckman et al., 2018; Mountjoy, 2022).

insurance programs.¹⁰ Our paper studies the relative importance of different enlisted careers for improving long-run economic trajectories.

3 Occupations in the US Army

The range of Army occupations. The United States Military is the nation’s largest single employer with roughly 2.2 million Active Duty and Reserve service members and roughly 720,000 civilian employees.¹¹ To put this in context, the size of the military labor force is comparable in magnitude to the population of Chicago, Illinois, which is the third largest US city. More broadly, there are 17.6 million living veterans in the United States comprising 7% of the non-institutionalized population over 18. These numbers are even larger among men, where one in seven US male adults have military experience. As a result there are currently 7.3 million working age male veterans which equates to roughly 1 in 10 full time male workers (U.S. Bureau of Labor Statistics, 2025). Thus veterans, and the occupational experience they accumulate while serving, play a dominant role in the labor market.

Operating an organization of this magnitude requires a diverse and varied set of roles. As a result, the US Army has over 160 distinct military occupational specialties (MOS). Each MOS is classified by a two-digit number, known as a career management field (CMF), followed by a letter. Some occupations, like infantryman (11B) and cannon crew-member (13B, also known as field artillery) are primarily connected to the Army’s war-fighting mission and have no straightforward civilian analogues. However, many other occupations do.

In fact, most Army occupations correspond to common civilian careers. For example, the Army has career management fields in human resources (CMF 42), legal services (CMF 27), finance (CMF 36), logistics (CMF 92), medical services (CMFs 60 and 61), military police (CMF 31), and network/computer specialists (CMF 25). Nearly 15% of soldiers enlist in human resources, legal, finance, or logistics roles—comparable to the 18% in infantry (CMF 11) and 6.6% in field artillery (CMF 13). In fact, the modal soldier in our data ($\approx 60\%$) is employed in a non-combat occupation.

The Army recruiting process. The Army recruiting process over our study horizon followed a standard procedure.¹² Most new recruits are young men between the ages of 17 and 21. For example, in our applicant data, 78% are male and the average age is 20.7 (see Table 1). The enlistment process typically begins when a potential recruit first sits down for a meeting with an Army recruiter. The Army recruiter’s primary role is to provide information to potential recruits about the broad nature of Army service, including things like the

¹⁰For examples of research studying combat and peace-keeping deployments, see: Bäckström and Hanes (2023); Bruhn et al. (2024); Jacobson et al. (2008); Rohlfs (2010); Negrusa et al. (2014); Anderson and Rees (2015); Cesur et al. (2016); Cesur and Sabia (2016); Cesur et al. (2022); Hjalmarsson and Lindquist (2019). For examples of research studying military educational benefits, see: Barr (2015, 2019); Barr et al. (2021). For examples of research studying gender integration in the military see Dahl et al. (2021); Greenberg et al. (2024). For examples of research studying the impact of service on families, including children, see: Angrist and Johnson IV (2000); Lyle (2006); Engel et al. (2010). For examples of research on Veteran’s Affairs Disability Compensation, see: Bruhn et al. (2024); Autor et al. (2016); Coile et al. (2021); Angrist et al. (2010); Sabia and Skimmyhorn (2023); Silver and Zhang (2022).

¹¹According to Fortune Magazine, the largest private U.S. employer is Walmart, which has 2.2 million employees globally as of 2020 (Forbes Media LLC, 2024). Department of Defense current employment information is available in publicly posted workforce reports (U.S. Department of Defense, n.d.).

¹²All information here applies to enlisted Army personnel. Officers follow a different recruitment profile (ROTC or military academies) and are not in this study. We lack data on other military branches.

military pay-scale, benefits, and potential military occupations. Once ready to enlist, recruits attend a two-day appointment at a Military Entrance and Processing Station (MEPS) to verify eligibility and take cognitive and physical tests.¹³ Once the applicant has established their qualifications and selected their occupation (more detail on this in the next section), they will sign an enlistment contract, swear an oath of service, and begin their Army career.

Qualifying for an Army occupation. Specific occupations have additional requirements beyond general enlistment eligibility. Examples include security clearances (e.g. military intelligence) or gender restrictions (e.g. combat roles). The most important occupation-specific requirement typically comes from a series of test scores known informally as “line-scores.” These scores are derived from the Armed Services Vocational Aptitude Battery (ASVAB), which applicants take during their MEPS appointment. The ASVAB itself is comprised of 10 sub-tests covering different topics such as Arithmetic Reasoning, Paragraph Comprehension, General Science, and others (see table F.2). The Army uses the results of these sub-tests to create the 10 line scores that determine occupational eligibility: Clerical (CL), Combat Operations (CO), Electronics (EL), Field Artillery (FA), General Maintenance (GM), General Technical (GT), Mechanical Maintenance (MM), Operators and Food (OF), Surveillance and Communications (SC), and Skilled Technical (ST) (see Appendix Table F.3).

Every Army occupation has a set of occupation specific line-score cutoffs that applicants must clear in order to enlist in that occupation. For example, to join as military police (MOS 31B), applicants must have a score of 91 on the ST line score.¹⁴ As a result, there is substantial scope for the line scores to shape which occupations applicants choose upon enlisting. In practice, while soldiers may enter the MEPS with an MOS in mind, applicants only learn their official line scores during the two-day MEPS appointment. The list of occupations actually available to them based on their line scores, training slots, and other qualifications will only then populate into the computerized interface. Furthermore, applicants must wait at least one month before retaking the ASVAB, new test results are binding, and guidance counselors are incentivized to get qualified individuals enlisted.¹⁵ Consistent with these institutional features, the density and balance checks we will show in section 5 provide empirical evidence that applicants do not systematically retake the ASVAB in order to qualify for a preferred occupation.

Army occupational training and subsequent service. Conditional on qualifying for service and selecting a military occupation, applicants enlist in the Army for a period of 2-6 years.¹⁶ All soldiers begin their career by spending ten weeks at a basic training unit where they learn basic soldier skills like equipment maintenance, firing a weapon, and first aid. After the ten weeks of initial basic training, soldiers will continue on to an occupation-specific school to undergo Advanced Individual Training (AIT). Infantry and Armor soldiers stay at their basic training location and attend a training program called One Station Unit Training (OSUT). AIT

¹³Basic eligibility requirements include being age 17-35, U.S. citizenship or permanent residency, background check, high school diploma or equivalent, and minimum AFQT score. The AFQT determines eligibility to enlist but does not influence eligibility for specific occupations.

¹⁴As with other enlistment requirements, the Army may issue exceptions to these occupation specific line score requirements – we will see evidence of this when we plot the first stage variation in section 5.

¹⁵Army regulations over our study period also prohibit certain types of retesting “for the sole purpose of increasing aptitude area scores to meet standards prescribed for enlistment options or programs” (Department of the Army, AR 601-210, 1995).

¹⁶Most soldiers enlist for 3-4 years (mean 3.7, median 4); 2- and 6-year terms are uncommon.

varies widely in its length and intensity, ranging from four weeks to one year, though it most typically lasts 1-6 months.¹⁷

While in AIT, soldiers learn the practical skills they will need to perform an occupation. For example, as a combat medic (MOS 68W), soldiers learn to “administer emergency medical care in the field in both combat and humanitarian situations. Your training will allow you to serve as a first responder and triage illnesses and injuries to save lives, much like a paramedic in the civilian world,” (U.S. Army, n.d.a). Similarly, as a wheeled vehicle mechanic (MOS 91B), soldiers learn to “perform maintenance, repair, and recovery operations on wheeled vehicles and select armored vehicles that serve the Army in a variety of mission-critical roles... inspect, service, maintain, repair, replace, adjust, and test wheeled vehicles, material handling equipment systems, subsystems and components, and automotive electrical systems, including wiring harnesses and starting and charging systems,” (U.S. Army, n.d.b).

Army occupational training shares many features with sectoral job training.¹⁸ This is due to the contractual service obligation, which means that soldiers accumulate subsequent, extensive on-the-job experience after they complete training.¹⁹ Early separation from the Army typically involves negative consequences, including loss of benefits. Consequently, most soldiers spend the duration of their first contract actively employed in their occupation. Figure F.1 shows, for soldiers with different initial contract lengths, the share that remain in the military over time since enlistment.

While serving their first enlistment contract, soldiers receive a standardized pay, promotion, and benefits structure. Base pay depends solely on rank and years of service (and not on occupation). Benefits such as medical care, life insurance, and retirement plans, are virtually identical for all soldiers at every rank or paygrade.²⁰ Most soldiers leave after their first contract (see figure F.1): 63-67% of those with 2-4 year terms leave within 5 years, and 69-72% with 5-6 year terms leave within 7 years. However, around 7-9% remain for 20 years, at which point they qualify for a pension. Retention rates among those who reach 10-12 years of service tend to be high.

4 Data and Summary Statistics

4.1 Military Data Sources, Line Scores, and Sample Construction

We use data from various administrative records, including U.S. Army applicant records from the Military Entrance and Processing Command (MEPCOM) spanning from February 1992 to September 2014. We com-

¹⁷One to six months of AIT covers up to the 90th percentile of AIT training requirements in our data (see Appendix Figure F.2).

¹⁸Sectoral job training provides occupational training in high-demand fields to individuals who are marginally attached to the labor force, followed by job placement services that facilitate durable employer connections. Unlike other job training models where evidence is mixed (e.g. Heckman and Smith, 2000), sectoral training has shown positive impacts in randomized evaluations (Katz et al., 2022).

¹⁹Differences in the length of time spent in an early career experience between enlistees and civilians are smaller than many might predict. For example, if we take Army applicants in the tax data who do not enlist, and we define their comparable “early career experience” as the modal occupation over the 5 years following the ASVAB (three-digit SOC code – see section 4), we find that they spend around 15% less time in that early career experience when compared to enlistees. This difference is driven, in large part, by a left tail among civilians that ‘shops’ around more.

²⁰Other components of total compensation like the Basic Allowance for Housing (BAH), which varies by location and dependents, specialized pay (e.g. for deployment), as well as enlistment and reenlistment bonuses, can vary across occupations, but cross-MOS variation in these tends to be limited, especially relative to total compensation.

bine these data with U.S. Army administrative pay and service records (1992–2024) and federal tax records (1999–2024).

Our analysis sample is composed of Active-Duty Army applicants who took their ASVAB between two sets of dates: February 1992 – December 2001 and July 2004 — September 2014. Both the way line scores were calculated and the line score cutoffs changed in January 2002 and again in July 2004 following the ASVAB renorming.²¹ Following [Greenberg et al. \(2022\)](#), we impose several additional sample restrictions: we exclude applicants with prior military service, who took the ASVAB in high school as part of the Career Exploration Program, who have invalid AFQT scores (zero or missing), and who fall outside the typical 17–35 enlistment age range. The majority of our analysis is further limited to the 99.8% of remaining applicants whose records could ever be matched by social security number to any tax record.

Our regression discontinuity analysis relies on the ASVAB line scores as running variables. We note here that one consequence of the way line scores are constructed is that, with one exception, the post-2004 line scores have numerous support points and are effectively continuous. However, prior to 2002, the line scores are derived from raw scores using a series of conversion tables, making the scores discrete values. Moreover, because some line score values group together more (or fewer) underlying test results, the densities prior to 2002 exhibit a sawtooth pattern (see e.g. figures [F.6–F.7](#)), which has implications for how we specify our regression discontinuity models (see section [5.1](#)).²² Appendix [A](#) contains more details on the line score variables.

4.2 Federal Tax Records

We merge applicants to their federal tax records using the Social Security Number from their Army application. We successfully link 99.8% of applicants. The tax records give us third party reported earnings information (Form W-2, Form 1999-Misc, and beginning 2020, Form 1099-NEC) and industry (6-digit NAICS). Crucially, the data also provide occupation information. Taxpayers report their occupation at the bottom of their Form 1040 each year, and it is available primarily for those filing electronic returns. We use a crosswalk created internally at the Treasury to merge these string variables to 2010 three-digit Standard Occupational Classification (SOC) codes, with most string variables successfully linked to a SOC code when the individual is employed. We are able to merge approximately 50% of Army applicant-year observations to a SOC between 2005–2022, with the denominator including nonfilers and those not employed. Conditional on having positive earnings, tax filers can be missing an occupation due to a variety of reasons including a failure to report anything in the relevant field on the 1040. Importantly, we note that missingness for this variable (i.e. having earned income but no SOC) appears balanced across cutoffs when we predict it using our preferred regression discontinuity model (see section [5](#) for model details).²³ Thus, when we analyze impacts on career trajectories, we keep Army applicants without a linked SOC who have no earnings in our sample (as zeros) but exclude (as missing) Army applicants who have earned income but no linked SOC. Hence occupation variables will

²¹ Prior to February 1992 the data do not reliably contain the test score version information we need to construct line scores. We drop the 2002–June 2004 period from our analysis as during this short transitional period we are unable to reliably calculate and verify the necessary scores for our identification strategy.

²² This is similar to the running variable in [Greenberg et al. \(2022\)](#), specifically the AFQT score, which also exhibits a sawtooth pattern as a result of the conversion process applied to the raw scores.

²³ Specifically, a joint F-test of the null hypothesis that missing a SOC is balanced across all cutoffs yields a p-value of 0.725.

be estimated on the sample of those who have earned income with a linked SOC or are unemployed, while earnings and industry-variables are always estimated on the full sample.

4.3 Construction of outcome variables

The primary outcomes in our analysis are measures of career trajectory and earnings; however, we also explore secondary outcomes such as educational attainment. To account for the dynamic nature of career development, all of our outcome variables are defined in terms of years-since-ASVAB. We now discuss their construction in more detail.

Career trajectory. A natural way to measure impacts on career trajectory is via the likelihood of subsequently persisting into an identical or closely related occupation. Thus, our preferred outcome for measuring changes in career trajectory is defined as the share of years (e.g. 11-20 years after the ASVAB) for which individual i is observed in an occupation (three-digit SOC code) that is closely related to the corresponding Army occupations unlocked at a line score cutoff. Constructing this variable requires linking Army occupations to their closely related civilian counterparts. To accomplish this, we create a Military Occupational Specialty (MOS) to three-digit SOC crosswalk based on several internal Army crosswalks and via one-by-one manual comparisons of the descriptions of MOS duties to those reported for each SOC code by the BLS. Given the fact that several civilian occupations can reasonably be linked to an MOS, we crosswalk each MOS to a “primary” SOC code and a “secondary” SOC code. To capture notions of career progression, we also maintained any relevant supervisory SOC codes.²⁴ For various combat occupations like infantry or field artillery, for which there is no natural civilian counterpart, we link the primary SOC code to the broad military category (55-0) and, for their secondary SOC code, we link them to law enforcement (33-3). Appendix table F.14 contains the complete crosswalk.

Since each line score cutoff can determine eligibility for multiple military occupations, we construct cutoff-specific, occupation-group outcomes. The natural way to do this is to create the outcome based on whether the individual is in any of the primary, secondary, or supervisory occupations corresponding to the full set of military occupations “unlocked” by that cutoff.²⁵ In practice and in the interest of power, since several military occupations at each cutoff are negligible in size or not substantially impacted by the cutoff, we do this for the military occupations where the cutoff creates a sufficiently large impact on take-up.²⁶ Thus, if an

²⁴For example, the “49” category of SOC codes contains installation, maintenance, and repair occupations. Within that broad category, there are numerous, more specific occupations such as 49-2 (“Electrical and Electronic Equipment Mechanics, Installers, and Repairers”) and 49-3 (“Vehicle and Mobile Equipment Mechanics, Installers, and Repairers”). One of those more specific three-digit SOC occupations is 49-1 which denotes “Supervisors of installation, maintenance, and repair workers.” Thus, if we link an Army occupation to any three-digit SOC code in the 49 broad category, we additionally retain the supervisory category in our crosswalk in order to capture natural career progression. In all, this results in each MOS having a mapping to a primary and secondary SOC code as well as any associated supervisory SOC codes.

²⁵Our results are not sensitive to this choice. In appendix figure F.8, we show that our key results on career trajectory continue to hold if we instead define this outcome variable using only the primary plus supervisory SOC codes. In appendix table F.8, we also show robustness to this definition by using a crosswalk that is built by a large language model trained to link military occupations to civilian occupations using only the skills and tasks contained in official Army documents – see discussion of robustness in section 6.1.

²⁶Specifically, we base our outcomes on the MOS at each cutoff that either has an MOS-specific, first stage coefficient greater than or equal to 0.003 or whose MOS-specific, first stage coefficient is $\geq 10\%$ of the total first stage across all MOS at the cutoff. In appendix table F.8, we show robustness to several other sensible ways of defining this variable, including a continuous measure built by feeding MOS / SOC descriptions from official documentation into a large language model. In this table, we also show results

applicant's 3-digit SOC falls into any of the codes associated with these military occupations at a given cutoff, then the variable is coded as 1 (denoting career continuation), and otherwise it is 0.

Income. Earned income is compiled from third party reported information returns available between 1999 and 2024. The measure is primarily derived from the wage income subject to Medicare tax reported by employers on Form W-2. We also add in contractor income reported by contractor businesses on Form 1099-MISC (and beginning in 2020, Form 1099-NEC). Our earned income measure is in 2010 real dollars, winsorized at the 99% level. To account for the fact that those in service also receive compensation that would typically be included as taxable wages in the civilian sector but are exempt from taxation for servicemembers, we follow [Greenberg et al. \(2022\)](#) and use the Army personnel records to add these additional forms of compensation into our measure of income for individuals who are on active duty.²⁷

Additional outcomes. We supplement our primary data sources with a number of other data sets that allows us to measure concepts like post-secondary educational attainment; veterans affairs disability receipt; occupation-specific measures of task and skill intensity; and occupation-specific measures of the prevalence of credentials or certification requirements. See appendix [B.2](#) for discussion of the construction of these supplementary outcomes, with appendix table [F.4](#) giving a concise summary.

4.4 Summary Statistics

Summary statistics for U.S. Army applicants are presented in Table [1](#). The data shows that applicants are typically young (≈ 21 years old), predominantly male (78%), and not currently attending college (93% have not attended even a semester of college). In fact, among the three quarters of applicants who are not enrolled in high school at the time they take the ASVAB, 14% do not even have a high school diploma. Compared to a nationally representative sample of 17–23 year-olds from the 2000 Current Population Survey (CPS), the Army applicants in this study are more likely to be Black (22% versus 15%) and less likely to be Hispanic (11% versus 15%). The applicants are also broadly representative in terms of ability: the average applicant has scores that fall at the 51st percentile of the AFQT score distribution, which is a cognitive aptitude test that is nationally normed. In the long run, the average applicant in our data will make an average of \$32,342 annually, over the 11-20 years after the ASVAB date. Forty-one percent of our applicant sample ultimately enlists.

We next present summary statistics on career trajectories. Specifically, we explore how long-run career paths vary by initial military occupation. These summary statistics provide context for understanding the different types of Army occupations and common three-digit SOC codes in our data, as well as to help validate the crosswalk that aligns the military occupations with comparable civilian careers.

Descriptively, we find that many of the soldiers serving in specific military occupations go on to have analog-

from a measure based on SOC-specific weights derived from the first stage impact that each cutoff has on entering the Army in a relevant military occupation. In all cases, the results are robust.

²⁷Military pay that is not reported on the W-2 as subject to Medicare tax encompasses various allowances and payments, such as the Basic Allowance for Housing (BAH), Basic Allowance for Subsistence (BAS), and additional compensation for deployment or foreign assignments, including Hardship Duty Pay, Imminent Danger Pay, Hazardous Duty Pay, and Family Separation Allowances. We do not account for other benefits, such as health coverage, retirement contributions, or G.I. Bill tuition and related housing allowances, some of which may have more natural civilian counterparts. Military pay not subject to taxation accounts for 17–25% of servicemembers' *total* compensation (see Appendix B of [Greenberg et al., 2022](#)).

gous careers as civilians in the long-run. We show this in two ways. First, in Appendix Figure F.3, we present Sankey plots that highlight long-run transition probabilities at the more aggregated CMF level. Across almost all CMFs, we observe that soldiers are most likely to end up in related civilian occupations. For example, we find that soldiers working in signal support occupations (computers/electronics) are most likely to eventually work in computer occupations (about 11%); those in medical fields end up in civilian healthcare occupations (about 15%); and soldiers working in careers as military police officers most often go on to roles in law enforcement (about 14%). Interestingly, soldiers in combat arms occupations are most likely to end up in law enforcement and construction (about 9%). Next, we disaggregate the CMFs into transition probabilities for specific occupations (i.e. MOS) by plotting them as a heatmap. Appendix figure F.4 paints a similar picture: soldiers involved in military information technology (all MOS beginning with CMF 25 such as 25P, 25U, 25B, etc.) end up in computer occupations (SOC 15-1); health-care related occupations (CMF 68) go on to careers in health-care (SOC 29-2), Human Resource soldiers (MOS 42A) end up in Human Resources (SOC 13-1), many of the mechanics (CMF 91) end up in maintenance, mechanic, and repair occupations (SOC 49-3 and 49-9), and so forth. The outlined cells on the figure correspond to the primary and secondary SOCs in our MOS-SOC crosswalk. Hence, this figure also reassuringly shows that our MOS-SOC crosswalk frequently maps MOS to the SOC codes that soldiers are likely to end up in.

Thus, the descriptive evidence reveals a strong correlation between initial military occupations and the corresponding long-run civilian career. However, this descriptive evidence is surely driven, in part, by selection: those with an interest in computers are presumably more likely to both choose that occupation upon enlisting in the Army, and to continue to choose similar occupations in the long-run after they have exited the service. Thus, to disentangle selection from causal effects, we now turn to the regression discontinuity analysis.

5 Empirical Strategy

Our empirical strategy for identifying the causal effect of different military occupations is a fuzzy regression discontinuity design. As discussed in section 3, the Army determines eligibility for different occupations using thresholds in a series of 10 test scores known as line scores. Intuitively, our identification strategy will compare the outcomes of soldiers who barely cleared one of these cutoffs to soldiers who fell just short. If this variation is as good as random, then we can sharply identify the reduced form causal effect that eligibility for certain army occupations has on long run outcomes. If the only way that eligibility affects the outcome is by changing the likelihood an applicant enlists in an occupation “unlocked” by the cutoff, then we can use this variation as an instrument for enlisting in that occupation in a fuzzy RD framework.

5.1 Research design

Formally, we estimate reduced form effects specific to each cutoff using the following sharp regression discontinuity model:

$$Y_i = \gamma Z_i + f(X_i) + \epsilon_i \quad (1)$$

Where Y_i is the long run outcome of applicant i measured at some specific point in time (e.g. time spent in an occupation 11-20 years after the ASVAB date);²⁸ X_i is applicant i 's line score from the relevant subtest; $Z_i = \mathbb{1}(X_i > c)$ is a binary variable indicating whether applicant i 's line score cleared the cutoff (c);²⁹ and ϵ_i is a residual. The parameter of interest from equation 1 is γ . It represents the expected difference in mean outcomes for applicants with a line score precisely at the RD threshold.

The key assumption we rely on for causal identification is that the conditional expectation function mapping the running variable into the outcome would have otherwise been continuous provided clearing the threshold did not change Army occupational eligibility. If this assumption holds, then γ represents the reduced form causal effect of occupational eligibility on the outcome of interest (Lee and Lemieux, 2010). This assumption has a number of testable implications which we explore in section 5.2.

Our baseline model specifies $f(\cdot)$ as a piece-wise quadratic polynomial within a bandwidth of 10 test score points around the threshold. We fix the bandwidth to be common across cutoffs because, as discussed in section 4, prior to 2002 the running variable is discrete and has a saw-tooth pattern and hence it is not possible to select an optimal bandwidth for cutoffs over that time period.³⁰ In several cases the 10 point bandwidth overlaps another cutoff in our data. In order to avoid mis-specification, in these cases we constrain the bandwidth to be smaller than 10 such that it never crosses another cutoff. When it is not possible to have a bandwidth larger than 5, we switch to a linear polynomial to improve power. However, our key findings are robust to varying the bandwidth and the order of the polynomial, including using all quadratic or all linear polynomials and a tighter bandwidth (see Table 3). Our preferred set of results use all cutoffs that have first stage effects exceeding 0.01 for our primary 11-20 earnings outcome. Thirty-five cutoffs meet this requirement and, with the exception of robustness, we focus on these cutoffs throughout.³¹ Standard errors are robust to heteroskedasticity.

Moving beyond the reduced form effect of occupation eligibility, we also explore the impact of joining an Army occupation unlocked by each cutoff using the following fuzzy regression discontinuity model:

$$Y_i = \beta D_i + g(X_i) + \eta_i \quad (2)$$

$$D_i = \pi Z_i + h(X_i) + u_i \quad (3)$$

²⁸Because tax data is only available 1999-2024, when we examine outcomes at different time horizons (e.g. earnings 11-20 years out) we weight each individual by the number of years she could potentially be observed in tax records.

²⁹In practice, there are multiple line scores and sometimes also multiple cutoffs per line-score; however, because our preferred empirical approach involves estimating one regression discontinuity model per cutoff, we suppress the implied indices on the line score and cutoff variables, as well as the associated model parameters, for notational simplicity.

³⁰In the post-2004 cases where we use a bandwidth of 10, the optimal bandwidth selected using the method of Calonico et al. (2014) tends to be close to (and indeed slightly higher than) 10 at the vast majority of cutoffs, with a median of around 13 and a mode of 11.

³¹Cutoffs with very small first stages typically apply to rare occupations. We restrict to cutoffs with larger first stages to facilitate computational tractability – in many cases, constructing proper standard errors for the 35 cutoff sample requires stacked regressions involving over 10 million observations (see table 1 for summary stats on the stacked sample) – and for ease of visualization, but we note here that our results are not sensitive to this decision. As discussed in more detail when we cover robustness in section 6.1, Appendix Table F.7 contains results using an expanded set of cutoffs constructed by requiring the first stage F-stat to be in excess of 10. The findings are similar.

Where D_i is an indicator that takes a value of 1 if the applicant enlists in the Army in an occupation they became eligible for as a result of clearing cutoff c ; the functions h and g are specified identically to f from the reduced form model; u_i and η_i are residuals; and the remaining variables are as described previously. The parameter of interest from this model is β , which we estimate using two-stage least-squares.

Causal identification in equation 2 requires that two more key assumptions hold in addition to the assumptions necessary for causal identification of the reduced form. The first additional assumption is that there is a non-zero first-stage. In other words, occupation eligibility does, in fact, change the likelihood an applicant enlists in the Army in the relevant occupation. We provide evidence that this assumption holds in section 5.2. The second additional assumption is an exclusion restriction. This requires that the only reason occupational eligibility affects the outcome is by changing the likelihood of enlisting in that occupation. We believe this assumption is reasonable in our setting given the institutions under consideration. The line scores themselves affect no other aspect of military service and are not used in the civilian world in any capacity. Nevertheless, we will also present results from over-identification tests that are consistent with exclusion.

5.2 Evidence in support of identifying assumptions

First Stages. We first present evidence that the first stage effect of crossing an eligibility threshold on occupation take-up is strong. Panel A of figure 2 plots the first stage discontinuity pooled across cutoffs.³² Thus, we can see visually how, on average, crossing a line-score threshold creates a discontinuity in the likelihood of entering a military occupation unlocked by that cutoff. Appendix figure F.5 further provides cutoff specific first stage discontinuity plots, along with the corresponding point estimates and standard errors. Across all cutoffs, there are small (1-6 p.p.) but visually apparent and statistically precise discontinuities in the likelihood of entering a military occupation unlocked by that specific cutoff.³³ In fact, all of the 35 cutoffs we leverage for identification have first stage F-statistics above 150, which is well in excess of commonly used thresholds for diagnosing weak instruments (Andrews et al., 2019; Lee et al., 2022; Angrist and Kolesár, 2024). Thus, the data reveal that the line score thresholds substantially impact the likelihood of entering specific military occupations.

No Manipulation. Having established a set of first stages, we next present evidence consistent with no manipulation. If individuals manipulate the running variable in order to obtain the treatment, this can invalidate the continuity assumption that underlies causal inference in a regression discontinuity design (McCrary, 2008). To examine this potential confound, we begin by presenting standard density or “bunching” checks. The

³²In practice, we limit this pooled figure to cutoffs where we can observe five support points in the distribution of the running variable before and after the cutoff. This is because, in a small number of cases (3 cutoffs), there is either not enough support on one side of the running variable to extend to five points, or we are forced to truncate the running variable in order to avoid “blowing through” another nearby cutoff. Including these complicated cutoffs in the pooled figure could create a misleading visual impression of “jumps” due to the changing nature of the sample used to create the binned average at each point. This restriction avoids that problem. Evidence of strong first stages for these additional cutoffs can be found in Appendix figure F.5.

³³It is also interesting to probe why the magnitudes of the first stages vary across cutoffs. For example, we might expect occupations that produce more attractive outcomes to be more appealing to applicants and hence crossing line-score cutoffs associated with these occupations to produce larger compliance rates. Appendix Figure F.18 reveals statistically insignificant relationships between the first stage size and various occupation features (though slopes are often modestly positive where expected to be). The main exception is MOS size, with larger military occupations having significantly larger first stages. This is consistent with the fact that applicants can only select an occupation if the Army has a training seat available for that occupation, which is more likely for large occupations.

histograms of all of the running variables in both periods are presented in Figures F.6–F.7. As discussed in section 4, over the 1992-2001 period the line scores are discrete, with some line score values grouping together more (or fewer) underlying test results. This leads to the “sawtooth” appearance of the density charts. Despite this, it is visually clear that there is no obvious score manipulation at the cutoffs in both the 1992-2001 and 2004-2014 periods. While the discrete scores for pre-2002 years also make formal density checks unavailable, we can collapse the densities to the integer level and run our baseline specification on the “counts.” When doing so, we fail to reject the null of no manipulation at all pre-period cutoffs (the smallest p-value is 0.12). For the 2004-2014 period, where the line scores are continuous, the figures also report the bias-adjusted p-values from [Cattaneo et al. \(2018\)](#)’s density test using their optimal bandwidth and a local linear, triangular kernel specification. In all cases, we fail to reject continuity of the density at conventional significance levels.

Balance. We also check for balance across the available set of time-of-application covariates in the Army applicant data, as well as for various tax data covariates measured in the year prior to application, exactly as in [Greenberg et al. \(2022\)](#). These covariates are available for applicants who enlisted as well as those who did not. The applicant data covariates include age, gender, race (white, Black, and Hispanic), and education (in high school, less than high school diploma, high school diploma, and some higher education). In the tax data we examine tax filing, earnings, employment, post-secondary attendance, and marriage in the year prior to Army application. For each of these variables, we employ our baseline specification (see Section 5). Since there are 35 cutoffs checked across 14 covariates totaling 490 regression coefficients, the resulting table is excessively large and is included in tables F.5 and F.6 in the Appendix. As a visual check, we plot the distribution of p-values across all 490 estimates in figure 1. Consistent with a balanced treatment, the distribution of p-values appears uniform. This interpretation is further supported by a joint test of balance across all covariates and cutoffs, which fails to reject ($P = 0.173$). Thus, the data are consistent with the RD variation being balanced across observable covariates.

Exclusion. To probe the exclusion restriction, we present results from an over-identification test that jointly examines instrument validity (i.e. exclusion) and LATE homogeneity (i.e. linear functional form) across cutoffs ([Wooldridge, 2010](#); [Angrist and Hull, 2023](#)). Intuitively, the test asks whether the relationship between the reduced form estimates at each cutoff and their associated first stages is well characterized by a line (consistent with LATE homogeneity) through the origin (consistent with exclusion). Our model passes this test ($P = 0.306$), which is consistent with exclusion being satisfied. See appendix table F.7 for the results of this test, and also the discussion of robustness in section 6.1 for more precise details on how the test is implemented.

6 Long Run Effects of Early Career Occupational Experience

This section presents our key finding: that early life occupational experiences have important long-run effects on career outcomes and economic well-being. We begin by showing that early career occupational experiences cause important changes in long-run career choice. In particular, we find a substantial degree of path dependence. On average, entering a specific military career increases an applicant’s likelihood of being on a similar career path 11-20 years later by 18.6 p.p., with limited evidence of heterogeneity across cutoffs.

Next, we turn to wages, where we find a substantial amount of treatment effect heterogeneity. The data reject both that the causal effects from each individual RD are equal to each other, and that they are all identically zero. However, we find that the heterogeneous treatment effects are well predicted by implied differences in occupational wage premia: early career experiences that, in the long-run, move applicants out of low-wage occupations and into high-wage occupations create economic success. Long-run changes in industries and industry-level wages, occupational non-routine task intensity and educational attainment also predict causal improvements in earnings. However, we find limited evidence that causal effects on features specific to military careers (e.g. likelihood of enlisting, combat exposure) are predictive of impacts on earnings.

We conclude by asking which types of early career experiences produce the largest gains. We find that military occupations connected *ex-ante* to high-wage career paths also have large impacts on earnings. Under strong assumptions about the nature of treatment effect heterogeneity, we are also able to identify specific groups of military careers that are most responsible for long-run success. We find that occupations related to computers, electrical work and electronic repair; construction, law-enforcement, and medical fields produce the largest improvements in economic well-being.

6.1 Path dependence in the labor market

Career trajectory. We begin by quantifying the effect of early career occupational experiences on long-run career choice. While early career occupational experience may influence later life outcomes in a number of ways, a natural starting point for understanding how initial occupation shapes long run outcomes is the extent to which they increase the likelihood of pursuing a related career path over the long-run.

To quantify this, we estimate our reduced form model (equation 1) for each cutoff separately. The outcome variable for each of these separate regressions is the share of time 11-20 years after the date they took the ASVAB that individual i is observed in an occupation linked to an Army occupation unlocked by the cutoff. For example, obtaining a score of 85 on line score OF is necessary for applicants to be eligible for Army occupations 92G (Culinary Specialist) and 88M (Motor Transport Operator). These occupations are naturally linked to the civilian occupations of Cooks and Food Preparation Workers (SOC 35-2), Food Service Managers (a subset of SOC 11-9), Motor Vehicle Operators (SOC 53-3) and Material Moving Workers (SOC 53-7). In this case, the outcome variable would be equal to the share of years an individual is observed in those occupations (SOC codes 35-2, 11-9, 53-3 and 53-7), or their supervisory counterparts (here, 35-1 and 53-1), or, if still in the Army, in an MOS linked to any of these SOC codes, 11-20 years after the date that they took the ASVAB test. See Appendix B for more detail.

For policy purposes, it is important to distinguish between two possible mechanisms that could lead to path dependence in our setting. First, it could be that path dependence arises as the result of a durable connection with an employer. For example, some of the applicants who are quasi-randomly assigned to the mechanic occupation could find that they enjoy working as a mechanic in the Army during their initial term of service. Thus, the cutoff may cause them to stay in that occupation with the same employer for an extended period of time. These “retention effects” are real causal consequences of our treatment and, as such, our baseline estimates will incorporate them. However, a second channel that could lead to path dependence would be the accumulation of occupation specific skills / human capital, credentials, or information. Since these factors are

not specific to the employer, we might expect them to extend across firms in the labor force in a more general way. To explore the second channel, we also estimate treatment effects on the likelihood an applicant is in an occupation unlocked by the cutoff at a *non-Army employer*.³⁴ We expect this to be a lower bound on the skills / information / credentialing channels, since these channels undoubtedly increase the likelihood of staying in the occupation within the same employer as well – it is just not possible for us to completely separate them out using our experimental variation alone.³⁵

Panels (b) and (c) of Figure 2 provide pooled, reduced form evidence that the cutoffs do, indeed, shape long run-career paths. In panel (b), the x-axis is the line score running variable, centered to be zero at the cutoff. The Y-axis denotes the probability, 11-20 years after the ASVAB, of being in an occupation closely related to one unlocked at the cutoff. Each dot is a conditional mean, pooled across cutoffs, at a point in the support of the line score distribution. Panel (c) is identical to panel (b), except it redefines the Y-variable in terms of occupations at non-Army employers only. In both panels, we observe clear visual evidence of a discontinuity in the likelihood of continuing on a closely related career path. However, as is the case with any instrumental variable design, the magnitude of the reduced form is hard to interpret without accounting for the associated first stage.

To better understand the magnitude, panels (a) and (b) of Figure 3 display the reduced form and first-stage estimates together in a “visual-IV” (V-IV) plot (Angrist and Pischke, 2009; Walters, 2015; Angrist and Hull, 2023). Each point in this figure corresponds to a distinct line-score cutoff. The y-axis corresponds to the reduced form estimate from our regression discontinuity model (equation 1). The x-axis corresponds to the first-stage estimates from model 2. Panel (a) shows our baseline results that include within employer retention. Panel (b) restricts to non-Army employers and hence excludes within-employer retention.

The V-IV slope from panels (a) and (b) of Figure 3 is our preferred way to summarize the long-run career impact of a typical Army occupation. Up to a weighting, the V-IV slope in panels (a) and (b) of figure 3 is equivalent to stacking the micro-data used for the 35 cutoff specific models implied by equation 2, and then estimating a 2SLS regression on the stacked data (see Appendix D for more discussion). Thus, our estimate is naturally interpreted as the causal effect that joining a typical army occupation has on the likelihood of continuing along a closely related career path later in life.³⁶ We also find little evidence that treatment effects

³⁴In appendix figure F.11 we also estimate treatment effects on the likelihood an applicant is in an occupation unlocked by the cutoff at any *non-military employer*, thereby also zeroing out cases where people transferred to other services or whose primary job was in the reserves. Results are very similar.

³⁵In practice, four of the line score cutoffs apply to primarily combat Arms occupations like Infantry or Field Artillery which have no direct civilian counterpart and are crosswalked to a military SOC as their primary SOC (and law enforcement as their secondary SOC). As a result, civilian path dependence is not as well defined at these cutoffs (especially when using primary SOC only). Hence, we opt to drop these 4 cutoffs when examining non-Army path dependence. However, this decision is not consequential. We find very similar results if we retain them with path dependence being based on their secondary SOCs (see appendix figure F.9). This is because crossing thresholds like the combat CO87 cutoff makes applicants substantially more likely to be in law enforcement occupations.

³⁶We prefer the V-IV slope to the slope of the corresponding 2SLS regression on the underlying stacked micro data because of its simple, transparent interpretation and because of its computational tractability. We estimate the V-IV slope using OLS with the reduced form estimates as the dependent variable and the first stage estimates as the independent variable, and we construct standard errors using the delta method and the full variance-covariance matrix of the estimated parameters, which is itself calculated on the stacked micro-data using a person-id level cluster and is therefore robust to typical forms of error dependence. In practice, we obtain similar results from 2SLS estimated directly the micro-data (see appendix table F.7).

on career path are heterogeneous across cutoffs. A formal test of the joint hypothesis that the IV estimates are all equal to one another yields a p-value of 0.341 (see Table 2). This suggests that path dependence is a relatively constant feature of early career experiences and not particular to certain types of occupations.

Another natural way to summarize the impact of these early career experiences on career trajectories is by directly examining the individual IV estimates. Panels (c) and (d) of figure 3 are identical to (a) and (b), except that we replace the reduced form estimates on the Y-axis with the corresponding IV estimate. This allows us to see visually how the treatment effects vary with average compliance rates (i.e. the size of the first stage). It also allows us to visualize “average” persistence across occupations by reporting the mean of the individual fuzzy RD estimates as a horizontal line.

By either measure, we find that early career occupational experiences create path dependence. The V-IV slope on the estimates that include retention within the Army suggest that the typical early life occupational experience causes an 18.6p.p. increase in the likelihood of being found in a closely related occupation 11-20 years later. Retention within-employer accounts for a little over one-third of this effect. For our estimates that exclude the within employer retention channel, we find an 11.8p.p. increase in the likelihood of being found in a closely related occupation at a non-Army employer over the 11-20 year time horizon.³⁷ Turning to panels (c) and (d), we find that the mean-IV suggests that the causal impact on career path is 18.3p.p. with employer retention, and 10.1p.p. without employer retention. Panel (a) of table 2 summarizes the key point estimates and provides corresponding standard errors.³⁸

Dynamics. Figure 4 plots our preferred summary measure of the impact on career path over time. The Y-axis corresponds to the V-IV slope coefficient. The X-axis corresponds to different time horizon’s relative to date of the ASVAB. Circles correspond to estimates of the causal effect on pursuing a related career. Triangles correspond to estimates that exclude within-employer retention and therefore represent the causal effect of pursuing a related career as a civilian. Whiskers denote 95% confidence intervals.

The dynamic pattern reveals large increases in Army occupational experiences in the short run that give-way to causal effects on the relevant non-Army occupations over time. In years 1-5, we see a 62 p.p. increase in the share of years spent on the relevant career path; however, virtually all of this is due to takeup of the relevant military occupation. By years 6-10, after every soldier in our sample has had the opportunity to leave the Army, we see a large decline in the estimates that include employer retention effects coupled with an 11 p.p. increase in the causal effect on pursuing that career at a non-Army employer. And by years 16-20, at which point nearly all soldiers in our sample have exited the military, we find that the two series have

³⁷One way to benchmark the impact on civilian career path is to take the applicants who did not enlist and define their comparable early career experience as their modal occupation over the 5 years immediately following the ASVAB, and then calculate the share of individuals still in that occupation in each year after ASVAB. We find that about 20% of these individuals are still in their same three-digit SOC code 11-20 years later. This suggests that our point estimates for civilian path dependence (11.8pp for 11-20) are large compared to a relevant civilian benchmark.

³⁸We also report results on our preferred V-IV measure of occupational persistence for Black and white subsamples in table F.10. Results, as expected, are substantially noisier, but overall, 11-20 occupational persistence is generally not statistically distinguishable across race and close to the full sample findings. However, consistent with Black soldiers staying longer in the military, occupational persistence results for Black applicants are largely driven by persistence in the military as opposed to civilian persistence (Greenberg et al., 2022), while much of the observed effect for white applicants occurs within non-Army employers (at least based on point estimates).

nearly converged, showing that impacts which exclude within employer retention dominate over longer time horizons.

Push versus pull effects. We find evidence that early career experiences in the military also “pull” applicants out of very different long-run counterfactual careers. Unlike the “push” effect highlighted in our results on path dependence, where it is natural to look for long-run effects on the likelihood individuals remain in a career unlocked by the cutoff, there are no institutional details to guide us towards the types of military careers individuals at each cutoff are most likely to be diverted away from. Instead, we will rely on the “pseudo first-stage” variation created by the causal effect each cutoff has on the likelihood of entering *every* military occupation. Intuitively, if the data reveal that the cutoff tied to careers in information technology causes a decline in the likelihood an applicant enters a military career as a police officer, we also expect to find a corresponding reduction in the likelihood that individual goes on to be a civilian police officer 20 years later.

Appendix figure F.12 displays this visually. Each point is an MOS-cutoff pair (e.g. MOS i and cutoff j). The x-axis denotes the pseudo-first stage estimate of how cutoff j affects the probability of entering the Army in MOS i . Then the Y-axis denotes the causal effect cutoff j has on the probability of being in the corresponding non-Army occupation for MOS i in the long-run. In the figure, green diamonds represent cutoff-mos pairs in which the MOS is the “focal” MOS unlocked by cutoff j . We would expect all green diamonds to be in the top-right quadrant, since (as we’ve shown) crossing a cutoff increases the probability of entering the Army in the focal MOS *and* increases the probability of being in the corresponding civilian occupation 11-20 years later. Points in the bottom-left quadrant indicate a cutoff that reduces your likelihood of being in an unrelated MOS (black points), and which also appears to reduce the probability of being in the corresponding civilian occupation 11-20 years later. The solid black line is the line of best fit. The dashed blue line is the best fit line among points with a positive pseudo-first stage. The dotted red line is the best fit line among points with a negative pseudo-first stage.

The figure suggests that, over the long-run, applicants are less likely to be found in the civilian careers connected to the military occupations they were diverted away from. To see this, first note that because most of the “focal” military occupations unlocked at each cutoff (green points) have positive and relatively large first stages, the slope of the line fit to MOS-cutoff pairs with positive first stages ranges from 0.104 to 0.123 depending on the specification (see appendix table F.9), which effectively replicates our baseline V-IV results for non-Army careers (see table 2). Next, observe that the regression line fit to MOS-cutoff pairs with *negative* pseudo-first stages (indicative of short-run diversion) has a slope that ranges from 0.077 to 0.105, suggesting that applicants are less likely to be in the civilian careers they are diverted away from. Importantly, these “pull” effects will have implications for our analysis of wages in the next section. However, before turning to wages, we finish this section with a discussion of robustness.

Robustness. The findings in this section are robust to a variety of alternative specifications. To begin, we note that the key findings on path dependence are robust to including controls and to varying the bandwidth and polynomial order of the underlying RD estimates. Table 3 shows how our key results change when using all quadratic polynomials, using a tighter, maximum bandwidth of 5 everywhere, using only linear polynomials

with a maximum bandwidth of 5, and to controlling for the time of application covariates in Table F.5. In all cases, we find substantively similar results.

Next, we note that our MOS to SOC code crosswalk is limited in two ways. First, it is narrowly defined, linking each MOS to, at most, two civilian careers (the “primary” and “secondary” SOC described in section 4). However, it could be that the skills and human capital developed during an occupational experience apply to a broader range of occupations in varying degrees. The second limitation is that our definition of being in a related career weights each SOC code equally when, in fact, cutoffs can impact multiple MOS in varying degrees, each linked to a different SOC code – in that case, if there are heterogeneous first stages at the MOS level, then our baseline estimates would mechanically underestimate the long run causal effect of landing in a related career since we should expect the cutoff level impacts to be larger for SOC codes linked to those MOS most affected by the cutoff. Thus, to assess the robustness of our baseline approach to more flexible ways of linking military and civilian careers, we develop two alternative measures of being in a related career.

For our first alternative measure, we trained a large language model to produce occupational distances in the form of long-run transition probabilities. Importantly, we formed these linkages using *only* the nature of the tasks and skills required for each military occupation as described in official Army documents (see appendix C for more details). We then re-estimated our baseline model re-defining the outcome to be share of years in a SOC code linked by the large language model to an MOS unlocked at the cutoff. Appendix figure F.10 has the results, which are almost identical to our baseline model. Thus, our results are robust to broader ways of linking military occupations to SOC codes.

For our second alternative measure, at each cutoff we built SOC-specific weights using the MOS-specific first stage estimates at each cutoff. More specifically, the SOC specific weights reflect the share of the full first stage attributed to the MOS’s linked to that SOC code at each cutoff (see Appendix C for a precise mathematical definition). We then construct our outcome variable as the weighted share of time, 11-20 years after the ASVAB date, that an applicant was working in an occupation tied to an MOS unlocked at that cutoff. Table F.8 contains the results, with the different outcome variables standardized for comparability across these different measures. As expected, we find larger estimates when using the weighted first stages – in fact, the point estimates suggest that our baseline could actually be understating the causal effect a typical military occupation has on the likelihood of entering a closely related career by 38%. However, we caution against interpreting this finding too strongly, since the confidence intervals suggest the two estimates are not statistically different.

Next, we note the fact that the same underlying subtests are sometimes used to construct different line-scores, which could create problems with the exclusion restriction. This would happen if, for example, getting one additional question correct at the information technology cutoff also increased the likelihood that the applicant cleared the cutoff for mechanic, thereby increasing the likelihood that applicant went on to enter the mechanic occupation in the military. Appendix figure F.12 (described in the preceding sub-section) suggests that this is not happening in practice. There are a total of 4,544 non-focal cutoff specific “pseudo first-stage” estimates. Of these non-focal estimates, only 5.4% (244) formally reject the hypothesis that the estimate is weakly less than zero at the 5% significance level. This means that, in aggregate, crossing a cutoff appears to move new

recruits into an unrelated MOS no more frequently than the 5% we would expect by chance alone. This is consistent with the cutoffs only having “push” effects on the expected, focal military occupations.

Our decision to limit the analysis to cutoffs with a first stage in excess of 0.01 could, in theory, create issues with our statistical inference. To address this, we also show results using an expanded set of cutoffs determined via a more traditional screen for weak instruments that requires the F-stat of the first stage to exceed 10. Because many of these cutoffs have substantially smaller, less powerful first stages than those used in our baseline model, we use the analogous two-stage least squares micro regression to replicate our main findings, rather than directly estimating the slope of the implied V-IV line.³⁹ This approach will reduce the influence of less powerful cutoffs in the final average, since two-stage least squares on the micro data will implicitly weight by the strength of the first stage (Angrist and Imbens, 1995; Angrist et al., 1996). Table F.7 contains the results. In all cases, the findings are similar to our baseline model.

Finally, we note here that our preferred specification passes a natural over-identification test implemented on the micro-data in the spirit of Kline and Walters (2016). Intuitively, if the V-IV slope is a sensible way to summarize the causal variation in Figure 3, then the estimates should be well characterized by a line (consistent with homogeneous effects / linear functional form) through the origin (consistent with exclusion). As shown in Angrist and Hull (2023), this is effectively the assumption that is tested in a classic over-ID test whenever there are more instruments than endogenous regressors – as is the case when estimating the slope of the V-IV line directly on the micro-data (see appendix D). Thus, we implement the over-ID test and display the results in table F.7. In all cases we fail to reject, which is consistent both with exclusion and with the idea that we are not ignoring important sources of heterogeneity (e.g. from different LATEs or driven by different counterfactuals) by summarizing this variation with the slope of the V-IV line.

6.2 Early career occupational experience and long-run economic well-being

We now quantify the impact that these early career experiences have on long run earnings. To begin, we note that the causal effects on earnings are highly-heterogeneous across cutoffs. Panel (b) of Table 2 presents the results of two joint tests. The first tests the null hypothesis that the IV-estimates are identically zero across cutoffs for three different labor market outcomes: wages, working, and log wages. In all three cases, we reject: these early career experiences are generating consequential changes in earnings. The second joint tests asks whether the causal effects across cutoffs are all identical to one another. Once again, we reject. This suggests that the magnitude of the causal effects vary substantially across the different early career experiences shifted by different cutoffs.

One potentially important source of heterogeneity in this setting is the varying nature of the long-run counterfactual career across cutoffs. In the preceding section, we established that crossing a line-score threshold shifts applicants into- and out-of occupations in the long-run that are closely related to the Army occupations they are shifted into- and out-of in the short-run. Thus, if we find that the information technology cutoff improves earnings, but the mechanic cutoff does not, that could be due to the fact that careers in information technology tend to pay well, *or* it could be due to the fact that the counterfactual career that the applicants to mechanic were “pulled” out-of simply had a much higher average wage.

³⁹See Appendix D for more discussion.

For this reason, our preferred approach to studying the heterogeneous impact on earnings will involve benchmarking the causal effect on earnings to the corresponding changes in occupation-specific average wages. Causal impacts on occupational average wages directly measure the implied earnings differential between the careers that a cutoff tended to shift applicants out-of and in-to. Thus, it is a straightforward way to characterize the nature of the career shifts documented in the preceding section that does not rely on our cross-walk, and hence flexibly accounts for both the “push” and “pull” effects discussed in section 6.1. To measure occupation wages, we take an applicant’s reported SOC code and assign them the in-sample, leave-one-out, average three-digit SOC code earnings in that given year.⁴⁰ Those in a military SOC code receive average military earnings, which does not vary across military occupations. We repeat this for each year since ASVAB and then construct average 11-20 occupation wages. We then measure causal effects on this outcome by using “average occupation wage” as the outcome variable in model 2. Thus, the causal effect on this variable gives exactly the change in occupation specific wage premia implied by how the applicant’s career has shifted 11-20 years post-ASVAB.

Benchmarking long-run causal effects on wages to long-run causal effects on occupation-wages therefore yields an apples-to-apples comparison that accounts for the heterogeneous nature of the long-run counterfactual career across cutoffs. Intuitively, we expect to find that military occupational experiences which, 11-20 years later, have caused applicants to move out of low-paying occupations and into-high paying occupations will also have larger effects on realized earnings.

Figure 5 contains our key result. In this figure, each point is a cutoff. The Y-axis denotes the 2SLS estimates from model 2 and hence corresponds to the causal effect of that early career occupational experience on realized earnings. The X-axis denotes the corresponding 2SLS estimate on leave-out average occupation wages. The dashed line is the 45-degree line, and the solid line is the line of best fit. Appendix Figure F.13 gives an identical plot, but using the reduced form estimates on the X- and Y-axes instead of the 2SLS estimates.

Figure 5 reveals the highly heterogeneous impact of different early career occupational experiences on long-run wages. Cutoffs that in the long-run end up shifting people away from low-wage occupations and into high-wage occupations also cause large increases in realized earnings, and vice-versa. For example, the fuzzy RD estimates of early career experiences in infantry/construction or computers/electronics suggest they increase earnings by \$23,805 (se=9,186) and \$11,514 (se = 7,638), respectively. These causal effects are mirrored in the corresponding changes in occupational average wages (\$25,279 and \$10,214). At the bottom end, we find that military careers related to warehousing/HR actually *reduce* earnings by \$13,600 (se = 5,336), with causal effects that are also mirrored in the corresponding changes in occupational wage premia (-\$7,516).⁴¹ And for many military occupational experiences, like those related to food service/truck driving (\$3944, se=7,320), we find no detectable impact on either earnings or occupational average wages. Note that these null effects occur despite the fact that these same cutoffs generate consequential changes in career path. This is consistent with

⁴⁰As robustness, we also report similar findings using SOC-specific average wages from a random sample of all Americans in the tax data, as opposed to the in-sample, leave-out average (see figure F.14). We also show comparable results using leave-out average industry wages, based on 6-digit EIN numbers mapped to NAICS, which are observed for everyone with a W-2 regardless of the occupation they reported on their Form 1040 (see figure F.15).

⁴¹In the military, warehousing occupations are held by enlisted personnel (not officers) and are known as “logistics.”

the cutoffs that generate null effects generating changes across occupations with otherwise similar average occupational earnings.⁴²

This finding is also mirrored in raw regression discontinuity plots. Panels (d) and (e) of figure 2 display pooled, reduced form discontinuities split out according to whether the cutoff is in the top 20% or bottom 20% of the distribution of causal effects on the occupation wage variable. Thus, Panel (d) displays the discontinuity for the 20% of cutoffs that created the largest moves away from lower-paying occupations and towards higher-paying occupations. Panel (e) displays the discontinuity for the 20% of cutoffs that created the largest moves from higher-paying to lower-paying occupations. In both cases, we find precisely estimated discontinuities that echo the broader pattern highlighted in figure 5. Panel (e) also makes evident that we estimate earnings *decreases* at several cutoffs. This could occur if applicants do not anticipate which occupations are more economically lucrative (e.g. information is limited) or if they value non-pecuniary career features (e.g. amenities).

The tight correspondence between changes in occupational earnings premia and the actual average effect on earnings in figure 5 suggests that much of the heterogeneity in wage effects across cutoffs is well explained by the way that different early life occupational experiences change career paths. Indeed, the implied changes in occupation wage premia can explain an enormous amount of the heterogeneity in the actual causal effect on earnings across cutoffs. For example, a simple regression of the fuzzy RD estimates on earnings on the fuzzy RD estimates on occupational average wages yields an adjusted R^2 of 0.638, showing that over 60% of the causal variation in earnings across cutoffs is explained by simply examining the change in occupational average earnings implied by the underlying causal effect on career path (see Table 2).

Perhaps more strikingly, the slope coefficient from a univariate regression of the causal earnings gains on causal changes in occupational average wages is indistinguishable from one ($\beta = 0.975$). This suggests that changes in occupational average wages are also effectively forecast unbiased for actual causal effects on earnings. In other words, the point estimates suggest that each dollar of wage premia translates essentially 1-1 into an additional dollar of actual, realized earnings. However, we caution against over-interpreting this result – the associated standard error is 0.208, and hence we do not have the precision necessary to rule out some amount of forecast bias. That said, even the lower bound of the 95% confidence interval would suggest that differentials in average occupational wages translate into average causal effects on earnings at a rate of 56 cents on the dollar. At a minimum, this finding suggests that changes in occupation-specific wage premia are tightly linked to economic well-being.

This tight correspondence between occupational average wages and causal effects on realized earnings begs a natural follow-on question: do the cutoffs cause other consequential changes in the lives of applicants that are similarly connected to the impact on earnings? And how much of the tight link between earnings and average occupational wages can be explained by the fact that changes in occupational earnings may be correlated with changes in other important outcomes impacted by the cutoffs? However, before turning to these questions, we

⁴²We note, of course, that these individual point estimates should not be interpreted as isolating causal effects of specific occupations or compared directly to one another as a result of the heterogeneous nature of the counterfactual. To see why, suppose that combat/construction pulls you out of particularly low wage counterfactual long-run occupations, while the computer/electronic threshold does not. Even if the two occupations paid similarly, the effects on earnings and average occupation wages would be larger for combat/construction. We return to this point in Section 6.4. Here we simply seek to establish that the cutoffs that shift you out of low-paying occupations into higher-paying occupations in the long-run also yield comparable earnings gains.

first discuss robustness of the results in this section.

Robustness. These results are robust to a number of alternative modeling choices and variable definitions. For example, Table 3 presents estimates that are identical to panels (b) and (c) of Table 2, except that we vary the bandwidth and order of the polynomial in the RD models used to estimate the impact of the individual cutoffs on the key variables of interest. In all cases, the results are similar.

Our results are also robust to alternative ways of measuring both earnings and occupation-specific average wages. Panel (a) of appendix figure F.14 is identical to figure 5, except that we use an alternative measure of occupation wages constructed from a 50% random sample pulled from the universe of tax records for those aged 31-44. The fact that the findings are nearly identical shows that these results are not a quirk of the particular sample used to build our benchmark. Panel (b) of figure F.14 is also identical to figure 5, except that we use log-wages instead of wages when constructing the causal effects plotted on the y-axis, which nets out extensive margin effects. Thus, the results are broadly similar even if we condition on working.

Finally, we note here that measurement error does not impact these findings. The fact that the X-axis is constructed using an estimated variable makes it appear as though measurement error could matter here (e.g. through attenuation bias). In Appendix Figure F.13, we show that we obtain similar results building this figure with the reduced form estimates (instead of the 2SLS estimates). The reduced-form on reduced-form version of this figure suggests that the slope is 0.999 (se = 0.155) with a corresponding adjusted R^2 of 0.702. This is significant because the reduced-form on reduced-form slope from Appendix Figure F.13 is numerically equivalent (up to a weighting) to a two-stage least squares regression on the micro-data that tests for forecast bias in the spirit of [Angrist et al. \(2017\)](#) and [Deming et al. \(2014\)](#). See Appendix D for a restatement of the proof and a discussion of the connection between the reduced-form on reduced-form plot (Appendix Figure F.13) and the “IV-on-IV” plot (Figure 5).⁴³ Importantly, none of the variables involved on the left hand side or right hand side of the micro-regression are estimated quantities; hence, they cannot be impacted by estimation error from the pre-step. However, out of an abundance of caution, we also estimate the forecast bias slope coefficient directly using the corresponding micro data regression (see Appendix Table F.7). Our key findings are virtually unchanged. Moreover, the over-id p-value from the micro-data regression suggests that we cannot reject a model in which occupation wage gains translate into earnings gains approximately dollar-for-dollar *at each cutoff* ([Sargan, 1958](#); [Angrist et al., 2016](#)). Hence, our core result is robust to approaches that cannot be impacted by estimation error from the pre-step.

6.3 Other predictors of long-run success

The finding that long-run changes in implied occupational wage premia can explain a substantial amount of the heterogeneity across cutoffs in the impact on realized earnings begs a natural follow-up question: “Do early career experiences in the military cause other changes that predict economic success?” To answer this question, we use both the tax data as well as several supplementary data sources including the ONet ([Autor et al., 2003](#)) database of occupation specific skills, the National Student Clearinghouse (NSC), the

⁴³If the first stage is orthogonal to the underlying causal effects, then the “IV-on-IV” slope in Figure 5 is also statistically equivalent, up to a weighting, to the two-stage least squares micro-data regression (see appendix D for mathematical details). We find evidence consistent with this orthogonality condition in our data: Panel (a) of Appendix Figure F.18 reveals little relationship between 2SLS causal effects on earnings and the corresponding first-stages.

American Community Survey (ACS), the BLS, the Current Population Survey (CPS), and state / federal regulations. The combined data allows us to capture long-run outcomes related to the industry of employment (e.g. average industry wages, still in the Army, and working for the government); as well as SOC-based, long-run career outcomes like occupational non-routine cognitive analytical task intensity (a proxy for high-skill work); occupational routine manual task intensity (a proxy for low-skill work); and measures of bargaining power, specifically the share of workers in an occupation who belong to a union. We also use the combined data to explore educational outcomes like college attendance and degree attainment. Finally, we return to the internal Army personnel data to measure features of the MOS specific to training (length, cost, college credits, etc.) provided in AIT. We also assign each MOS (based on its primary SOC) a certification index that measures if the occupation typically uses or requires credentials. Last, we also measure a host of other Army-specific concepts that could vary across military occupations unlocked by the cutoffs, including total years served; combat MOS designation; the share of years on active duty 11-20 years after the ASVAB date; receipts of Veteran's Affairs Disability Compensation (VADC); joining the Army at all; length of time spent deployed to a combat zone; and initial contract length. See appendix B for precise variable definitions mapped to the specific data sources we used to build them, with appendix table F.4 giving a concise summary.

We find that changes in long-run average industry earnings, holding a government (non-military) job, making a career out of the military, high-skill task content, post-secondary degree attainment, and the likelihood of joining an MOS related to a civilian occupation that requires academy training (e.g. the police or firefighter academy) also explain some of the heterogeneity across cutoffs; however, changes in other outcomes, such as those that are idiosyncratic to military service, do not. To explore the connection between other outcomes and the earnings heterogeneity, we estimate univariate regressions identical to those used to generate the results in panel (c) of table 2, except that the causal effect on “occupation-earnings” is replaced with the causal effect on the alternative occupational features described earlier in this section.

Figure 6 contains the results. Each bar displays adjusted R^2 values from one of these univariate regressions. Corresponding scatter-plots, similar to figure 5, are contained in appendix exhibits F.15 – F.17. As a benchmark, the bar labeled “Occupation Wages” replicates the adjusted R^2 values from panel (c) of table 2. From the figure, we can see that changes in the average wages of the industry of employment, the likelihood of obtaining a government job, and making a career out of the military explain 51%, 28%, and 19%, respectively, of the heterogeneity across cutoffs. We also find evidence that high-skill work matters. Non-routine task intensity, along with measures of post-secondary degree attainment, explain 13% and 11% of the variation. Changes in college attendance, the nature of low-skill work, and unionization rates, do not explain causal effects. And while we find evidence that starting a career linked to occupations that require academy training (e.g. police / fire academy) matters, there is little evidence that cutoff-induced changes in other forms of certification predict long-run wage gains.

Changes in outcomes specific to Army provided occupational training (AIT) are less predictive of long-run economic success. An interesting theoretical question is whether the long-run impacts we detect are more closely connected to the training itself, or the subsequent on-the-job experience. To explore this, we examine variables such as the number of vocational or college credits the military recommends be awarded for the training; the length of time the training lasts; and its total cost. The adjusted R^2 values from these regressions

range from -0.01 to 0.05, demonstrating that they are less tightly linked to long-run causal wage gains. This is consistent with the institutional details described in section 3. Training itself typically lasts only 1-6 months, which is small relative to the total length of time a typical soldier will spend performing their occupation on-the-job (3-4 years). So, while it is not possible for us to be definitive on the question of training versus on-the-job experience given the nature of our quasi-experimental design, the results and institutional details are consistent with the idea that the subsequent, on-the-job experience is the primary channel.

Changes in outcomes specific to Army service are not predictive of causal impacts on long-run well-being. *Ex ante*, it is reasonable to believe that characteristics like combat exposure, receipt of disability compensation, the length of time served, or the likelihood of joining the Army *at all* could vary across Army occupations while also generating independent effects on long-run realized earnings. If that were the case, we would expect to find that causal effects on these features should be highly correlated with the realized causal effect on actual earnings. However, along virtually every outcome we have examined that is idiosyncratic to experiences while enlisted in the Army, the adjusted R^2 is negative, suggesting that they do no better than noise at explaining the heterogeneity across cutoffs in long-run wage effects.

The explanatory power of changes in some of the outcomes in figure 6 raises an important secondary question related to mechanisms. One reason average occupational earnings could translate into realized improvements in earnings is simply that they are correlated with changes in other characteristics (i.e. likelihood of employment, making a career out of the military, routine/non-routine tasks intensity, educational requirements, etc.). For example, it could be that occupation is important for long-run earnings only insofar as occupations tend to be clustered in different industries with very different average earnings or having very different risk for an unemployment spell.

Motivated by this question, we next ask whether changes in occupational average wages continue to translate into realized effects on earnings even among cutoffs that create similar changes in other variables like industry average wages. To implement this analysis, we estimate bivariate regressions of the form:

$$Y_c = \alpha + \beta Z_c + \pi X_c + e_c \quad (4)$$

Where Y_c is the 2SLS estimate (model 2) of the causal effect on realized earnings generated by the occupations unlocked at cutoff c ; Z_c is the corresponding estimate of the causal effect on occupation-specific average wages; and X_c is a control variable that corresponds to the estimated causal impact on another outcome (e.g. average industry wages).

Table 4 contains the results. Panel (A) gives the unconditional relationship as a baseline. Panel (B) gives results conditional on the corresponding change in the likelihood of working. Panel (C) gives results conditional on corresponding changes in industry features. Panel (D) gives results conditional on corresponding changes in career outcomes. Panel (E) gives results conditional on corresponding changes in educational outcomes. Panel (F) gives results conditional on corresponding changes in certification requirements. Panel (G) gives results conditional on corresponding changes in training features. Panel (H) gives results conditional on

corresponding changes in military specific outcomes.

In all cases, the estimates suggest that improvements in occupational wage premia continue to be tightly linked to causal effects on realized earnings. Even in the case of variables like industry average wages and the likelihood of working, we observe that improvements in occupational earnings premia continue to translate into long-run impacts on economic well-being at a rate of over 80 cents on the dollar. Moreover, we also note that 1 is contained in the 95% confidence interval of virtually every model. However, we caution against over-interpreting this fact, since many of the associated confidence intervals are large, such that we also cannot rule out substantial deviations from 1 in several of the conditional relationships. That said, the fact that we continue to find a tight link between changes in occupation wages and earnings even conditional on the corresponding changes in these other consequential variables demonstrates the substantial importance of occupation-specific wage premia for explaining economic well-being.

Taken together, the results up to this point paint a clear picture. In the long-run, early career experiences are important for economic well-being. They have a large impact on subsequent career paths. To the extent that these long-run career shifts also move individuals into higher (lower) earnings occupations, they also substantially improve (reduce) economic well-being. Impacts on average industry wages, the likelihood of obtaining a government job, making a career out of the Army, high-skill work, degree attainment, and academy training are also indicative of long-run economic success.

6.4 Which types of early career experiences are most consequential for long-run economic well-being?

In this subsection, we present two final exercises that better characterize the Army occupations that put young adults on the path to success. Notably, our results in the preceding sections demonstrate that early career experiences have consequential, long-run implications for earnings when they move individuals across long-term career paths with different average earnings. However, these findings pertain to the long-run career impacts of early occupational experiences and do not directly characterize which Army occupations most improve economic well-being. The exercises in this section aim to shed light on this question.

In our first exercise, we assign an *ex-ante*, long-run expected wage value to each Army occupation, based on the average wages in the occupations crosswalked to that MOS. In the spirit of figure 5, we then ask whether we can predict long run earnings gains at each cutoff using only information about whether or not that cutoff shifts you from a military occupation with lower-*expected* future earnings and into a military occupation with higher-*expected* future earnings. We find that expected future earnings translates into actual changes in earnings at a rate of 78 cents on the dollar.

In our second exercise, we estimate a “partially pooled” multiple endogenous variable model of the causal effect of different military careers on long-run earnings. Under strong assumptions about the nature of treatment effect heterogeneity, this will allow us to identify the impact of groups of *specific* military careers on long run outcomes relative to a consistently defined counterfactual. We find that Army careers in military occupations related to computers/electronics, combat/construction, and health/police/intelligence all substantially improve long-run earnings relative to not joining the Army. However, Army careers in logistics management that often

involve extensive work in warehouses and that tend to be linked to below average paying civilian careers do not improve long-run earnings relative to not joining the Army.⁴⁴

Do early career occupations linked to high-wage career paths improve long-run well-being? To answer this question, we first need to measure the expected future earnings of the occupations moved at each cutoff. We do so by linking each MOS to the average occupation-wage across all SOC codes mapped to that MOS. This procedure will be noisier and less precise than our long-run analysis for two reasons. First, in our long-run analysis, we benchmark the overall impact on long-run wages to the impact on long-run occupation wages, which is based on contemporaneously reported SOC-codes and therefore does not rely on the linkages between military and civilian careers embodied in our crosswalk. Additionally, the long-run occupational effects potentially capture more complex and dynamic occupational changes than what happens at entry. Nevertheless, it is instructive to ask whether the long-run earnings gains are predictable ex-ante.

To this end, we calculate “military occupation implied future earnings” as the average occupation-wage across all SOC codes mapped to each MOS. This means that if an MOS is linked to a primary, secondary, and a single supervisory SOC code, we assign it the simple mean of the three SOC-level average long-run occupation wages, and we assign each soldier that value for their respective entry MOS. Thus, causal effects on this variable measure the extent to which early career experiences in the military induced by a cutoff shift individuals from a low expected wage career path and into a high expected wage career path.

Figure 7 displays the results. In this figure, each observation is one of our 35 cutoffs. The Y-axis is the corresponding 2SLS estimate from model 2 that captures the impact of joining an occupation unlocked by that cutoff on average earnings 11-20 years later. The X-axis is the corresponding estimate of the causal effect on MOS-level implied future earnings. Thus, the X-axis reveals whether the military occupations unlocked at a given cutoff moved applicants out of a career path with low expected future earnings and into a career path with high expected future earnings. The solid line represents the line of best fit. The dashed line represents the 45-degree line.

Figure 7 reveals that the cutoffs which move applicants into expected higher wage careers also improve long-run realized earnings. We find a strong, positive relationship that lines up closely with the 45-degree line. In fact, the slope coefficient is 0.78 (se = 0.335), suggesting that expected future earnings translate into actual future earnings at a rate of 78 cents on the dollar. Unsurprisingly, given the additional sources of measurement error in this exercise, we find that the amount of variation in wage effects that can be explained by changes in implied future earnings is 17.3% – substantial, but not nearly as large as what we find in the long run.

To get a sense of which MOS have high implied future earnings based on this procedure, Figure F.19 collapses the MOS-level future earnings to the larger Career Management Field (CMF) level, weighting by the number of people in each MOS. We take the 15 largest CMFs from Table F.1 and pool the various combat CMFs together. At the top end, Army occupations in combat arms and police, signal corps (i.e. computers/electronics),

⁴⁴Importantly, military careers in logistics for enlisted soldiers should not be confused with the similarly named role performed by officers. For officers, who enter the Army with a four-year college degree, military logistics is more closely aligned with supply chain management. For enlisted soldiers, who generally do not enter the Army with a four-year college degree, logistics involves performing the day-to-day work on the warehouse floor that is necessary to execute the supply chain tasks as directed by those same officers.

and medical roles all have implied 11-20 earnings above \$43,000. On the lower end, with implied earnings between \$35,000 and \$39,000, are Army occupations in Transportation (e.g. drivers), Human Resources, and Quartermaster (e.g. warehouse operations and food preparation jobs). Our findings suggest that, on average, moving people out of the occupations with lower expected wage trajectory careers (e.g. warehouse operations) and into occupations with higher expected wage trajectory careers (e.g. computers/electronics) yields long-run earnings benefits.

Next, we probe whether other MOS-level characteristics can predict or help explain long-run earnings gains in the spirit of our earlier exercises (e.g. Figure 6 and Table 4). Appendix Table F.11 contains the results. When the causal effects on earnings are regressed on the causal effects on these additional characteristics, we typically find low adjusted R^2 's. This indicates that it is generally difficult to find other single-variate predictors (besides expected future wages) of long-run earnings gains. Our baseline slope coefficient of 0.78 and adjusted R^2 of 0.17 shows more sensitivity to the one-by-one inclusion of these other MOS-level characteristics than occupation-wages, which is unsurprising given the additional sources of measurement error, but if anything, the coefficient on implied future earnings tends to increase, and we typically can't reject a coefficient of 1.

While these results are, arguably, unsurprising given our findings related to the overall impact that early career experiences have on career trajectory, it is reassuring for policy that we can indeed predict future wage gains using these imperfect measures of eventual average expected wages. These results suggest that the commonly used heuristic of comparing average wage rates across occupations can serve as a good guide for designing active labor market policies that will put young adults on the pathway to success.

A “partially-pooled” multiple endogenous variable model. To this point, we have shown that early career experiences linked to higher-wage career paths tend to improve long-run well-being. However, this approach does not allow us to directly isolate the effects of starting out one’s career in one specific set of military occupations relative to another. This is because differences across cutoffs could be due to *either* real treatment effect heterogeneity (i.e. different short-run military occupations have different long run effects) *or* the heterogeneous nature of the counterfactual (i.e. different cutoffs pull applicants out of counterfactual short-run occupations with better / worse career prospects). As discussed previously, this complication is compounded by the fact that, as with most regression discontinuity designs, a number of our cutoff-specific estimates on earnings have large standard errors, which will make it difficult to distinguish signal from noise when comparing across cutoffs.

In this section, we address both of these challenges by developing a “partially-pooled” multiple endogenous variable model of the impact of specific military occupations. Under strong assumptions about the nature of treatment effect heterogeneity (which we discuss in more detail below), a jointly estimated two-stage least squares model will deliver treatment effect estimates that are “normed” to a consistent short-run counterfactual career – in our case, the reference group will be individuals who remain civilians after taking the ASVAB. To improve power, we will “average” the effects of similar occupations into a smaller number of more precise point estimates.

To accomplish this second goal, we group the military occupations in our data into nine mutually exclu-

sive career groups: Warehouse Operations/HR, Mechanics/Maintenance, Transport/Food Preparation, Utilities/Equipment Management, Health/Police/Intelligence/Other Specialists, Combat/Construction, Computers/Electronics, All Other Army Occupations, and Civilian (did not join the Army). These occupation groups are largely derived by taking the cutoff-based groups of military occupations used for our results in the first portion of the paper (see section 4) and then further consolidating by additionally grouping them across cutoffs whenever cutoffs occur in the same line score. This approach makes sense since, in theory, each ASVAB line score is meant to assess aptitude for a specific skill at baseline and hence the careers attached to it are comparable on that margin (see Appendix A for more discussion of the line scores). Importantly, our procedure results in multiple cutoffs tied to each occupation group which will allow us to run an over-identification test that allows us to check empirically whether these groups are sensible / consistent with the data. The list of MOS in each career group can be found in Table F.15, with more detailed discussion of the data construction, as well as robustness to these various choices, found in Appendix E.

We estimate the following multiple endogenous variable two-stage least squares model:

$$Y_i = \beta D_i + f(\mathbf{X}_i) + \epsilon_i \quad (5)$$

$$D_{ig} = \Pi_g Z_i + h_g(\mathbf{X}_i) + \eta_{ig}, \quad g = 1, 2, \dots, 8 \quad (6)$$

Where Y_i is 11-20 earnings; $D_i = \{D_{ig}\}$ is a vector of 8 treatment indicators, one for each of the 8 occupation groups with “remaining a civilian” constituting the omitted category; $\mathbf{X}_i = \{X_{il}\}$ is a vector of the various line scores; and f maps line scores into the series of piecewise linear and quadratic polynomials around the various RD thresholds as implied by our baseline, individual RD models. We instrument for the treatment variables with 8 first stage equations of the form shown in 6, where D_{ig} takes a value of 1 if applicant i enlists in any occupation contained in group g ; $Z_i = \{Z_{ic}\}$ is a vector of indicator variables that take a value of one if individual i crosses line score threshold c ; and h_g maps line scores into the series of piecewise linear and quadratic polynomials around the various RD thresholds as implied by our baseline, individual RD models. In order to increase the power of our model, we also include the AFQT thresholds from [Greenberg et al. \(2022\)](#). Since these cutoffs generate a strong first stage on joining the Army, they create helpful extensive margin variation that affords us the power to benchmark these estimates relative to a consistent civilian counterfactual.

This model identifies an interpretable causal parameter under strong assumptions. In addition to the standard RD assumptions outlined earlier in our analysis of the individual thresholds, this approach also requires that treatment effects as defined by our eight career groups are constant and linear. This assumption could fail in two ways. First, it could be that different occupations within a group have different causal effects. If any of our instruments cause individuals to move across occupations with very different impacts on earnings *within* a group, this would violate the implicit exclusion restriction embedded in our constant linear effects assumption. Second, it could be that treatment effects are heterogeneous in ways that are not captured by the eight treatment variables we have defined. This would be a form of model mis-specification, since it implies a

violation of the functional form assumption (i.e. linearity) imposed by our model.

In support of these assumptions, we implement an over-identification test. Our method of pooling military careers within these different groups leaves us with more instruments than endogenous regressors; hence we are overidentified. Intuitively, if different subsets of instruments produce drastically different estimates of the vector of causal effects (β in model 5), then it must be due either to an exclusion restriction violation or a functional form violation. This is effectively what is tested using a classic Sargan-Hansen J-test (Wooldridge, 2010). We implement this test and fail to reject ($P = 0.14$). Up to sampling variation, this is consistent with the pooled model being sufficiently rich to account for the most important forms of heterogeneity.⁴⁵

Figure 8 displays the results from estimating our multiple endogenous variable 2SLS model (Appendix Table F.13 contains the corresponding point estimates). Each bar in Figure 8 is the estimated causal effect (β) from model 5 with the civilian (non Army-joiners) mean added back-in for scale – thus the height of each bar represents the implied long-run earnings of each military occupation group, with causal effects encoded in differences between the bar height and the horizontal line (which represents the civilian mean). Occupation groups with treatment effects that are statistically significant (relative to being a civilian) are bolded.

First, and broadly consistent with Greenberg et al. (2022), we see that joining the Army appears to be, on average, beneficial. The Army occupational categories are jointly significant ($P = 0.013$) and the point estimates suggest that most occupation groups make applicants better off than they would have been had they remained a civilian. Among these broad Army occupational groups, there is also evidence that the different Army career groups have different long-run impacts. The group with the smallest (and negative, $-\$3,610$) estimated earnings differential is the Warehouse Operations (i.e. quartermaster jobs in logistics and supply) and Human Resources group. For example, this point estimate is statistically different from the three largest point estimates for Health/Police/Intelligence ($\$8,160$; with a p-value on the difference = 0.007), Combat/Construction ($\$8,418$; with a p-value on the difference = 0.022), and Computers/Electronics ($\$10,679$; with a p-value on the difference = 0.073). The differences between Warehouse Operations/HR and these top 3 groups are thus on the order of $\$12,000$ - $\$14,000$, though sometimes nosily estimated. Generally speaking, the military occupations in the top and bottom occupational groups considered here also tend to line up consistently with those in the top and bottom of implied earnings from the preceding exercise (Figure 7 and F.19). For example, the effects on these career groups are consistent with Army occupations in the quartermaster corp and transportation being towards the lower end, and with combat arms, police, signal corps, intelligence and medical occupations towards the higher end. While these results are necessarily suggestive given the stronger assumptions and often large confidence intervals, this model nevertheless broadly supports our key finding: that shifting people out-of and into occupations with higher expected earnings can increase their economic well-being in the long-run.

⁴⁵We note here that the first-stage estimates from this model are also sensible. Appendix Table F.12 shows the 8 first stages (equation 6). We always find positive and significant effects where expected, indicating that all the instruments are indeed pushing people into the relevant Army career groups. Also as expected, the ‘off-diagonal’ effects are generally negative or zero and smaller, indicating that the cutoffs pull people from a more diffuse set of counterfactual Army occupations. The AFQT cutoffs induce entry into most, but not all, Army career groups, including the < 4% of the sample in other Army jobs. Appendix Table F.13 shows what happens when we drop the category for other army jobs and instead consider the omitted category to be civilian or other army job; broad patterns are stable.

7 Conclusion

In this paper, we provide evidence that early career occupational experiences in the military can improve long-run outcomes for the millions of young-adults that are not destined for college and otherwise difficult to reach. We find that early career occupational experiences generate a substantial amount of path dependence, with point estimates that suggest a 19p.p. increase in the likelihood of being observed in an identical or closely related occupation as much as 20 years later. These long-run career shifts have important implications for earnings, with long-run earnings changes that are well-predicted by changes in occupational earnings premia. As a result, early experience in short-run careers that are linked to high-wage career paths appears to constitute a key channel for economic success. Thus, a key take-away is that it's not too late to improve long-run economic outcomes for a large group of less-educated young adults that generally face more limited economic opportunities.

The military training we study bears close resemblance to a wide array of educational programs, vocational training opportunities, and active labor market policies across the globe. Most closely related is sectoral job training which, much like the military, involves a concentrated period of training followed by practical, on the job experience. Consistent with our results, a small number of randomized controlled trials have also found this model successful, provided the program focused on high-paying careers like IT and nursing ([Katz et al., 2022](#)). There is also promising recent evidence that large scale active labor market programs can work when individuals are induced to complete them and ultimately change careers, findings that are consistent with our results ([Humlum et al., 2023](#)). More broadly, countries like Austria, Germany, and Switzerland enroll as much as 60% of students in vocational tracks meant to provide students with occupation specific skills they can take directly to the labor market. In Norway, for example, recent work has found that disadvantaged young men in particular can benefit from this style of learning ([Bertrand et al., 2021](#)).

Thus, our analysis of military occupations points to rich space for future work on the careers and well-being of the non-college educated. For example, our results suggest that continued study of active labor market policies could yield transformative and broadly applicable policy interventions. It is also possible that lighter-touch interventions, such as targeted employer subsidies (e.g. [Dillon et al., 2025](#)) or information interventions meant to move young adults into successful careers, could also be effective. Our findings suggest that the key is to design these programs in ways that lead to durable connections with high-quality careers that pay well – how, precisely, to accomplish that goal is an open question.

References

Abraham, Lisa, Christine Mulhern, Isaac M Opper, Prateek Puri, and Elizabeth A Roth, *Workers' First Jobs and Long Term Wage Growth: Variation by Gender and Race/Ethnicity*, RAND, 2024.

Acemoglu, Daron and David Autor, “Skills, tasks and technologies: Implications for employment and earnings,” in “Handbook of labor economics,” Vol. 4, Elsevier, 2011, pp. 1043–1171.

— **and Pascual Restrepo**, “Automation and new tasks: How technology displaces and reinstates labor,” *Journal of economic perspectives*, 2019, 33 (2), 3–30.

— **and** —, “Unpacking skill bias: Automation and new tasks,” in “AEA Papers and Proceedings,” Vol. 110 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2020, pp. 356–361.

Altonji, Joseph G, Lisa B Kahn, and Jamin D Speer, “Cashier or consultant? Entry labor market conditions, field of study, and career success,” *Journal of Labor Economics*, 2016, 34 (S1), S361–S401.

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel, “The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects,” in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 5, Elsevier, 2016, pp. 305–396.

Anderson, D Mark and Daniel I Rees, “Deployments, Combat Exposure, and Crime,” *The Journal of Law and Economics*, 2015, 58 (1), 235–267.

Andrews, Isaiah, James H Stock, and Liyang Sun, “Weak instruments in instrumental variables regression: Theory and practice,” *Annual Review of Economics*, 2019, 11, 727–753.

Andrews, Rodney J., Scott A. Imberman, Michael F. Lovenheim, and Kevin M. Stange, “The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability,” *Review of Economics and Statistics*, 2025. Forthcoming.

Angrist, Joshua and Michal Kolesár, “One instrument to rule them all: The bias and coverage of just-ID IV,” *Journal of Econometrics*, 2024, 240 (2), 105398.

Angrist, Joshua D., “Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records,” *American Economic Review*, 1990, 80 (3), 313–336.

—, “Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants,” *Econometrica*, 1998, 66 (2), 249–288.

— **and Guido W. Imbens**, “Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity,” *Journal of the American Statistical Association*, 1995, 90 (430), 431–442.

Angrist, Joshua D and John H Johnson IV, “Effects of Work-related Absences on Families: Evidence from the Gulf War,” *ILR Review*, 2000, 54 (1), 41–58.

— **and Jörn-Steffen Pischke**, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2009.

— **and Peter Hull**, “Instrumental variables methods reconcile intention-to-screen effects across pragmatic cancer screening trials,” *Proceedings of the National Academy of Sciences*, 2023, 120 (51), e2311556120.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, “Identification of Causal Effects Using Instrumental Variables,” *Journal of the American Statistical Association*, 1996, 91 (434), 444–455.

Angrist, Joshua D, Peter D Hull, Parag A Pathak, and Christopher R Walters, “Leveraging lotteries for school value-added: Testing and estimation,” *The Quarterly Journal of Economics*, 2017, 132 (2), 871–919.

Angrist, Joshua D., Peter Hull, Parag A. Pathak, and Christopher R. Walters, “Interpreting Tests of School VAM Validity,” *American Economic Review: Papers and Proceedings*, 2016, 106 (5), 388–392.

—, **Stacey H. Chen, and Brigham R. Frandsen**, “Did Vietnam Veterans Get Sicker in the 1990s? The

Complicated Effects of Military Service on Self-Reported Health," *Journal of Public Economics*, 2010, 94 (11-12), 824–837.

—, —, and **Jae Song**, "Long-Term Consequences of Vietnam-Era Conscription: New Estimates Using Social Security Data," *American Economic Review*, 2011, 101 (3), 334–338.

Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano, "Ex ante returns and occupational choice," *Journal of Political Economy*, 2020, 128 (12), 4475–4522.

Autor, David, "Skills, Education, and the Rise of Earnings Inequality among the 'Other 99 Percent,'" *Science*, 2014, 3.

—, **Claudia Goldin, and Lawrence F Katz**, "Extending the race between education and technology," in "AEA papers and proceedings," Vol. 110 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2020, pp. 347–351.

Autor, David H and Michael J Handel, "Putting tasks to the test: Human capital, job tasks, and wages," *Journal of labor Economics*, 2013, 31 (S1), S59–S96.

—, **Frank Levy, and Richard J Murnane**, "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 2003, 118 (4), 1279–1333.

—, **Mark Duggan, Kyle Greenberg, and David S Lyle**, "The Impact of Disability Benefits on Labor Supply: Evidence from the VA's Disability Compensation Program," *American Economic Journal: Applied Economics*, 2016, 8 (3), 31–68.

Bäckström, Peter and Niklas Hanes, "The Impact of Peacekeeping on Post-deployment Earnings for Swedish Veterans," *Scandinavian Working Papers in Economics No. 1010*, 2023.

Barr, Andrew, "From the battlefield to the schoolyard: The short-term impact of the Post-9/11 GI Bill," *Journal of Human Resources*, 2015, 50 (3), 580–613.

—, "Fighting for education: Financial aid and degree attainment," *Journal of Labor Economics*, 2019, 37 (2), 509–544.

—, **Laura Kawano, Bruce Sacerdote, William Skimmyhorn, and Michael Stevens**, "You can't handle the truth: The effects of the post-9/11 GI bill on higher education and earnings," Technical Report, National Bureau of Economic Research 2021.

Bertrand, Marianne, Magne Mogstad, and Jack Mountjoy, "Improving educational pathways to social mobility: evidence from Norway's reform 94," *Journal of Labor Economics*, 2021, 39 (4), 965–1010.

Bingley, Paul, Petter Lundborg, and Stéphanie Vincent Lyk-Jensen, "The opportunity costs of mandatory military service: Evidence from a draft lottery," *Journal of Labor Economics*, 2020, 38 (1), 39–66.

Borgschulte, Mark and Paco Martorell, "Paying to avoid recession: Using reenlistment to estimate the cost of unemployment," *American Economic Journal: Applied Economics*, 2018, 10 (3), 101–127.

Bruhn, Jesse, Kyle Greenberg, Matthew Gudgeon, Evan K Rose, and Yotam Shem-Tov, "The effects of combat deployments on veterans' outcomes," *Journal of Political Economy*, 2024, 132 (8), 2830–2879.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, 82 (6), 2295–2326.

Card, David and Ana Rute Cardoso, "Can Compulsory Military Service Raise Civilian Wages? Evidence from the Peacetime Draft in Portugal," *American Economic Journal: Applied Economics*, 2012, 4 (4), 57–93.

Case, Anne and Angus Deaton, "The great divide: education, despair, and death," *Annual Review of Economics*, 2022, 14 (1), 1–21.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinu-

ity," *The Stata Journal*, 2018, 18 (1), 234–261.

Cesur, Resul, Alexander Chesney, and Joseph J Sabia, "Combat Exposure, Cigarette Consumption, and Substance Use," *Economic Inquiry*, 2016, 54 (3), 1705–1726.

— and **Joseph J Sabia**, "When War Comes Home: The Effect of Combat Service on Domestic Violence," *Review of Economics and Statistics*, 2016, 98 (2), 209–225.

—, —, and **Erdal Tekin**, "Post-9/11 War Deployments Increased Crime among Veterans," *The Journal of Law and Economics*, 2022, 65 (2), 279–310.

Coile, Courtney, Mark Duggan, and Audrey Guo, "To Work for Yourself, for Others, or Not At All? How Disability Benefits Affect the Employment Decisions of Older Veterans," *Journal of Policy Analysis and Management*, 2021, 40 (3), 686–714.

Conlon, John J. and Dev Patel, "What Jobs Come to Mind? Stereotypes about Fields of Study," 2025. Unpublished manuscript, August 31, 2025.

Corak, Miles and Patrizio Piraino, "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, January 2011, 29 (1), 37–68.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment," *The quarterly journal of economics*, 2013, 128 (2), 531–580.

Cullen, Julie Berry, Gordon B. Dahl, and Richard De Thorpe, "Job Mismatch and Early Career Success," 2025.

Dahl, Gordon B, Andreas Kotsadam, and Dan-Olof Rooth, "Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams," *The Quarterly journal of economics*, 2021, 136 (2), 987–1030.

Deming, David J, "The growing importance of social skills in the labor market," *The Quarterly Journal of Economics*, 2017, 132 (4), 1593–1640.

— and **Kadeem Noray**, "Earnings dynamics, changing job skills, and STEM careers," *The Quarterly Journal of Economics*, 2020, 135 (4), 1965–2005.

—, **Justine S Hastings, Thomas J Kane, and Douglas O Staiger**, "School choice, school quality, and postsecondary attainment," *American Economic Review*, 2014, 104 (3), 991–1013.

Department of the Army, *Military Occupational Classification and Structure* United States Department of the Army, Washington, DC: Headquarters, Department of the Army, December 2022. Certified current as of 20 December 2022.

Department of the Army, AR 601-210, "Army Regulation 601-210: Active and Reserve Components Enlistment Program," Technical Report, U.S. Department of the Army, Washington, DC 1995. Multiple versions issued between 1992 and 2014.

Dillon, Eleanor W., Lisa B. Kahn, Joanna Venator, and Michael Dalton, "Do Workforce Development Programs Bridge the Skills Gap," NBER Working Paper 34012, National Bureau of Economic Research July 2025.

Engel, Rozlyn C, Luke B Gallagher, and David S Lyle, "Military Deployments and Children's Academic Achievement: Evidence from Department of Defense Education Activity Schools," *Economics of Education Review*, 2010, 29 (1), 73–82.

Fadlon, Itzik, Frederik Plesner Lyngse, and Torben Heien Nielsen, "Causal effects of early career sorting on labor and marriage market choices: A foundation for gender disparities and norms," Technical Report, National Bureau of Economic Research 2020.

Forbes Media LLC, “America’s Top Private Companies,” <https://www.forbes.com/lists/largest-private-companies/> 2024. Accessed: 2025-10-21.

Gathmann, Christina and Uta Schönberg, “How general is human capital? A task-based approach,” *Journal of Labor Economics*, 2010, 28 (1), 1–49.

Greenberg, Kyle, Matthew Gudgeon, Adam Isen, Corbin Miller, and Richard Patterson, “Army Service in the All-Volunteer Era,” *The Quarterly Journal of Economics*, November 2022, 137 (4), 2363–2418.

—, **Melanie Wasserman, and E Anna Weber**, “The Effects of Gender Integration on Men: Evidence from the US Military,” Technical Report, National Bureau of Economic Research 2024.

Haeck, Catherine and Jean-William Laliberté, “Careers and Intergenerational Income Mobility,” *American Economic Journal: Applied Economics*, 2025, 17 (1), 431–458.

Heckman, James and Jeffrey Smith, “The Sensitivity of Experimental Impact Estimates: Evidence from the National JTPA Study,” in David Blanchflower and Richard Freeman, eds., *Youth Employment and Joblessness in Advanced Countries*, Chicago: University of Chicago Press for NBER, 2000, pp. 331–356.

Heckman, James J., John Eric Humphries, and Gregory Veramendi, “The Causal Effects of Education on Earnings, Health, and Smoking,” *Journal of Political Economy*, 2018, 126 (S1), S197–S246.

Hershbein, Brad and Lisa B Kahn, “Do recessions accelerate routine-biased technological change? Evidence from vacancy postings,” *American Economic Review*, 2018, 108 (7), 1737–1772.

Hjalmarsson, Randi and Matthew J Lindquist, “The Causal Effect of Military Conscription on Crime,” *The Economic Journal*, 2019, 129 (622), 2522–2562.

Huckfeldt, Christopher, “Understanding the Scarring Effect of Recessions,” *American Economic Review*, 2022, 112 (4), 1273–1301.

Humlum, Anders, Jakob Munch, and Mette Rasmussen, “What works for the unemployed? Evidence from quasi-random caseworker assignments,” 2023.

Imbens, Guido W. and Wilbert van der Klaauw, “Evaluating the Cost of Conscription in The Netherlands,” *Journal of Business & Economic Statistics*, 1995, 13 (2), 207–215.

Jacobson, Isabel G, Margaret AK Ryan, Tomoko I Hooper, Tyler C Smith, Paul J Amoroso, Edward J Boyko, Gary D Gackstetter, Timothy S Wells, and Nicole S Bell, “Alcohol Use and Alcohol-related Problems Before and After Military Combat Deployment,” *Journal of the American Medical Association*, 2008, 300 (6), 663–675.

Kahn, Lisa B, “The long-term labor market consequences of graduating from college in a bad economy,” *Labour economics*, 2010, 17 (2), 303–316.

Kambourov, Gueorgui and Iourii Manovskii, “Occupational mobility and wage inequality,” *The Review of Economic Studies*, 2009, 76 (2), 731–759.

— and —, “Occupational specificity of human capital,” *International Economic Review*, 2009, 50 (1), 63–115.

Katz, Lawrence F, Jonathan Roth, Richard Hendra, and Kelsey Schaberg, “Why do sectoral employment programs work? Lessons from WorkAdvance,” *Journal of Labor Economics*, 2022, 40 (S1), S249–S291.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, “Field of study, earnings, and self-selection,” *The Quarterly Journal of Economics*, 2016, 131 (3), 1057–1111.

Kline, Patrick and Christopher R Walters, “Evaluating public programs with close substitutes: The case of Head Start,” *The Quarterly Journal of Economics*, 2016, 131 (4), 1795–1848.

Lee, David S and Thomas Lemieux, “Regression discontinuity designs in economics,” *Journal of economic literature*, 2010, 48 (2), 281–355.

—, **Justin McCrary, Marcelo J Moreira, and Jack Porter**, “Valid t-ratio Inference for IV,” *American Economic Review*, 2022, 112 (10), 3260–3290.

Lyle, David S, “Using Military Deployments and Job Assignments to Estimate the Effect of Parental Absences and Household Relocations on Children’s Academic Achievement,” *Journal of Labor Economics*, 2006, 24 (2), 319–350.

McCrary, Justin, “Manipulation of the running variable in the regression discontinuity design: A density test,” *Journal of econometrics*, 2008, 142 (2), 698–714.

Mountjoy, Jack, “Community colleges and upward mobility,” *American Economic Review*, 2022, 112 (8), 2580–2630.

Neal, Derek, “The Link between Ability and Specialization: An Explanation for Observed Correlations between Wages and Mobility Rates,” *The Journal of Human Resources*, 1998, 33 (1), 173–200.

—, “The Complexity of Job Mobility among Young Men,” *Journal of Labor Economics*, April 1999, 17 (2), 237–261.

Negrusa, Sebastian, Brigitte Negrusa, and James Hosek, “Gone to War: Have Deployments Increased Divorces?,” *Journal of Population Economics*, 2014, 27 (2), 473–496.

Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz, “The short-and long-term career effects of graduating in a recession,” *American Economic Journal: Applied Economics*, 2012, 4 (1), 1–29.

Pavan, Ronni, “Career Choice and Wage Growth,” *Journal of Labor Economics*, July 2011, 29 (3), 549–587.

Reeves, Richard V., *Of Boys and Men: Why the Modern Male Is Struggling, Why It Matters, and What to Do about It*, Brookings Institution Press, 2022.

Rohlf, Chris, “Does Combat Exposure Make You a More Violent or Criminal Person? Evidence from the Vietnam Draft,” *Journal of Human resources*, 2010, 45 (2), 271–300.

Roys, Nicolas A and Christopher R Taber, “Skill prices, occupations, and changes in the wage structure for low skilled men,” Technical Report, National Bureau of Economic Research 2019.

Rubinstein, Yona and Yoram Weiss, “Post schooling wage growth: Investment, search and learning,” *Handbook of the Economics of Education*, 2006, 1, 1–67.

Sabia, Joseph J and William Skimmyhorn, “How Do Combat Deployments Affect Veterans’ Health and Labor Market Outcomes? Evidence From the U.S. Army,” Center for Health Economics and Policy Studies Working Paper, 2023.

Sanders, Carl and Christopher Taber, “Life-cycle wage growth and heterogeneous human capital,” *Ann. Rev. Econ.*, 2012, 4 (1), 399–425.

Sargan, J. D., “The Estimation of Economic Relationships Using Instrumental Variables,” *Econometrica*, 1958, 26 (3), 393–415.

Silver, David and Jonathan Zhang, “Impacts of Basic Income on Health and Economic Well-Being: Evidence from the VA’s Disability Compensation Program,” Working Paper 2022.

Staiger, Matthew, “The intergenerational transmission of employers and the earnings of young workers,” *Washington Center for Equitable Growth Working Paper*. <https://equitablegrowth.org/working-papers/the-intergenerational-transmission-of-employers-and-the-earnings-of-young-workers/> [Accessed 22 June 2022], 2021.

Topel, Robert H and Michael P Ward, “Job mobility and the careers of young men,” *The Quarterly Journal of Economics*, 1992, 107 (2), 439–479.

Traiberman, Sharon, “Occupations and import competition: Evidence from Denmark,” *American Economic Review*, 2019, 109 (12), 4260–4301.

U.S. Army, “Combat Medic Specialist (68W) — Career: Science & Medicine, Intensive Care,” <https://www.goarmy.com/careers-and-jobs/career-match/science-medicine/intensive-care/68w-combat-medic-specialist.html> n.d. Accessed: 2025-10-21.

— , “Wheeled Vehicle Mechanic (91B) — Career: Mechanics & Engineering, Test & Repair,” <https://www.goarmy.com/careers-and-jobs/career-match/mechanics-engineering/test-repair/91b-wheeled-vehicle-mechanic.html> n.d. Accessed: 2025-10-21.

U.S. Bureau of Labor Statistics, “Employment Situation of Veterans—2024,” News Release USDL-25-0377, U.S. Bureau of Labor Statistics, Washington, DC March 2025. Released March 20, 2025.

U.S. Department of Defense, “DoD Personnel, Workforce Reports & Publications,” <https://dwp.dmdc.osd.mil/dwp/app/dod-data-reports/workforce-reports> n.d. Accessed: 2025-10-21.

Wachter, Till Von, “The persistent effects of initial labor market conditions for young adults and their sources,” *Journal of Economic Perspectives*, 2020, 34 (4), 168–194.

Walters, Christopher R, “Inputs in the production of early childhood human capital: Evidence from Head Start,” *American Economic Journal: Applied Economics*, 2015, 7 (4), 76–102.

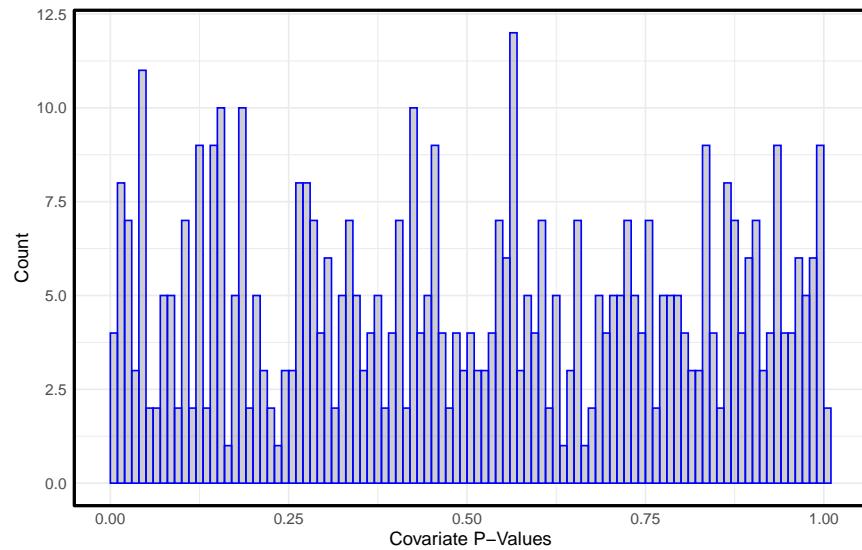
Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., Cambridge, MA: MIT Press, 2010.

Zimmerman, Seth D., “The Returns to College Admission for Academically Marginal Students,” *Journal of Labor Economics*, 2014, 32 (4), 711–754.

Zimmerman, Seth D., “Elite colleges and upward mobility to top jobs and top incomes,” *American Economic Review*, 2019, 109 (1), 1–47.

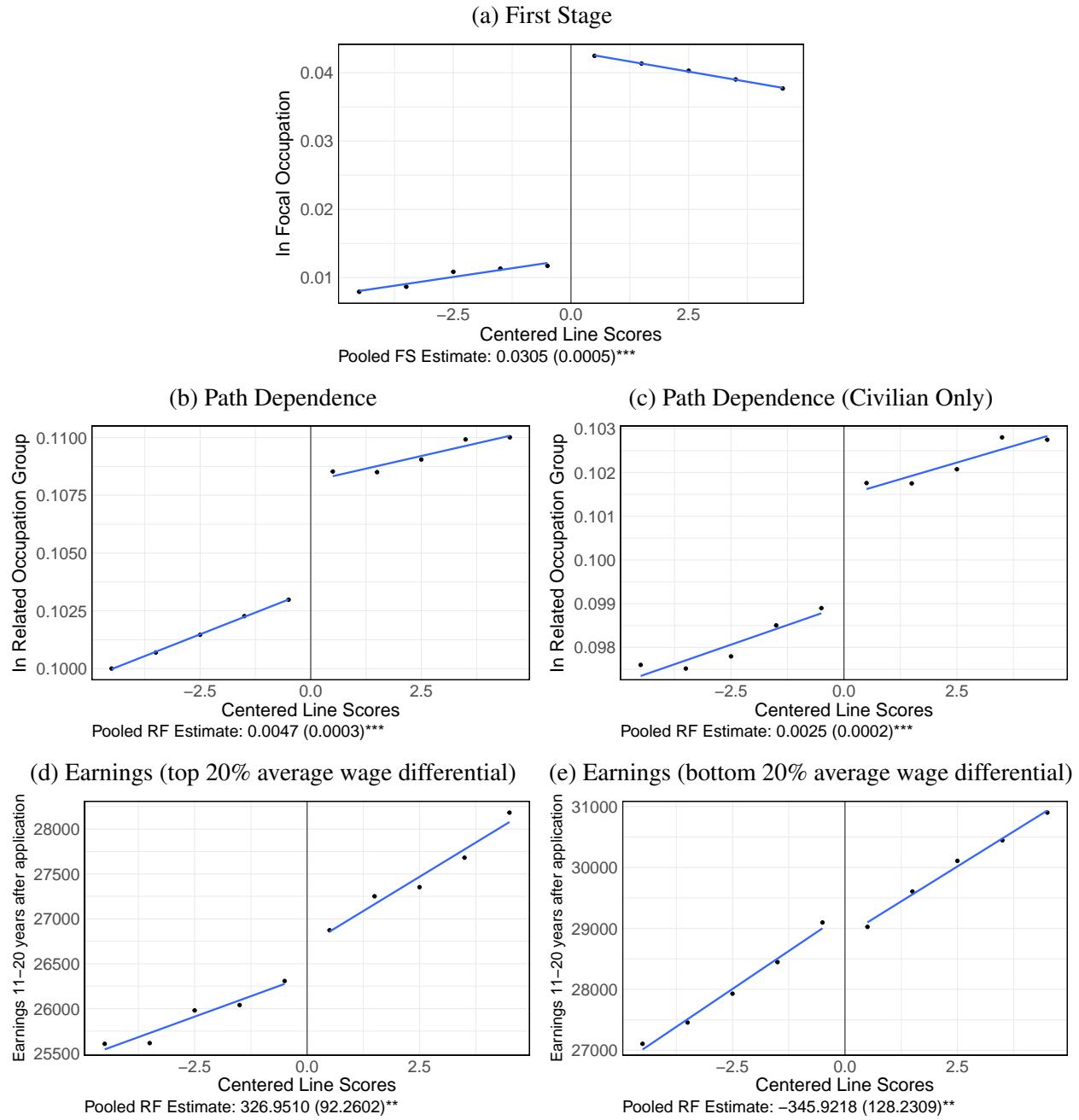
Figures

Figure 1: Covariate Balance



Notes: This Figure plots the $35 \times 14 = 490$ p-values for the reduced-form effects of crossing each of the 35 cutoffs on each of the 14 baseline covariates. The underlying point estimates and standard errors are reported in Tables F.5 and F.6. Tables F.5 and F.6 also report the p-value from a test of balance conducted jointly across all covariates and cutoffs. We fail to reject that the baseline covariates are balanced across the thresholds ($P = 0.173$).

Figure 2: Pooled First Stage and Reduced Form Plots



Notes: This figure plots discontinuities for various key outcomes, including first stage take-up of the relevant military occupations (panel A); career continuation 11-20 years after ASVAB (panels B and C); and wages (panels D and E) pooled across all cutoffs with at least five support points in the running variable before and after the cutoff. For panel A, the outcome variable used to construct the Y-axis is defined to equal 1 if the applicant enlisted in an occupation they became eligible for as a result of clearing the cutoff. For panel B, the outcome is the share of years the applicant worked in an occupation unlocked as a result of clearing the relevant cutoff. Panel C is identical to panel B, except that it only counts years employed in the relevant occupation for a non-Army employer. For Panel D, the outcome is earnings, but the sample of cutoffs is further restricted to the 20% of cutoffs that create the largest implied average occupational wage differential between the careers applicants are pushed into or pulled out of in the long run. Hence, these are cutoffs where we would expect to see positive impacts on earnings based on the implied career change. Panel E is identical to panel D, except restricted to the bottom 20% of cutoffs in terms of implied average occupational wage differential. A complete set of figures corresponding to the individual first-stage and reduced form effects for all 35 cutoffs, along with associated point estimates and standard errors, are contained in Appendix Figure F.5.

Figure 3: Causal Effects on Having a Related Occupation 11-20 Years Later

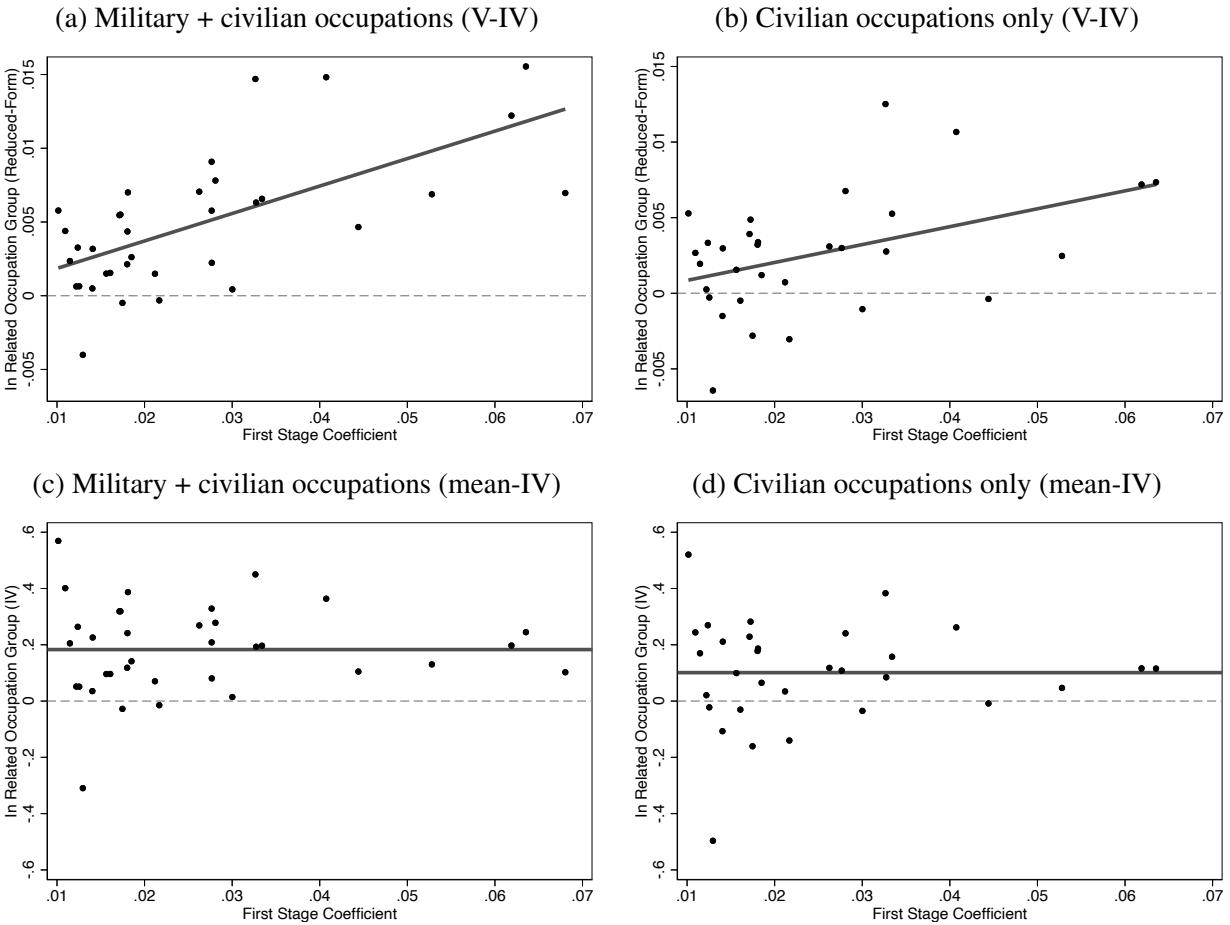
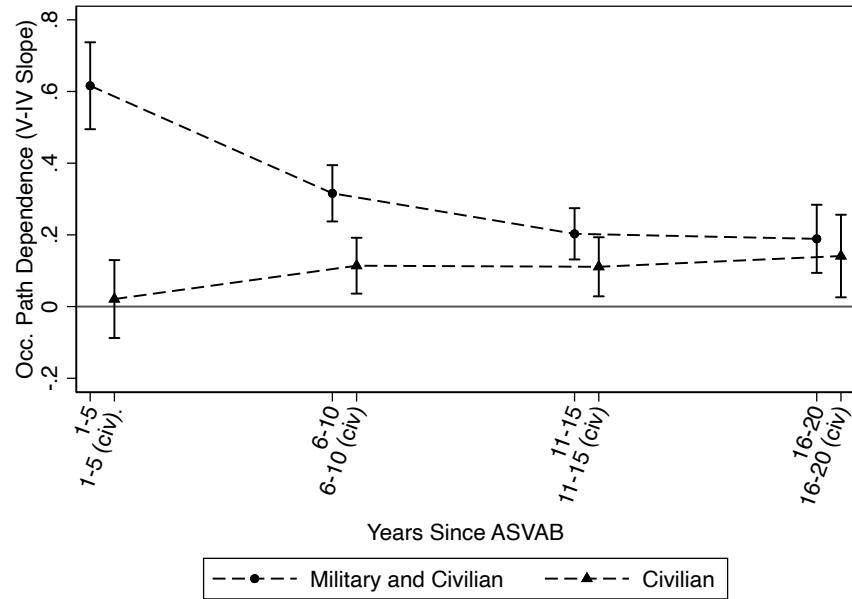
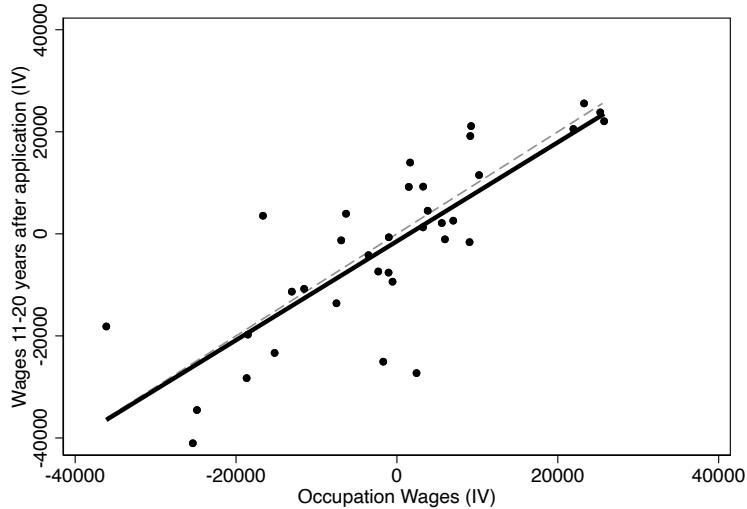


Figure 4: The Dynamics of Path Dependence



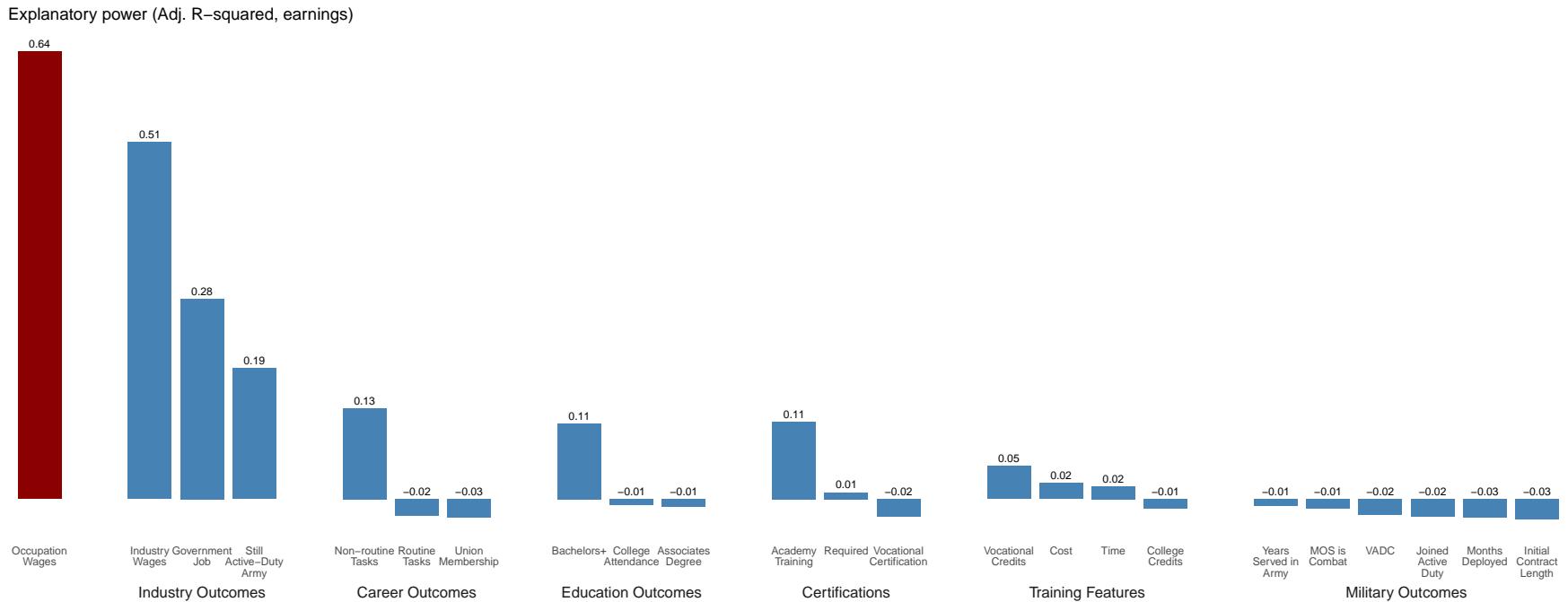
Notes: This figure plots our preferred measure of occupational path dependence (V-IV slope) for different time horizons: 1-5 years, 6-10 years, 11-15 years, and 16-20 years post-ASVAB. The circles represent the path dependence effect including both Army and non-Army employers. The triangles represent estimates restricted to non-Army employers only. 95% confidence intervals, based on standard errors constructed using the delta method from the full variance-covariance matrix of reduced form and first stage estimates in each time horizon, are depicted.

Figure 5: Realized Earnings and Long-run Shifts in Average Occupational Earnings



Notes: This figure plots causal effects on realized earnings against causal effects on occupation wage premia. Each point is a cutoff. The Y-axis denotes the year 11-20 IV effect of the military occupations unlocked by that cutoff on average own wages. The X-axis denotes the year 11-20 IV effect of the military occupations unlocked by that cutoff on leave-out average occupation wages. Thus, the X-axis directly measure which military careers tend to move individuals out of low-paying occupations and into high-paying occupations (as well as the converse) over the long-run. The occupation variable used to construct the causal effects on the X-axis reflect full sample, leave-out average wages based on recorded three-digit SOC codes. We find similar results using a version of this variable constructed from a 50% random sample pulled from the universe of tax records for those aged 31-44 (see Appendix figure F.14). The line of best fit is shown in solid black. The dashed gray line corresponds to the 45-degree line. Table 2 provides the slope, standard error of the slope, and adjusted R^2 of the best fit line. As robustness, Appendix Figure F.13 plots an identical figure, but where we replace the X- and Y-axes with the corresponding reduced form effects. Appendix Table F.7 also gives corresponding estimates of the slope directly estimated from a 2SLS regression on the micro-data. See Appendix D for more discussion on the connection between this plot, the reduced-form version, and the micro-data 2SLS regression.

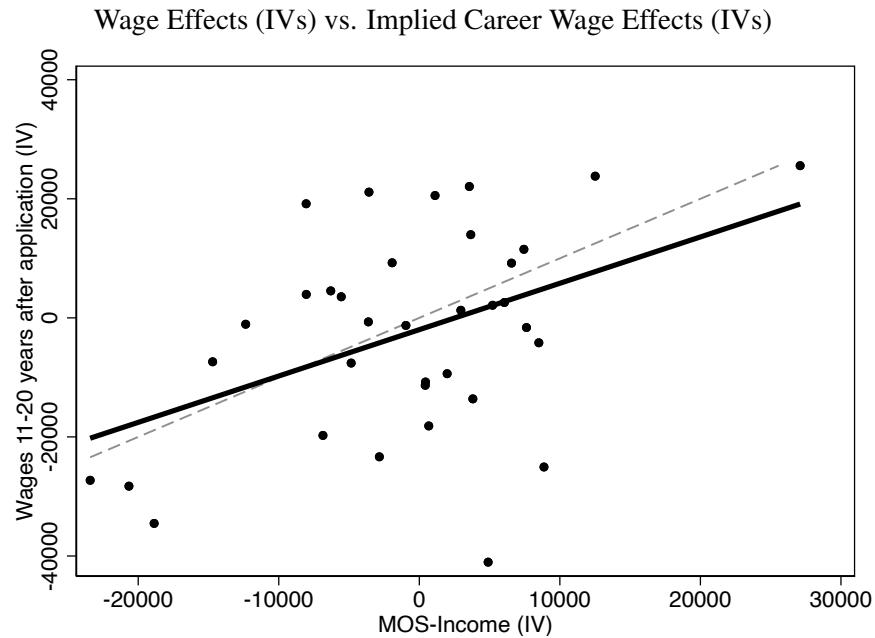
Figure 6: What Explains the Heterogeneous Wage Effects?



44

Notes: This figure displays adjusted R^2 values from univariate regressions that predict the cutoff-specific IV effects on year 11-20 average earnings with the corresponding cutoff specific IV effect on various other outcomes. For example, the bar labeled “Occupation Wages” replicates the adjusted R^2 value from the “Wages v. Occ Wages” column in panel (c) of table 2; the adjusted R^2 values of the remaining bars are generated identically, except that the causal effect on “occupation-earnings” which appears on the X-axis of figure 5 is replaced with the causal effect on an alternative outcome. See appendix figures F.15 – F.17 for the individual scatter-plots underlying these R^2 values. “Industry Wages” is the (leave-out) mean earnings of others working in the same six-digit industry as reported by the employer on form W2. “Government Job” is the share of time, 11-20 years after ASVAB, spent working for a government, non-military employer. “Still Active-Duty Army” is the share of time, 11-20 years after the ASVAB, that an applicant was still considered on active duty according to the Army pay data. “Non-routine Task” intensity (a proxy for high-skill work) and “Routine Task” intensity (a proxy for low-skill work) are constructed from the ONet database as described in Autor et al. (2003). “Union Membership” is taken from the CPS and measures the share of individuals with the same occupation who reported belonging to a union. Each of these is based on recorded SOC codes over years 11-20 after the ASVAB. “Bachelors+” is an indicator constructed from the NSC data that measures attainment of any bachelors or advanced degree by 2020. “College Attendance” is identical, except that it measures enrollment in any post-secondary institution (rather than attainment). “Associates Degree” is identical, except that it measures only associates degrees. “Academy Training”, “Required”, and “Vocational Certification” are indices built using data from the CPS, BLS, and state level regulations and attached to each individual’s initial MOS based on the corresponding primary SOC-code in our crosswalk. “Academy Training” measures the share of individuals working in a 3-digit SOC code that required training in an Academy (e.g. a police or fire academy, defined as training lasting more than 3 months at a specialized institution). “Required” measures the share of individuals working in a 3-digit SOC code that require any form of certificate for employment. “Vocational Certification” measures the share of individuals working in a 3-digit SOC code that required a vocational certificate for employment. “Vocational Credits” is built from internal Army data and measures the number of vocational credits the military recommends be awarded to a soldier after completion of the occupation specific training for their MOS. “Cost” is built from internal Army data and measures the estimated cost of the Army provided occupational training for each MOS. “Time” is built from internal Army data and measures the number of weeks soldiers spend in occupation specific training for their MOS. “College Credits” is built from internal Army data and measures the number of college credits the Army recommends be awarded to a soldier after completion of the occupation specific training for their MOS. “Years Served in Army” measures the number of years the soldiers spent in the active-duty Army according to Army pay data. “MOS is Combat” is an indicator that takes a value of one if the applicant enlisted in an Army combat occupation. “VADC” is an indicator that takes a value of 1 if the individual received any Veterans Affairs Disability Compensation as of 2019. “Joined Active Duty” is an indicator that takes a value of one if the individual ever enlisted in the active-duty Army. “Months Deployed,” measures the number of months the individual received hazardous duty pay within the first four years after the ASVAB, which is only available to soldiers deployed to dangerous war-zones. “Initial Contract Length” measures the number of years the applicant was initially contractually obligated to serve. See appendix B.2 for more detailed discussion of each of these variables and how they are constructed.

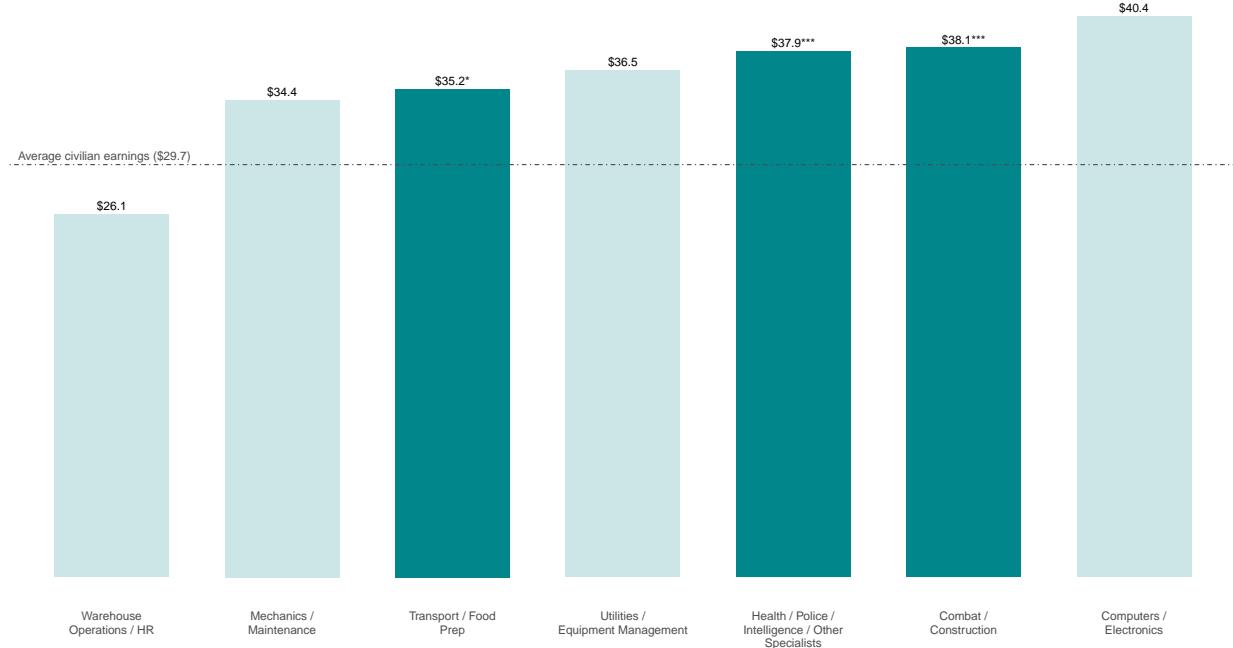
Figure 7: Realized Earnings and Short-run Shifts in Career Path



Notes: This figure plots causal effects on realized earnings, 11-20 years after the ASVAB, against causal effects on the expected change in future wages implied by the corresponding short-run impact on military career. In this figure, each point is a cutoff. The Y-axis denotes the year 11-20 IV effect of that cutoff on average own wages. The X-axis denotes the IV effect on the expected future earnings of the military occupation (as measured by average, long-run, leave-out earnings of all civilian SOC codes linked to an MOS unlocked by the cutoff). Thus, the X-axis quantifies whether the military occupations at that cutoff moved applicants onto a higher- or lower-expected wage career path. The line of best fit is shown in solid black. The dashed gray line corresponds to the 45-degree line. The slope coefficient is 0.779 (se = 0.335), the constant is insignificantly different from 0, and the Adjusted R^2 is 0.173. The standard error is calculated using the delta method from the full covariance-variance matrix of the two IV estimates at each cutoff.

Figure 8: Implied Long-run Earnings from joint 2SLS model

Average long-run annual earnings (USD, thousands)



Notes: This figure displays estimates from the model $Y_i = \beta D_i + f(\mathbf{X}_i) + \epsilon_i$ where Y_i is earnings; $D_i = \{D_{ig}\}$ is a vector of 8 treatment indicators, one for each of the 8 occupation groups with “remaining a civilian” constituting the (9th) omitted category; $\mathbf{X}_i = \{X_{il}\}$ is a vector of the various line scores; and f maps line scores into the series of piecewise linear and quadratic polynomials around the various RD thresholds as implied by our baseline, individual RD models. We instrument for the treatment variables with 8 first stage equations of the form: $D_{ig} = \Pi_g \mathbf{Z}_i + h_g(\mathbf{X}_i) + \eta_{ig}$ where D_{ig} takes a value of 1 if applicant i enlists in any occupation contained in group g ; $\mathbf{Z}_i = \{Z_{ic}\}$ is a vector of indicator variables that take a value of one if individual i crosses the line score threshold c ; and h_g maps line scores into the series of piecewise linear and quadratic polynomials around the various RD thresholds as implied by our baseline, individual RD models. In support of the additional restrictions on treatment effect heterogeneity necessary to identify the joint model (over and above the impact of the individual RDs), we note that it passes an over-ID test ($P = 0.144$) that jointly tests for exclusion restriction and functional form violations. Each bar in this figure is the estimated causal effect from our multiple endogenous variable model (β) with the civilian mean added back-in for interpretability – thus the height of each bar represents the long run earnings of each military occupation group, with causal effects encoded in differences between the bar height and the horizontal line representing the civilian mean. We omit “Other Army Job” from the figure. All estimates can be found in Appendix Table F.13. Bars with treatment effects (relative to civilian) that are statistically different from zero at the 10% level are in bold, with stars denoting conventional 1, 5, and 10% levels.

Tables

Table 1: Summary Statistics

	Full Sample	Stacked Sample	Complier Means
<i>Panel (a): Demographic variables, measured prior to ASVAB</i>			
Age	20.7	20.4	20.1
Male	78%	73%	74%
White	59%	53%	45%
Black	22%	27%	35%
Hispanic	11%	13%	15%
In High School	25%	27%	14%
No High School Diploma	14%	16%	14%
High School Diploma	53%	53%	69%
Some College +	7%	5%	3%
AFQT Score	51.3	44.3	48.4
<i>Panel (b): Outcome variables, measured 11-20 years after ASVAB</i>			
Income	\$32,342	\$30,042	\$35,225
Occupation Wages	\$32,636	\$31,028	\$35,036
Occupational Persistence	–	0.102	0.014
Occupational Persistence (Civilian Only)	–	0.094	0.016
Observations	2,164,404	10,344,845	10,344,845

Notes: This table reports means for demographic variables as measured at baseline (i.e. prior to the ASVAB) and key outcome variables, as measured 11-20 years after the ASVAB. We report means for the Army applicant data (“Full Sample”), and then the stacked subsample used in our main analysis, which includes all observations included across all 35 line-score cutoffs. The final column reports untreated complier means for the stacked sample, estimated by a 2SLS regression on the outcome $-Y(1 - D)$ at each cutoff and then averaged across cutoffs. Mutually exclusive education categories are defined as follows: “In High School” equals 1 for those still enrolled in high school at the time of application; “No High School Diploma” takes a value of 1 for those no longer in high school, but who have a GED, credential near completion, or less than high school completion; “High School Diploma” takes a value of one for those who have earned a high school diploma but not attended any college; and “Some College+” takes a value of one for anyone who has attended at least one semester of college. “Income” denotes average earned income, 11-20 years after the ASVAB, as compiled from third party reported information returns (as described in Appendix B), and is reported in 2010 real dollars. “Occupation Wages” is constructed by taking an applicant’s reported SOC code, and then assigning them the in-sample, leave-one-out, average three-digit SOC code earnings of others with the same SOC code in that year. We then average over years 11-20 after the ASVAB, identically to our construction of the income variable. “Occupational Persistence” is the share of time (11-20 years after the ASVAB) for which an individual is observed in an occupation (three-digit SOC) that is linked by our crosswalk to the corresponding Army occupations unlocked at a line score cutoff. Note that this variable is defined in relation to specific cutoffs; hence, it is only defined for the column corresponding to the stacked sample and does not have an entry in the first column. “Occupational Persistence (Civilian Only)” is defined identically, except we only count an individual as being in a linked occupation if they are employed at a non-Army employer.

Table 2: Summary of Key Results

Panel (a): Occupational Path Dependence			
	Military & Civ.	Civilian Only	
<i>V-IV:</i>			
Slope Coef.	0.186*** (0.034)	0.118*** (0.040)	
Slope Coef. (no constant)	0.186*** (0.020)	0.108*** (0.022)	
<i>Mean-IV:</i>			
Average IV Effect	0.183*** (0.029)	0.101*** (0.030)	
<i>Joint tests:</i>			
$P(F \geq f \mid H_0 : \beta_c = 0 \forall c)$	<0.001	0.001	
$P(F \geq f \mid H_0 : \beta_c = \beta \forall c)$	0.341	0.301	
Panel (b): Heterogeneous Wage Effects			
	Wages	Work	log(Wages)
$P(F \geq f \mid H_0 : \beta_c = 0 \forall c)$	<0.001	0.023	0.005
$P(F \geq f \mid H_0 : \beta_c = \beta \forall c)$	<0.001	0.018	0.004
Panel (c): Earnings versus Occupation Wages			
Slope Coef.	0.969*** (0.208)		
Adj. R^2	0.635		

Notes: This table summarizes the key results and relationships described in Section 6. Panel (a) summarizes our results on path dependence. It first reports the V-IV slopes and heteroskedasticity robust standard errors (constructed using the delta method and using the full covariance-variance matrix of reduced form and first stage estimates) from Figure 3 panels (a) and (b). Below this, it reports the same objects but for a regression without a constant (as is implied by the IV exclusion restriction), which yields efficiency gains. It then reports the average of the IV effects and associated standard error (generated from a stack of all 35 IV regressions) from panels (c) and (d) of Figure 3, as well as joint tests that these are jointly equal to zero and jointly equal to each other. In panel (b), this Table reports joint tests for whether or not the 35 IV estimates for wages, working, and $\ln(\text{wages})$ are jointly zero or jointly equal. Panel (c) reports the post-estimation OLS slopes, heteroskedasticity-robust standard errors (constructed using the delta method and using the full covariance-variance matrix of IV estimates), and adjusted R^2 between wages and occupation wages depicted in Figure 5.

Table 3: Robustness of Key Results

Panel (a): Occupational Persistence										
	(1) Baseline	(2) Baseline	(3) Quadratic	(4) Quadratic	(5) BW \leq 5	(6) BW \leq 5	(7) Linear, BW \leq 5	(8) Linear, BW \leq 5	(9) Controls	(10) Controls
	Military & Civ.	Civilian Only	Military & Civ.	Civilian Only	Military & Civ.	Civilian Only	Military & Civ.	Civilian Only	Military & Civ.	Civilian Only
V-IV:										
Slope Coef.	0.186*** (0.034)	0.118*** (0.040)	0.180*** (0.039)	0.116*** (0.046)	0.224*** (0.045)	0.126*** (0.049)	0.150*** (0.029)	0.084*** (0.034)	0.179*** (0.034)	0.111*** (0.039)
Slope Coef. (no constant)	0.186*** (0.020)	0.108*** (0.022)	0.192*** (0.022)	0.113*** (0.025)	0.187*** (0.025)	0.091*** (0.027)	0.165*** (0.017)	0.090*** (0.019)	0.187*** (0.020)	0.112*** (0.022)
IV Persistence Effects:										
Average IV Effect	0.183*** (0.029)	0.101*** (0.030)	0.197*** (0.034)	0.110*** (0.035)	0.158*** (0.033)	0.068*** (0.034)	0.170*** (0.026)	0.091*** (0.027)	0.191*** (0.028)	0.126*** (0.033)
$P(F \geq f H_0 : \beta_c = 0 \forall c)$	< 0.001	0.001	< 0.001	0.018	< 0.001	0.034	< 0.001	0.012	< 0.001	0.001
$P(F \geq f H_0 : \beta_c = \beta \forall c)$	0.341	0.301	0.393	0.625	0.657	0.401	0.652	0.529	0.528	0.438
Panel (b): IV Wage Effects										
	(1) Baseline	(2) Quadratic	(3) BW \leq 5	(4) Linear, BW \leq 5	(5) Controls					
	Wages	Wages	Wages	Wages	Wages					
$P(F \geq f H_0 : \beta_c = 0 \forall c)$	< 0.001	0.001	0.005	0.006	< 0.001					
$P(F \geq f H_0 : \beta_c = \beta \forall c)$	< 0.001	< 0.001	0.004	0.007	< 0.001					
Slope Coef. on Wages vs. Occ Wages	0.969*** (0.208)	0.951*** (0.275)	0.936*** (0.238)	0.890*** (0.207)	0.976*** (0.228)					
Adj. R^2	0.635	0.583	0.573	0.622	0.658					

Notes: This table reports robustness checks on the main results in Table 2. In panel (a), the first two columns simply restate our baseline results on path dependence (with and without Army retention) for comparison purposes. Columns (3) and (4) report results where we modify the baseline model by using all quadratic polynomials to estimate the underlying regression discontinuity at each cutoff. Columns (5) and (6) report results where we modify the baseline model by using a maximum bandwidth of 5 to estimate the underlying regression discontinuity at each cutoff. Columns (7) and (8) report results where we modify the baseline model to use all linear polynomials and a maximum bandwidth of 5 to estimate the underlying regression discontinuity at each cutoff. Columns (9) and (10) report results where we modify the baseline model to include stack-specific controls for all the time of application covariates in Table F.5. In panel (b) we report the p-values relating to wage IV effects, as well as the “IV-on-IV” slope coefficient for wages versus average occupation-wages for the same set of alternative specifications reported in panel (a). Standard errors are robust to heteroskedasticity and constructed with the delta method using the corresponding full variance-covariance matrix.

Table 4: Stability of Correlation between Wages and Occupation Wages

	Estimate (SE)	Adj. r-squared
<i>Panel A: Baseline</i>		
No controls	0.97*** (0.208)	0.635
<i>Panel B: Extensive Margin Control</i>		
Working	0.86*** (0.234)	0.627
<i>Panel C: Industry Outcome Controls</i>		
Industry Wages	0.81** (0.275)	0.631
Government Job	0.84*** (0.149)	0.673
Still Active-Duty Army	0.92*** (0.183)	0.628
<i>Panel D: Career Outcome Controls</i>		
Occupation Non-routine Tasks	0.92*** (0.152)	0.632
Occupation Routine Tasks	0.98*** (0.128)	0.644
Occupation Union Membership	1.04*** (0.130)	0.659
<i>Panel E: Education Outcome Controls</i>		
Bachelors+	0.92*** (0.123)	0.665
College Attendance	0.98*** (0.137)	0.626
Associates Degree	0.96*** (0.137)	0.633
<i>Panel F: Certification Requirement Controls</i>		
Academy Training	0.93*** (0.161)	0.631
Any Certification Required	0.96*** (0.143)	0.625
Vocational Certification	0.97*** (0.138)	0.626
<i>Panel G: Training Feature Controls</i>		
Vocational Credits	0.96*** (0.141)	0.624
Total Cost	0.95*** (0.143)	0.631
Time (Weeks)	0.95*** (0.137)	0.635
College Credits	0.97*** (0.139)	0.624
<i>Panel H: Military Outcome Controls</i>		
Years Served in Army	0.97*** (0.142)	0.624
MOS is Combat	0.98*** (0.146)	0.624
VADC	0.97*** (0.127)	0.632
Joined Active Duty	0.97*** (0.134)	0.629
Months Deployed	0.98*** (0.133)	0.644
Initial Contract Length	0.97*** (0.134)	0.629

Notes: This table presents results from bivariate regression models of the form $Y_c = \alpha + \beta Z_c + \pi X_c + e_c$, where Y_c is the 2SLS estimate (model 2) of the causal effect on realized earnings generated by the occupations unlocked at cutoff c ; Z_c is the corresponding estimate of the causal effect on (leave-one-out) average occupation earnings; and X_c is a control variable that corresponds to the estimated causal impact on another outcome (e.g. leave-one-out average industry wages). Panel (A) replicates the unconditional results from table 2 as a baseline. Panels B-H give results conditional on corresponding changes in the extensive margin (B); industry features (C); career outcomes (D); educational outcomes (E); certification requirements (F); training features (G); and military specific outcomes (H). See appendix B.2 for precise variable definitions. Standard errors are robust and calculated directly from equation 4.

Online Appendix

A Details of Constructing Line Scores for Army Occupations

Our analysis relies on enlistment records that allow us to observe a new enlistee's first Military Occupational Specialty (MOS) and the corresponding "line scores," that determine military occupation eligibility.⁴⁶

We recover line scores for each applicant from their "standard scores", which measure how the applicant performed on each of the various ASVAB sub-tests. The sub-tests that comprise the ASVAB are shown in Table F.2. We observe individual-level "standard scores" (or the corresponding raw sub-test scores from which these are derived) for all applicants.⁴⁷

Line scores are then constructed based on different combinations of the standard scores. Table F.3 describes the line scores. Prior to 2002, these line scores were constructed by summing select standard scores and then using a conversion table to convert this sum into the final line score. For example, the General Maintenance (GM) line score was the sum of the Mathematics Knowledge (MK), Electronic Information (EI), Auto and Shop (AS), and General Science (GS) standard scores, converted via a table to the final GM line score. The final line scores have a mean of 100 and standard deviation of 50. After 2004, the line scores are weighted averages of all standard scores. For example, the GM score is

$$\approx 0.24 \cdot GS + 0.46 \cdot AR + 0.45 \cdot MK + 0.29 \cdot MC + 0.30 \cdot EI + 0.50 \cdot AS + 0.22 \cdot (WK + PC) - 23.36.$$

One consequence of the way line scores are constructed and used is that post-2004 line scores are effectively continuous, but prior to 2002 scores are discrete and exhibit a sawtooth pattern (see e.g. figures F.6–F.7).⁴⁸ This occurs because conversion tables map integer sums of scaled scores into integer line score values in such a way that sometimes a single sum of scaled scores corresponds to a single line score value, but other times multiple sums of scaled scores correspond to the same line score value.

Line scores changed over the course of our sample frame. Both current and historical line scores for each 'current-year' MOS are outlined in DA-PAM 611-21 Chapter 10-3-2. Since some MOS designation codes change over time, we map MOS at each point in time to their 'current-year' MOS using a crosswalk. There are roughly 150 distinct entry MOS in both periods. Across all 10 line scores there are roughly 40 cutoffs per-period, but as explained in Section 5, we will examine a subset of these that produce sizable enough first stages.⁴⁹

⁴⁶The Armed Forces Qualification Test (AFQT) determines whether applicants are able to enlist in the Army at all, but otherwise does not determine eligibility for any particular job in the Army. In section 6.4, our joint-2SLS model also leverages the two AFQT cutoffs, at the 31st and 50th percentiles of a nationally-normed distribution, that generate discontinuities in an applicant's eligibility to enlist. The Army rarely accepts applicants with AFQT scores below 31, sometimes requires GED recipients to achieve a score of 50 or higher, and often requires applicants to score 50 or higher to receive enlistment bonuses. Following Greenberg et al. (2022), whenever we use AFQT as a running variable, we use applicants' *first* AFQT on file from their initial visit to the MEPS.

⁴⁷For the 1992-2001 period we observe raw test scores. We use time-varying and test-version specific conversion tables to translate raw scores into standard scores for each individual in our data. For the 2004-2014 period we observe standard scores directly in the data.

⁴⁸Given that the underlying standard scores prior to 2002 are also generated from conversion tables and are discrete, there is no way to construct a more continuous line score. The AFQT score percentiles in all periods and the GT score post-2004, which is based on only the sum of VE and AR, are also discrete.

⁴⁹In practice many cutoffs either apply to very uncommon or small, singleton jobs, or apply to jobs for which another cutoff tends to be far more binding. Cutoffs can also be adjacent (e.g. MM 87/88), which in practice cannot be separately identified, reducing the effective number of cutoffs. In such cases, we count this as one cutoff (treating all jobs affected by either MM 87 or 88) and treat it as occurring at 87 with a doughnut-hole between 87-88.

B Data and Outcome Variable Details

Here we describe how we construct each of our key and auxiliary outcome variables from the underlying source data.

B.1 Primary Outcomes

Path Dependence in an identical or closely related occupation, 11-20 years after application. Our preferred outcome for measuring changes in career trajectory is defined as the share of years (e.g. 11-20 years after the ASVAB) for which individual i is observed in an occupation (three-digit SOC) that is closely related to the corresponding Army occupations unlocked at a line score cutoff. In the tax data, for 2005-2022, we are able to measure three-digit Standard Occupational Classification (SOC) codes for a subset of taxpayers. This comes from the fact that taxpayers can report their occupation at the bottom of their Form 1040 each year. We use a crosswalk created internally at the Treasury to merge these string variables to 2010 three-digit SOC codes, with most string variables successfully linked to a SOC code when the individual is employed. We are able to merge approximately 50% of Army applicant-year observations to a SOC between 2005-2022, with the denominator including nonfilers and those not employed. Conditional on having positive earnings, tax filers can be missing an occupation due to a variety of reasons including a failure to report anything in the relevant field on the 1040. Importantly, we note that missingness for this variable (i.e having earned income but no SOC) appears balanced across cutoffs when we predict it using our preferred regression discontinuity model. Specifically, a joint F-test of the null hypothesis that missing a SOC is balanced across all cutoffs yields a p-value of 0.725. Thus, when we analyze impacts on occupational trajectories, we keep Army applicants without a linked SOC who have no earnings in our sample (as zeros) but exclude (as missing) Army applicants who have earned income but no linked SOC.

In order to measure persisting in an identical or closely related occupation, we have to map Army Occupations (MOS) to SOC codes. To accomplish this, we create an MOS to three-digit SOC crosswalk. Given the fact that several civilian jobs can reasonably be linked to an MOS, we crosswalk each MOS to a “primary” SOC code and a “secondary” SOC code. To capture notions of career progression, we also maintained any relevant supervisory SOC codes.⁵⁰ In all, this results in each MOS having a mapping to a primary and secondary SOC code as well as any associated supervisory SOC codes. For various combat occupations like infantry or artillery, for which civilian counterparts are less obvious, we link the primary SOC code to the broad military category (55-0) and, for their secondary SOC code, we link them to law enforcement (SOC 33-3). Appendix table F.14 contains the complete crosswalk.

Since each line score cutoff can determine eligibility for multiple military occupations, we construct cutoff-specific, occupation-group outcomes. The natural way to do this is to create the outcome based on whether the individual is in any of the primary, secondary, or supervisory occupations corresponding to the full set of military occupations “unlocked” by that cutoff. In practice and in the interest of power, since several military occupations at each cutoff are negligible in size or not substantially impacted by the cutoff, we do this for the military occupations with sufficiently large MOS-specific first stages. Specifically, we base our outcomes on the MOS at each cutoff that either has an MOS-specific, first stage coefficient greater than or equal to 0.003 or (to capture at least some MOS at cutoffs with smaller overall first stages) whose MOS-specific, first stage coefficient is $\geq 10\%$ of the total first stage across all jobs at the cutoff. If an applicant’s 3-digit SOC falls into any of the codes associated with these military occupations at a given

⁵⁰For example, the “49” category of SOC codes contains installation, maintenance, and repair occupations. Within that broad category, there are numerous, more specific occupations such as 49-2 (“Electrical and Electronic Equipment Mechanics, Installers, and Repairers”) and 49-3 (“Vehicle and Mobile Equipment Mechanics, Installers, and Repairers”). One of those more specific three-digit SOC occupations is 49-1 which denotes “Supervisors of installation, maintenance, and repair workers.” Thus, if we link an Army occupation to any three-digit SOC code in the 49 broad category, we additionally retain the supervisory category in our crosswalk in order to capture natural career progression.

cutoff, then the variable is coded as 1 (denoting career continuation), and otherwise it is 0. However, our results are robust to alternative ways of linking MOS to SOC-codes (see Appendix C).

As a concrete example, obtaining a score of 85 on line score OF is necessary for applicants to be eligible for Army occupations 92G (Culinary Specialist) and 88M (Motor Transport Operator).⁵¹ These occupations are naturally linked to the civilian occupations of Cooks and Food Preparation Workers (SOC 35-2), Food Service Managers (a subset of SOC 11-9), Motor Vehicle Operators (SOC 53-3) and Material Moving Workers (SOC 53-7). In this case, the outcome variable would be equal to the share of years an individual is observed in those occupations (SOC codes 35-2, 11-9, 53-3 and 53-7), or their supervisory counterparts (in this case, 35-1, 53-1), or, if still in the Army, in an MOS linked to these same SOC codes, 11-20 years after they date that they took the ASVAB test.

We construct this measure for each year since application and then take the average for each applicant across all 11-20 years after application. Because for some cohorts we cannot observe SOCs in all years since application due to the data coverage being limited between 2005 and 2022, we weight all regressions by the number of years between 11-20 for which we have tax data and the individual either had a SOC code or had no earnings.

Civilian-Only Path Dependence, 11-20 years after application. For our baseline measure of occupational path dependence, if you are still in the Army in any given year, we take the MOS you have in that year and crosswalk it to its primary SOC, so that being in the same or a related occupation takes a value of 1 if you are in the same MOS as you started in (or switched to an MOS that is still linked to a SOC that counts at that cutoff). However, we are also interested in persistence that is not driven by within employer retention. Hence, we also explore non-Army path dependence. To do so, we simply replace the outcome variable = 0 for everyone who is still in the Army in the given calendar year, so that values of 1 can only come from years in which you are in a related non-Army SOC. We expect this to be a lower bound on the skills / information / credentialing channels, since these channels undoubtedly increase the likelihood of staying with the same employer as well – it is just not possible for us to completely separate them out using our experimental variation alone.

Alternative Path Dependence Variables: We also present robustness to several other sensible ways of defining this variable, including using only the primary plus supervisory codes as opposed to the secondary codes (Appendix Figure F.8), not counting path dependence at any military employer (Appendix Figure F.11), cross-walking combat specific military occupations like the infantry to SOC codes related to law-enforcement (Appendix Figure F.9), and an alternative measure built by feeding MOS / SOC descriptions from official documentation into a large language model (Appendix Figure F.10). Details for the LLM-based approach are in Appendix C.

Earnings, 11-20 years after application. Earned income is compiled from third party reported information returns available between 1999 and 2024. The measure is primarily derived from the wage income subject to Medicare tax reported by employers on Form W-2. We also add in contractor income reported by contractor businesses on Form 1099-MISC (and beginning in 2020, Form 1099-NEC). Our earned income measure is in 2010 real dollars, winsorized at the 99% level. Those with no earned income are set to zero. Note that self-employed income reported on a 1099 is the only source of self-employed income included in this measure. Here we are relying strictly on employer-provided income reports (including the 1099) so as not to restrict our sample only to individuals who have filed their own individual tax returns.

To account for the fact that those in service also receive compensation that would typically be included as taxable wages in the civilian sector but are exempt from taxation for servicemembers, we follow [Greenberg](#)

⁵¹In practice, the cutoff is also attached to several additional Army MOS; however, we ignore them here for clarity of exposition.

et al. (2022) and use the Army personnel records to add these additional forms of compensation into our measure of income for individuals who are on active duty.⁵²

We construct earnings in each calendar year since application (as determined by the date of the ASVAB test, which is year 0 after application) for each applicant from 11 years after application up to 20 years after application and then average these numbers. Because for some cohorts we cannot observe income in all years since the ASVAB due to the outcome data being limited to between 1999 and 2024 (for example, for the 2010 applicant cohort we observe earnings 11-14 years out, as 15 years out would correspond to 2025), we weight all regressions by the number of years between 11-20 for which we have tax data for that applicant.

Average Occupation Wages, 11-20 years after application. To measure average occupation wages, we take an applicant's reported SOC in each year since application (SOC codes are observed from 2005-2022) and assign them the in-sample, leave-one-out, average three-digit SOC earnings in that given year.⁵³ As with our path dependence measures, we set occupation wages equal to 0 if you are unemployed, and missing if you are employed but cannot be mapped to SOC codes. Importantly, we do not find evidence that crossing a threshold is related to missingness – a joint F-test across all 35 cutoffs that predicts missingness with our preferred RD specification yields a p-value of 0.725; hence, the missingness appears balanced across cutoffs.

We note that this measure is purely based on the reported SOC; there is no cross-walking of MOS to SOCs involved. In almost all cases, individuals with a SOC that we identify as still in the Army using the Army pay data report a string in the E-filed 1040 which the treasury linked to a military SOC code (e.g. the string “soldier” gets mapped to the broad military SOC code category of 55). This means that, for this measure, everyone still in the military receives average military earnings regardless of their military occupation – so most of the variation in average occupation wages arises from cutoff induced variation in civilian occupations.

As with earnings, we construct average occupation wages in each calendar year since application, for each applicant, from 11 to 20 years after application and then average these. We weight all regressions by the number of years between 11-20 for which we have tax data and the individual either had a SOC code or had no earnings.

Alternative Occupation Wage Variables. As robustness, we also report similar findings using SOC-specific average wages from an 50% random sample of Americans aged 31-44 in the tax data, as opposed to the in-sample, leave-out average (Panel A of Appendix Figure F.14). We also show comparable results using leave-out average industry wages – Average Industry Wages – based on 6 digit EIN numbers mapped to NAICS, which are observed for everyone with a W-2 regardless of the occupation they reported on their Form 1040 or even whether they filed (Panel A of Appendix Figure F.15).

Initial military occupation implied future earnings In Section 6.4 we calculate “military occupation implied future earnings” as the average occupation-wage across all SOC codes mapped to each MOS. This means that if an MOS is linked to a primary, secondary, and a single supervisory SOC code, we assign it the simple mean of the three SOC-level average long-run (11-20) occupation wages. Unlike many of our other exercises, this exercise relies on our MOS-specific crosswalk being accurate at the MOS-level. This

⁵²Military pay that is not reported on the W-2 as subject to Medicare tax encompasses various allowances and payments, such as the Basic Allowance for Housing (BAH), Basic Allowance for Subsistence (BAS), and additional compensation for deployment or foreign assignments, including Hardship Duty Pay, Imminent Danger Pay, Hazardous Duty Pay, and Family Separation Allowances. We do not account for other benefits, such as health coverage, retirement contributions, or G.I. Bill tuition and related housing allowances, some of which may have more natural civilian counterparts. Military pay not subject to taxation accounts for 17-25% of servicemembers' *total* compensation (see Appendix B of Greenberg et al., 2022).

⁵³Earnings are constructed similarly to our baseline, with one adjustment: since observing occupation already requires the filing of a tax return, we also include self-employment income reported on Form 1040 for a more comprehensive measure of earned income (though the results are very similar with their exclusion, due in part to the small share of self-employed).

procedure also relies on an equal weighting across the 1-4 linked SOCs for each MOS. Moreover, since we only begin observing SOCs in 2005 in the tax data, we can't easily construct SOC average wages for those who don't join the Army. For these civilians, we instead regress average occupation wages on AFQT score and use the implied fitted value. The result of this procedure is to have, for every individual, based solely on the MOS they first joined (or if they didn't join, based on AFQT-specific predicted occupation wages), an 'implied future earnings' amount.

B.2 Secondary Outcomes

In this sub-section we provide additional details regarding variable construction for our secondary outcomes. Appendix Table F.4 provides a concise summary of the auxiliary data sources used to build these variables.

Tasks Characteristics

Variables: Non-routine tasks, Routine tasks

Using O*Net Database 20.1 (October 2015 release), we follow the procedure used in [Acemoglu and Autor \(2011\)](#) to construct non-routine and routine task measures at the 6-digit SOC level. We then use 2015 ACS data to get weight estimates for each 6-digit SOC code, which are joined with the O*Net 6-digit SOC codes to do a weighted aggregation to the 3-digit SOC code level.

Years: 2015

Data Sources: ACS 2015 from IPUMS; O*Net 20.1 (October 2015 release) including Work Activities and Work Context modules.

Union Membership

Variable: Union Member

Using CPS variable UNION from IPUMS extract of CPS, we create a union variable that is 1 if the survey respondent is a union member and 0 otherwise. We then use a weighted aggregation to the 3-digit SOC level.

Years: 2015

Data Sources: CPS ASEC 2015 downloaded from IPUMS; OCCSOC to SOC conversion table from IPUMS, downloaded 06 AUG 2024 at <https://usa.ipums.org/usa/volii/occtooccsoc18.shtml>.

Education Variables

Variables: Bachelors+, College Attendance, Associates Degree

We use NSC data connected to each individual applicant starting from 1992 through 2020. Bachelors+ indicates whether an individual received a Bachelor's degree or higher by 2020. College Attendance captures whether an individual attended any higher education by 2020. Associate's Degree indicates receipt of an Associates degree by 2020.

Years: 1992–2020

Data Sources: Internal Army Personnel records linked with National Student Clearinghouse.

Military Variables

Variables: Years Served in Army, MOS is Combat, Still Active Duty, Joined Active Duty, Months Deployed, Initial Contract Length

We use internal personnel data from 1992 through 2024 to construct these variables. Years Served in Army represents the aggregate number of years served in the active-duty Army between 1992 and 2024. MOS is Combat is an indicator variable equal to 1 if the applicant entered the Army under an MOS classified as a Combat role, specifically Infantry (CMF 11), Combat Engineers (CMF 12), Field Artillery (CMF 13), Air Defense Artillery (CMF 14), or Armor (CMF 19). Still Active Duty equals 1 if currently in active duty Army service. Joined Active Duty equals 1 if the individual joined the Army as active duty. Months Deployed represents the aggregate number of months deployed while in the Army. Initial Contract Length measures the length in years of the initial contracted time to serve in the Army.

Years: 1992–2024

Data Sources: Internal Army personnel records.

VA Disability (VADC)

Variable: VADC

Using Veterans Affairs data linked with internal Army personnel records, this is an indicator variable that takes a value of 1 if the individual received any Veterans Affairs Disability Compensation as of 2019.

Years: 1992–2020

Data Sources: Internal Army personnel records linked with VA internal data.

AIT Training Courses and Costs

Variables: Training weeks, Cost estimates

Using internal Army records on training course requirements and cost calculations, we construct a variable recording the number of weeks a new recruit of each MOS is required to receive occupation-specific training through their first 3 years in the Army (some MOS have additional training sequences after initial AIT). Note that new enlistees in the combat arms MOS (such as infantry) complete “One Station Unit Training” (OSUT) rather than AIT. Thus, we have counted the time spent in OSUT beyond basic training as the AIT equivalent for these MOS.

This data is supplemented by internal personnel records recording weeks spent at training as cross-validation and supplementing where training requirements may have changed over the years. We also use these internal cost records to estimate the total cost of training for each MOS. These cost estimates include physical equipment costs, instructor costs, and individual pay costs.

Years: 2022, supplemented by personnel records from 1992–2024

Data Sources: Internal Army cost calculations for training requirements by MOS; internal Army personnel records.

Academic Credits

Variables: College credits, Vocational credits

The American Council on Education (ACE) works with the Department of Defense to generate course credit recommendations for military training. We use underlying data provided by ACE, which includes academic

college credit recommendations as well as vocational credit recommendations. We construct credit-hour recommendations that each MOS receives within their first term (usually ranks through E4).

Years: 2023 ACE extract

Data Sources: ACE course credit recommendations.

Certification Requirements by SOC

Variables: *Certification requirements, Vocational training, Academy training, Higher education requirements*

Using an LLM, we synthesize data for each 6-digit SOC on training, certification, and college education requirements. This is done at the state level where possible, pulling from various sources depending on the occupation. The primary source is the Bureau of Labor Statistics (BLS) “Certification and Licensing by Occupation” report and its “Professional Certifications and Occupational Licenses Analysis.” These are supplemented with reports and data compiled by various NGOs and research centers. Most sources do not list by SOC codes, so we rely on the LLM to create the necessary links between occupations and SOC codes.

A variable for each state-occupation pair indicates whether a special certification is required, vocation-specific training is required, academy training is required (such as police or fire academy, defined as training lasting more than 3 months at a specialized training institution), or higher education requirements (associates, bachelors, law degree, etc.).

Once compiled at the state level for each SOC, we use BLS estimates of employees in each occupation in each state to create a weighted aggregation of rough 6-digit occupation codes to 3-digit codes. These are then aggregated to the national level using state-level occupation employee weights. The aggregated data is interpretable as the fraction of workers in that occupation nationally that require certification, vocational training courses, college degrees, or academy training. The complete synthesized dataset is available in the replication package.

Primary Data Sources:

Bureau of Labor Statistics (BLS): Certification and Licensing by Occupation (<https://www.bls.gov/cps/cpsaat53.htm>); Occupational Employment and Wage Statistics (<https://www.bls.gov/oes/current/oessrcst.htm>); Professional Certifications and Occupational Licenses Analysis (<https://www.bls.gov/opub/mlr/2019/article/professional-certifications-and-occupational-licenses.htm>).

State Occupational Licensing Index (SOLI): Master Data Download (https://www.archbridgeinstitute.org/wp-content/uploads/2024/08/MasterData_SOLI2024.xlsx).

Institute for Justice: License to Work 3rd Edition 2024 (<https://ij.org/report/license-to-work-3/>); License to Work 2nd Edition 2017 (<https://ij.org/report/license-to-work-2/>); Occupational Licensing Research (<https://ij.org/issues/economic-liberty/occupational-licensing/>).

National Conference of State Legislatures (NCSL): National Occupational Licensing Database (<https://www.ncsl.org/labor-and-employment/the-national-occupational-licensing-database>); Occupational Licensing Legislation Database (<https://www.ncsl.org/labor-and-employment/occupational-licensing-legislation-database>).

Knee Regulatory Research Center (West Virginia University): Annual Licensing Database Snapshot 2024 (<https://csorwvu.com/annual-licensing-database-snapshot-2024-4424/>); Universal Licensing Reforms Survey

(<https://csorwvu.com/policy-brief-survey-of-universal-licensing-reforms-in-the-united-states-2024/>).

Additional Healthcare-Specific Sources: Federation of State Medical Boards

(<https://www.fsmb.org/step-3/state-licensure/>); National Governors Association Health Workforce Toolkit (<https://www.nga.org/state-health-workforce-toolkit/licensing-and-regulation/>).

C Robustness to other measures of career progression

Linkages with large language models. As robustness to our main specification, we also linked military occupations to civilian occupations using a large language model. Our procedure for linking utilized a two-stage prompt. In the first stage, we used the OpenAI API to submit the following prompt to chatGPT (model 4.0) for every military occupation specialty, one-by-one:⁵⁴

I am going to give you some information about an Army occupation. That information will include:

- (A) MOS: this will be the 3 character code the Army uses for the occupation.
- (B) Name: this will be the name of the Army occupation.
- (C) Major duties: this will be a description of the important duties soldiers in this occupation would perform while in the Army.
- (D) E1-E4 duties: this will be a description of the important duties a soldier in this occupation would perform while enlisted in a lower rank, such as the rank of private, private first class, or specialist.
- (E) E5 duties: this will be a description of the important duties a soldier in this occupation would perform while enlisted at the rank of sergeant.

I want you to create four lists using the information contained in the descriptions that I provide:

- (1) Military tasks: Create a list of tasks that someone in this military occupation will perform as part of their military job. Limit this to tasks that are very specific to the military and are not broadly used anywhere in the civilian labor force.
- (2) Civilian tasks: Create a list of tasks that someone in this military occupation will perform as part of their military job. But now describe all tasks that are also broadly used in the civilian labor force outside of the military.
- (3) Military skills: Create a list of skills that someone in this military occupation will need to perform the tasks in 1 and 2. Limit this response to skills that are very specific to the military and are not broadly useful in the civilian labor force.
- (4) Civilian skills: Create a list of skills that someone in this military occupation will need to perform the tasks in 1 and 2. But now describe all skills that are also broadly used in the civilian labor force outside of the military.

What is the difference between a task and a skill? A task is an action or duty that the individual will perform as part of their job. A skill is a trait or ability that will help them perform these tasks effectively.

Here are some examples:

- Managing a complicated filing system is a task. Organization and attention to detail are skills that might be helpful for this task.

⁵⁴The string of symbols requested in the LLM output ("\$#!") was to facilitate processing the text ex-post.

- Walking long distances with heavy equipment is a task. Physical fitness and endurance are skills that might be helpful for this task.
- Training junior soldiers is a task. Leadership and teamwork are skills that might be helpful for this task.
- Installing computer IT systems is a task. Logical thinking and computer coding are skills that might be helpful for this task.
- Repairing car engines is a task. Knowledge of tools and manual dexterity are skills that might be helpful for this task.

Please provide your response to items (1), (2), (3), and (4) formatted as follows:

(1) Military tasks: {YOUR RESPONSE TO 1}

\$#!

(2) Civilian tasks: {YOUR RESPONSE TO 2}

\$#!

(3) Military skills: {YOUR RESPONSE TO 3}

\$#!

(4) Civilian skills: {YOUR RESPONSE TO 4}

Let's start. Here is the information on the Army occupation:

To this general prompt, we appended MOS-specific military occupation descriptions. These descriptions were drawn from official Army documentation, specifically DA PAM 611-12 which outlines military occupational classification and structure ([Department of the Army, 2022](#)). As an example, here is the description that we appended to the end of the prompt for the plumber MOS:

(A) MOS: 12K

(B) Name: Plumber

(C) Major duties: The plumber installs and repairs pipe systems and fixtures, and petroleum pipeline systems. Duties for MOS 12K at each level of skill are:

(D) E1-E4 duties: Installs steel, plastic, and copper pipes, water closet with tank, urinals, water heaters, shower systems, water systems, waste systems, lavatories, and cross linked polyethylene (PEX) pipes. Repairs valves and fixtures control devices, steel, plastic, and copper pipes, cast iron pipes, and PEX pipes. Maintains plumbing tools, grades sewer lines, tests waste and water systems, clears waste systems stoppages, identifies plumbing materials, winterizes water systems, interprets plumbing prints and drawings, prepares a material takeoff list, and rescues an electrical shock victim.

(E) E5 duties: Installs cast iron pipes and draws a plumbing job plan. Coordinates construction planning and advises on plumbing construction. Inspect plumbing systems/components for proper installation and ensures employment of proper safety procedures.

And here is the response the large language model returned:

Based on the information provided for the Army occupation MOS 12K, Plumber, here are the tasks and skills categorized into military and civilian contexts:

(1) Military tasks:

- Installation and repair of petroleum pipeline systems.
- Winterizing water systems.
- Rescuing an electrical shock victim.
- Coordinating construction planning specific to military operations.
- Advising on plumbing construction within military settings.

\$#!

(2) Civilian tasks:

- Installing and repairing pipe systems and fixtures.
- Installing steel, plastic, copper, and PEX pipes.
- Installing water closets, urinals, water heaters, shower systems, and lavatories.
- Repairing valves, fixtures, control devices, and various types of pipes.
- Maintaining plumbing tools and equipment.
- Grading sewer lines and testing waste and water systems.
- Clearing waste system stoppages.
- Identifying plumbing materials.
- Interpreting plumbing prints and drawings.
- Preparing a material takeoff list.
- Drawing a plumbing job plan.
- Inspecting plumbing systems/components for proper installation.
- Ensuring employment of proper safety procedures.

\$#!

(3) Military skills:

- Knowledge of military-specific petroleum pipeline systems.
- Ability to perform emergency rescue operations, such as rescuing an electrical shock victim.
- Understanding of military construction planning and coordination.
- Familiarity with military safety protocols and procedures.

\$#!

(4) Civilian skills:

- Proficiency in plumbing installation and repair techniques.
- Ability to work with various materials like steel, plastic, copper, and PEX.
- Skills in using and maintaining plumbing tools.
- Competence in reading and interpreting plumbing blueprints and drawings.
- Problem-solving skills for clearing stoppages and repairing systems.
- Attention to detail for inspecting installations and ensuring safety.
- Planning and organizational skills for preparing material lists and job plans.

- Communication skills for advising and coordinating with others.

\$#!

Once we had results from this prompt for every MOS, we moved on to the second stage.

In the second stage, we opened a new instance of GPT 4.0 via the OpenAI API, and asked the large language model to estimate long-run MOS to SOC transition probabilities using only the civilian relevant tasks and skills. Note that opening a new, clean API instance was important, since it ensured that the response would not be contaminated by the first prompt and which could potentially undermine our goal of linking primarily on the basis of skills and abilities, and not on the basis of military job titles, etc.

Specifically, in the second phase, we submitted the following prompt, one-by-one, for each MOS:

I am going to give you a list of tasks and skills.

I want you to imagine someone who spent 5 years performing those tasks and developing those skills as part of the job they held in the Army from ages 20-25. Now, using primarily the tasks and skills that I will provide you, I want you to predict the 10-20 occupations that the individual is most likely to have 15 years later, at age 40, after they have left the Army. Please give your answer in the form of a six digit BLS SOC code, with the associated SOC title, and please assign a probability to each of these six digit SOC occupations that represents how likely that person is to be in that occupation at age 40.

Please provide your answer in the following format:

(1) {YOUR FIRST BLS SOC CODE} ^\$^ {TITLE OF YOUR FIRST BLS SOC CODE} ^\$^ {PROBABILITY YOU BELIEVE INDIVIDUAL WILL BE IN THIS SOC OCCUPATION AT AGE 40}

\$#!

(2) {YOUR SECOND BLS SOC CODE} ^\$^ {TITLE OF YOUR SECOND BLS SOC CODE} ^\$^ {PROBABILITY YOU BELIEVE INDIVIDUAL WILL BE IN THIS SOC OCCUPATION AT AGE 40}

\$#!

(3) {YOUR THIRD BLS SOC CODE} ^\$^ {TITLE OF YOUR THIRD BLS SOC CODE} ^\$^ {PROBABILITY YOU BELIEVE INDIVIDUAL WILL BE IN THIS SOC OCCUPATION AT AGE 40}

\$#!

and continuing the list in that format until you have provided the required information for all 10-20 occupations that you predict the individual is most likely to have 15 years later, at age 40, after they have left the Army.

Remember, it is important that you primarily use the information contained in the tasks and skills I provide you when predicting the future occupations.

To this prompt, we appended the MOS civilian tasks and skills found in stage 1 prior to submitting it via the API. Here is an example of what we sent for the plumber MOS:

Here are the types of tasks this individual performed:

- Installing and repairing pipe systems and fixtures.
- Installing steel, plastic, copper, and PEX pipes.
- Installing water closets, urinals, water heaters, shower systems, and lavatories.

- Repairing valves, fixtures, control devices, and various types of pipes.
- Maintaining plumbing tools and equipment.
- Grading sewer lines and testing waste and water systems.
- Clearing waste system stoppages.
- Identifying plumbing materials.
- Interpreting plumbing prints and drawings.
- Preparing a material takeoff list.
- Drawing a plumbing job plan.
- Inspecting plumbing systems/components for proper installation.
- Ensuring employment of proper safety procedures.

Here are the types of skills this individual developed:

- Proficiency in plumbing installation and repair techniques.
- Ability to work with various materials like steel, plastic, copper, and PEX.
- Skills in using and maintaining plumbing tools.
- Competence in reading and interpreting plumbing blueprints and drawings.
- Problem-solving skills for clearing stoppages and repairing systems.
- Attention to detail for inspecting installations and ensuring safety.
- Planning and organizational skills for preparing material lists and job plans.
- Communication skills for advising and coordinating with others.

Now please use those tasks and skills to do what I've asked."

To this prompt, the large language model returned the following response.

Based on the tasks and skills provided, the individual has developed a strong foundation in plumbing, installation, repair, and system inspection. This experience and skill set can translate into various civilian occupations. Here are the predicted occupations for this individual at age 40, along with the probabilities:

- (1) 47-2152 ^\$^ Plumbers, Pipefitters, and Steamfitters ^\$^ 0.25
\$#!
- (2) 49-9021 ^\$^ Heating, Air Conditioning, and Refrigeration Mechanics and Installers ^\$^ 0.15
\$#!
- (3) 47-1011 ^\$^ First-Line Supervisors of Construction Trades and Extraction Workers ^\$^ 0.10
\$#!
- (4) 47-3015 ^\$^ Helpers--Pipelayers, Plumbers, Pipefitters, and Steamfitters ^\$^ 0.08
\$#!
- (5) 47-4011 ^\$^ Construction and Building Inspectors ^\$^ 0.07

\$#!
(6) 49-9098 ^\$^ Helpers--Installation, Maintenance, and Repair Workers ^\$^ 0.06
\$#!
(7) 49-1011 ^\$^ First-Line Supervisors of Mechanics, Installers, and Repairers ^\$^ 0.06
\$#! (8) 47-2073 ^\$^ Operating Engineers and Other Construction Equipment Operators ^\$^ 0.05
\$#!
(9) 47-2111 ^\$^ Electricians ^\$^ 0.05
\$#!
(10) 47-3013 ^\$^ Helpers--Electricians ^\$^ 0.04
\$#!
(11) 49-9041 ^\$^ Industrial Machinery Mechanics ^\$^ 0.03
\$#!
(12) 47-3019 ^\$^ Helpers, Construction Trades, All Other ^\$^ 0.03
\$#!
(13) 47-2181 ^\$^ Roofers ^\$^ 0.02
\$#!
(14) 47-2221 ^\$^ Structural Iron and Steel Workers ^\$^ 0.02
\$#!
(15) 49-9071 ^\$^ Maintenance and Repair Workers, General ^\$^ 0.02
\$#!
(16) 47-2031 ^\$^ Carpenters ^\$^ 0.01
\$#!
(17) 47-2061 ^\$^ Construction Laborers ^\$^ 0.01
\$#!
(18) 49-9099 ^\$^ Installation, Maintenance, and Repair Workers, All Other ^\$^ 0.01
\$#!

These probabilities reflect the likelihood of transitioning into these roles based on the individual's experience and skills in plumbing and related tasks. The highest probability occupations are those directly related to plumbing and supervisory roles in construction and maintenance, leveraging both technical skills and leadership potential.

For the robustness check, we then aggregated these to the three-digit standard occupation classification code level – this rendered them comparable to our baseline results. We then estimated our baseline model, but with the outcome variable defined to be the share of years spent at a civilian employer in *any* SOC code linked by the LLM to an MOS unlocked at the cutoff. The results are contained in appendix table F.8.

MOS specific weights. As another robustness check, we estimated our baseline model with an outcome variable for “in related career” defined on the basis of the individual military occupation specialties most impacted by the cutoff. The intuition is that, if plumbers and cooks are both tied to the same threshold, but that threshold only impacts an applicant’s likelihood of joining the military as a plumber (but not cook), we should only expect to see long-run effects on the likelihood of that individual becoming a civilian plumber.

To implement this idea, we take each cutoff and estimate “pseudo” first stages on each military occupation unlocked by that cutoff. Next, we redefine our outcome variable to capture impacts on career trajectory by “upweighting” the SOC codes linked to MOS with the largest pseudo-first stage estimates. Specifically, for each cutoff we construct a list of linked SOC codes with weights that sum to 1, where SOC-level weights are generated as follows: First we give each $MOS \times$ linked SOC at that cutoff a weight of $\mu_{c,mos,soc} = w \frac{\pi_m}{\pi}$, where w is either 1/1, 1/2, 1/3, or 1/4 depending on whether that MOS is linked to 1 to 4 SOCs, π_m denotes the pseudo first stage effect that crossing cutoff c has on joining the army in MOS m as estimated using our baseline model, and π denotes the full first-stage estimate (defined as the sum of the MOS-specific first stages). Then we collapse to the cutoff-SOC level, summing these weights, so that if every MOS was only linked to a single SOC, each SOC would receive $\sum \frac{\pi_m}{\pi}$ weight, where the sum is over all the focal MOS linked to the SOC. At this point, we have a data set with cutoff-by-SOC level weights that sum to 1 across all the linked SOC codes within each cutoff: call them $\mu_{c,soc}$. We generate the final outcome weights as $\omega_{c,soc} = |M_{soc}| \mu_{c,soc}$, where $|M_{soc}|$ is the number of linked SOCs at the cutoff. This ensures that SOC codes linked to MOS that make up a larger share of the total first stage get a weight greater than 1 and SOC codes linked to MOS with a smaller share get weights less than 1. Our outcome variable thus takes on a value of 0 in each year you are not employed in a civilian SOC unlocked by the cutoff, and $\omega_{c,soc}$ for each year you are employed in the corresponding SOC. We average this over the 11-20 years after ASVAB date.

Because this upweighting changes the distribution of the outcome variable, we standardize the outcome variable at each cutoff to be mean 0 and standard deviation 1 before estimation. Results using this alternative measure of career path are contained in Appendix table F.8, column (3).

We also repeat the same upweighting procedure for the LLM-generated outcomes, but here also incorporate the LLM-generated probabilities for each three-digit SOC. We ensure these probabilities add up to 1 within MOS and then simply use these instead of the w (that was either 1/1, 1/2, 1/3, or 1/4) above. The rest of the process is the same and results using this measure of career path are contained in Appendix table F.8, column (4).

D Equivalence of main results to corresponding micro-data regressions

In the main text, our preferred method for summarizing the variation across cutoffs is to directly estimate the slope of the V-IV line in figure 3 (for our results on path dependence) and to directly estimate the slope of the causal effect on wages versus causal effect on occupation wages (“IV-on-IV”) line in figure 5. In this appendix, we give conditions under which these slopes are statistically equivalent to a properly specified two-stage least squares regression estimated directly on the micro data.

Our argument proceeds in two stages. First, we show that the slope of a line that connects reduced form effects from multiple “experiments” (equivalently, instruments) on one outcome to reduced form effects on another outcome is numerically equivalent (up to a weighting) to a properly specified 2SLS regression on the micro-data.⁵⁵ While this link is already well known (e.g. [Angrist and Hull, 2023](#)), we restate the argument here for convenience. Second, we link the slope of the “IV-on-IV” line (e.g. figure 5) to the slope of the corresponding “reduced-form on reduced-form” line developed in part one of the argument. Equivalence

⁵⁵Note that by viewing the “first-stage” as a reduced form outcome, this argument nests that canonical “V-IV” slope.

between the “IV-on-IV” slope and the slope from the micro data regression then follows immediately.

Slope of the “Reduced-form on Reduced-form” line and 2SLS on the micro data

We begin by rewriting the reduced form equation for each cutoff c and applying the Frisch–Waugh–Lovell theorem to residualize on $f_c(X_{ic})$:

$$\tilde{Y}_{ic} = \gamma_c \tilde{Z}_{ic} + u_{ic}, \quad \mathbb{E}[\tilde{Z}_{ic} | c] = 0. \quad (\text{D.1})$$

Analogously, define α_c for some alternative outcome via $\tilde{Y}'_{ic} = \alpha_c \tilde{Z}_{ic} + u'_{ic}$. Collect the reduced-form slopes as $\gamma = \{\gamma_c\}$ and $\alpha = \{\alpha_c\}$.⁵⁶ Next, stack the residualized micro data:

$$Y = \{\tilde{Y}_{ic}\}, \quad (\text{D.2})$$

$$X = \{\tilde{Y}'_{ic}\}, \quad (\text{D.3})$$

$$Z = [Z_1, \dots, Z_{35}], \quad (\text{D.4})$$

where each Z_c is the $N \times 1$ column equal to \tilde{Z}_{ic} for observations belonging to cutoff c and zero otherwise. Then $Z'Z = D = \text{diag}(Z'_1 Z_1, \dots, Z'_{35} Z_{35})$ is diagonal. For each c ,

$$\gamma_c = \frac{Z'_c Y}{Z'_c Z_c}, \quad \alpha_c = \frac{Z'_c X}{Z'_c Z_c}, \quad (\text{D.5})$$

so that $Z'Y = D\gamma$ and $Z'X = D\alpha$.

Consider the 2SLS model with one endogenous regressor:

$$Y = \beta^{2SLS} X + e, \quad (\text{D.6})$$

$$X = \pi Z + v. \quad (\text{D.7})$$

With one endogenous regressor, the 2SLS coefficient is

$$\hat{\beta}^{2SLS} = \frac{(Z'X)'(Z'Z)^{-1}(Z'Y)}{(Z'X)'(Z'Z)^{-1}(Z'X)} \quad (\text{D.8})$$

$$= \frac{(D\alpha)'D^{-1}(D\gamma)}{(D\alpha)'D^{-1}(D\alpha)} \quad (\text{D.9})$$

$$= \frac{\alpha'D\gamma}{\alpha'D\alpha} \quad (\text{D.10})$$

Which is the solution to the (through the origin) weighted least squares problem,

$$\arg \min_{\beta} \sum_c \omega_c (\gamma_c - \beta \alpha_c)^2, \quad (\text{D.11})$$

with weights given by

⁵⁶Observe that this notation nests that classic V-IV argument, since we can simply set Y'_{ic} to be an indicator for “treatment take-up.” In that case, α_c just becomes the experiment-specific or instrument-specific first stage.

$$\omega_c = Z'_c Z_c = \sum_{i \in c} \tilde{Z}_{ic}^2 = n_c \operatorname{Var}(\tilde{Z}_{ic} | c) \quad (\text{D.12})$$

i.e., the residual variance of the instrument within cutoff c . This shows that the slope of the RF-on-RF line is equivalent (up to a weighting) to 2SLS on the micro data.

Equivalence of “IV-on-IV” line to “reduced-form on reduced-form”. Can the slope of an “IV-on-IV” line (as in figure 5) be estimated from the same micro-data regression as the “reduced-form on reduced-form” 2SLS specification? Intuitively, if the experiment- or instrument-specific first stages are orthogonal to the underlying causal effects, then rescaling a “reduced-form on reduced-form” plot by these first stages should, in expectation, not change the underlying slope.

More formally, let π_c denote the associated first-stage from each cutoff, and let β_c denote the causal effect of the treatment itself on the first outcome, with β'_c denoting the causal effect of treatment on the second outcome. We are looking for conditions such that the “RF-on-RF” slope is equal to the “IV-on-IV” slope:

$$\frac{\operatorname{cov}(\pi_c \beta_c, \pi_c \beta'_c)}{\operatorname{var}(\pi_c \beta'_c)} = \frac{\operatorname{cov}(\beta_c, \beta'_c)}{\operatorname{var}(\beta'_c)} \quad (\text{D.13})$$

Observe that if there is no variation in the first stage (i.e. $\pi_c = \pi$) then the equality holds trivially, so without loss of generality we can assume that $\operatorname{var}(\pi_c) \neq 0$. Now suppose further that $\pi_c \perp (\beta_c, \beta'_c)$. Then we can rewrite this equality as:^{57,58}

$$\frac{\mathbb{E}(\pi_c^2) \operatorname{cov}(\beta'_c, \beta_c) + \operatorname{var}(\pi_c) \mathbb{E}(\beta') \mathbb{E}(\beta_c)}{\mathbb{E}(\pi_c^2) \operatorname{var}(\beta'_c) + \operatorname{var}(\pi_c) \mathbb{E}(\beta'_c)^2} = \frac{\operatorname{cov}(\beta_c, \beta'_c)}{\operatorname{var}(\beta'_c)} \quad (\text{D.14})$$

⁵⁷To derive the numerator of the left-hand side of this expression:

$$\begin{aligned} \operatorname{cov}(\pi_c \beta_c, \pi_c \beta'_c) &= \mathbb{E}(\pi_c \beta_c \pi_c \beta'_c) - \mathbb{E}(\pi_c \beta_c) \mathbb{E}(\pi_c \beta'_c) \\ &= \mathbb{E}(\pi_c^2 \beta_c \beta'_c) - (\mathbb{E}(\pi_c) \mathbb{E}(\beta_c)) (\mathbb{E}(\pi_c) \mathbb{E}(\beta'_c)) \\ &= \mathbb{E}(\pi_c^2) \mathbb{E}(\beta_c \beta'_c) - \mathbb{E}(\pi_c)^2 \mathbb{E}(\beta_c) \mathbb{E}(\beta'_c) \\ &= \mathbb{E}(\pi_c^2) (\operatorname{cov}(\beta_c, \beta'_c) + \mathbb{E}(\beta_c) \mathbb{E}(\beta'_c)) - \mathbb{E}(\pi_c)^2 \mathbb{E}(\beta_c) \mathbb{E}(\beta'_c) \\ &= \mathbb{E}(\pi_c^2) \operatorname{cov}(\beta_c, \beta'_c) + (\mathbb{E}(\pi_c^2) - \mathbb{E}(\pi_c)^2) \mathbb{E}(\beta_c) \mathbb{E}(\beta'_c) \\ &= \mathbb{E}(\pi_c^2) \operatorname{cov}(\beta_c, \beta'_c) + \operatorname{var}(\pi_c) \mathbb{E}(\beta_c) \mathbb{E}(\beta'_c) \end{aligned}$$

⁵⁸To derive the denominator of the left-hand side of this expression:

$$\begin{aligned} \operatorname{var}(\pi_c \beta_c) &= \mathbb{E}((\pi_c \beta_c)^2) - \mathbb{E}(\pi_c \beta_c)^2 \\ &= \mathbb{E}(\pi_c^2) \mathbb{E}(\beta_c^2) - \mathbb{E}(\pi_c)^2 \mathbb{E}(\beta_c)^2 \\ &= \mathbb{E}(\pi_c^2) (\operatorname{var}(\beta_c) + \mathbb{E}(\beta_c)^2) - \mathbb{E}(\pi_c)^2 \mathbb{E}(\beta_c)^2 \\ &= \mathbb{E}(\pi_c^2) \operatorname{var}(\beta_c) + (\mathbb{E}(\pi_c^2) - \mathbb{E}(\pi_c)^2) \mathbb{E}(\beta_c)^2 \\ &= \mathbb{E}(\pi_c^2) \operatorname{var}(\beta_c) + \operatorname{var}(\pi_c) \mathbb{E}(\beta_c)^2. \end{aligned}$$

From there, cross multiply and simplify to get:

$$\mathbb{E}(\beta'_c)\mathbb{E}(\beta_c)var(\beta') = cov(\beta'_c, \beta_c)\mathbb{E}(\beta'_c)^2 \quad (\text{D.15})$$

Which further simplifies to:

$$\mathbb{E}(\beta_c) - \frac{cov(\beta_c, \beta'_c)}{var(\beta'_c)}\mathbb{E}(\beta'_c) = 0 \quad (\text{D.16})$$

This condition just says that the intercept of the “IV-on-IV” plot must pass through zero. Thus, the slope of the “IV-on-IV” line can be estimated directly from the micro data whenever (1) the first stage is orthogonal to the underlying causal effect; and (2) the intercept of the IV-on-IV line passes through zero.

Regarding the first condition, we note here that we do not find a statistically significant relationship between the cutoff specific IV-estimates on wages and the corresponding first-stages. This is consistent with the independence assumption. See panel (a) of figure F.18.

Regarding condition (2), note that this is implied by the slightly stronger proportionality condition discussed in [Angrist and Hull \(2023\)](#):

$$\gamma_c = \beta\alpha_c \quad (\text{D.17})$$

As shown in [Angrist and Hull \(2023\)](#), the over-ID test in table F.7 is a direct test of proportionality (i.e. equation D.17). In all cases, we fail to reject. Thus, the proportionality assumption appears to be satisfied, and hence condition (2) is also satisfied in our data.

Discussion. We present results from the corresponding micro-data models for all of our key results in appendix table F.7. In all cases, the results are similar to our preferred method.

This robustness check is useful for three reasons. First, we note that the equivalence holds only up to a weighting. Thus, the robustness check shows that, in practice, our results are not sensitive to the weights. Second, it confirms that our results are not affected by classical or non-classical measurement error. This is because the micro-regression does not require a pre-estimation step and hence all variables involved in the micro-regression are measured without estimation error from the pre-step. Third, it provides a tractable way for us to incorporate cutoffs with small (but strong) first stages into the analysis. Our baseline specification uses only cutoffs with a first stage larger than 0.01. We do this in order to facilitate transparent analysis of the underlying RD variation – some of the cutoffs with a small first stage yield extremely large point estimates for earnings as a result of the imprecision in the reduced form, which renders plots like figures 3 and 5 illegible, dominates simple averages, etc. These issues are less important with 2SLS on the micro-data, since the influence of these cutoffs is reduced as a result of the weighting, which is proportional to the strength of the first stage ([Angrist and Pischke, 2009](#)). In theory, the size restriction from our preferred specification could introduce a bias. In practice, however, the micro-data regressions that include an expanded set of cutoffs meeting a more traditional F-stat screen of 10 yield nearly identical results (see appendix table F.7).

E “Partially-pooled” multiple endogenous variable model details

In this section, we explain how we develop and estimate our “partially-pooled” multiple endogenous variable model of the impact of specific military occupations. Our goal is to estimate the effect of starting one’s career in one of nine mutually exclusive career groups in a 2SLS model with 8 endogenous treatment variables (with civilian careers representing the 9th, omitted category) using threshold-crossing dummies as instruments.

We begin by discussing how we construct these 9 career groups. The smallest occupational groupings we can hope to identify at any given cutoff consist of the MOS’s unlocked by that cutoff. For example, CL90 (post-2004) unlocks several MOS broadly related to logistics: 42A, 68G, 68J, 92A, 92Y, and 92F. Prior to 2002, CL95 unlocked this same set of MOS. Hence, we can define a “warehouse operations / HR” career group that consists of these 6 MOS and the corresponding (single) instrument for this group will be whether you crossed either the CL90 cutoff (post-2004) or the CL95 cutoff (pre-2002).⁵⁹ In theory, we could then define a corresponding career group for every other ‘pair’ of cutoffs in the pre-2002 and post-2004 period, a process that would have resulted in 14 endogenous treatment variables. However, in order to improve power, we further group similar occupations together. For example, we combine the infantry and field artillery cutoff-unlocked MOS into one combat group. The corresponding instruments for this group are then whether or not you cross the infantry-related cutoff, or separately, whether or not you cross the field artillery cutoff. We further opt to group together the occupations unlocked by the various electronics cutoffs we studied (into Computer/Electronics), the occupations unlocked by the various mechanical maintenance cutoffs (into Mechanics/Maintenance), and the occupations unlocked by the various skilled technical cutoffs (into Health/Police/Intelligence/Other Specialists).

The end result of this pooling is that we have nine mutually exclusive career groups – Warehouse Operations/HR, Mechanics/Maintenance, Transport/Food Preparation, Utilities/Equipment Management, Health/Police/Intelligence/Other Specialists, Combat/Construction, Computers/Electronics, All Other Army Occupations, and Civilian (did not join the Army). The list of MOS that comprise each career group can be found in Table F.15, alongside the relevant focal cutoffs. Some of these career groups will be unlocked by a single corresponding instrument (e.g. whether you crossed the relevant CL cutoff in either period). Others are over-identified (e.g. the combat/construction group can be unlocked either by crossing the infantry-related cutoff or by the crossing the field artillery cutoff). In addition, in order help identify the coefficient on ‘all other army MOS’ and facilitate norming relative to being a civilian, we leverage the AFQT thresholds from Greenberg et al. (2022) for exceeding an AFQT of 31 or 50 based on your first AFQT on file. The end result of this process is that we seek to identify coefficients on 8 treatment indicators using 16 instruments.⁶⁰

With these definitions in mind, we estimate the following multiple endogenous variable two-stage least squares model:

$$Y_i = \beta D_i + f(\mathbf{X}_i) + \epsilon_i \quad (E.1)$$

$$D_{ig} = \Pi_g Z_i + h_g(\mathbf{X}_i) + \eta_{ig}, \quad g = 1, 2, \dots, 8 \quad (E.2)$$

⁵⁹Individuals only appear in one of these two application periods since we restrict to non-prior service applicants.

⁶⁰Note that these 16 instruments are all instruments for crossing a threshold in *either* the pre-2002 or post-2004 period, so they correspond to 32 cutoffs in our baseline specifications. 28 of these 32 are occupation cutoffs and the other 4 are AFQT cutoffs. We utilize 28 occupation cutoffs, as opposed to the 35 in our baseline, because we omit a handful of cutoffs that we had at baseline but only appeared in one of the pre-2002 or post-2004 periods without their ‘pair’ (since the pair had an insufficient first stage). As an example, we omit cutoffs like CL90 pre, which shifts only MOS 92F in the pre-2002 period and has an analogous CL86 cutoff in the post-2004 period but the first stage post-2004 is extremely small. In such cases, MOS 92F will be a part of the ‘other Army MOS’ category.

Where Y_i is 11-20 earnings; $D_i = \{D_{ig}\}$ is a vector of 8 treatment indicators, one for each of the 8 occupation groups with “remaining a civilian” constituting the omitted category; $X_i = \{X_{il}\}$ is a vector of the various line scores; and f maps line scores into the series of piecewise linear and quadratic polynomials around the various RD thresholds as implied by our baseline, individual RD models. We instrument for the treatment variables with 8 first stage equations of the form shown in E.2, where D_{ig} takes a value of 1 if applicant i enlists in any occupation contained in group g ; $Z_i = \{Z_{ic}\}$ is a vector of indicator variables that take a value of one if individual i crosses the line score threshold c ; and h_g maps line scores into the series of piecewise linear and quadratic polynomials around the various RD thresholds as implied by our baseline, individual RD models. All the running variables in both equations are interacted with whether the applicant is from the pre-2002 or post-2004 period and a control for post-2004 is included, so that we effectively fit RD models to each period-specific cutoff, even though the instrument is a single instrument for crossing the relevant threshold in either the pre- or post-period. The polynomials and instruments are always set to 0 outside the relevant RD windows (which in the case of multiple cutoffs per line score span all the cutoffs), and we control separately for being in each possible RD window.

We interpret this model through the lens of a constant-linear effects framework. As such, this model identifies interpretable causal parameters under strong assumptions. In addition to the standard RD assumptions outlined earlier in our analysis of the individual thresholds, this approach also requires that treatment effects as defined by our eight career groups are constant and linear. This assumption could fail in two ways. First, it could be that different occupations within a group have different causal effects. If different instruments move individuals across occupations within a group, this would violate the implicit exclusion restriction embedded in our constant linear effects assumption, since it would imply that our instruments generate causal impacts on earnings through a channel not captured by any of our endogenous variables. Second, it could be that treatment effects are heterogeneous in ways that are not captured by the eight treatment variables we have defined. This would be a form of model mis-specification, since it implies a violation of the functional form assumption (i.e. linearity) imposed by our model.

In support of these assumptions, we implement an over-identification test. Our method of pooling military careers within these different groups (e.g. pooling infantry and field artillery into one) leaves us with more instruments than endogenous regressors (the instrument for crossing the infantry cutoff and the instrument for crossing the field artillery cutoff in this case); hence we are overidentified. Intuitively, if different subsets of instruments produce drastically different estimates of the vector of causal effects (β in model E.1), then it must be due either to an exclusion restriction violation (e.g. the causal effect of being induced to swap occupations within a career group) or a functional form violation (e.g. a substantially different local average effect to some sub-population of compliers, which would imply that the linear constant effects model in E.1 is wrong). This is effectively what is tested using a classic Sargan-Hansen J-test (Wooldridge, 2010). We implement this test and fail to reject ($P = 0.144$), which suggests that our pooled model is sufficiently rich to account for the most important forms of heterogeneity in the observed data.

Appendix Table F.12 reports the 8 first stages (equation E.2). We always find positive and significant effects where expected, indicating that all the instruments are indeed pushing people into the relevant Army career groups and they do so at the magnitudes we anticipated based on the separate RDs. Also as expected, the ‘off-diagonal’ effects are generally negative or zero and smaller, indicating that the cutoffs pull people from a more diffuse set of counterfactual Army occupations. The AFQT cutoffs induce entry into most, but not all, Army career groups, including the < 4% of the sample in other Army jobs.⁶¹ As a separate point of interest, the AFQT thresholds *do not* tend to shift people into some of the more economically lucrative jobs in the army – jobs in computers and electronics jobs or in some of the other skilled technical specialist positions

⁶¹ Appendix Table F.13 shows what happens when we drop the category for other army jobs and instead consider the omitted category to be civilian or other army job; broad patterns are stable.

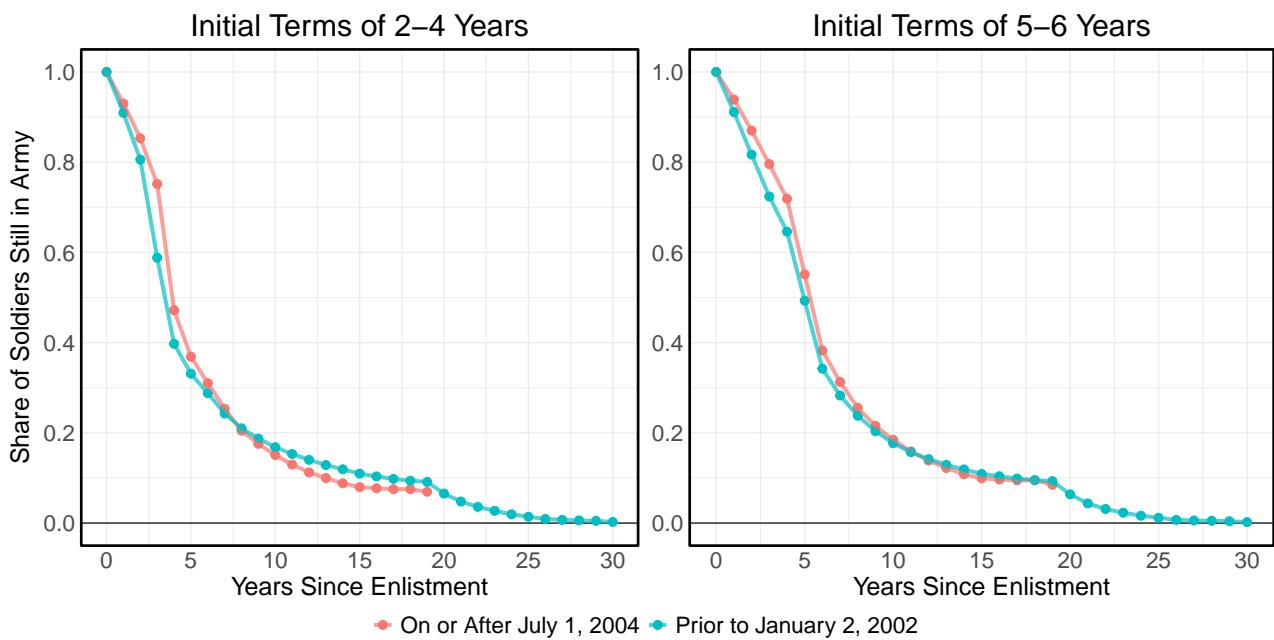
like health, police, intelligence. This suggests that the AFQT cutoff-induced variation in Greenberg et al. (2022) might under-estimate the ‘average’ effect of joining the Army on future earnings.

Appendix Table F.13 contains the corresponding estimates from estimating our multiple endogenous variable 2SLS model (Figure 8 presents the results visually). We discuss the resulting findings in the main text.

F Appendix Figures and Tables

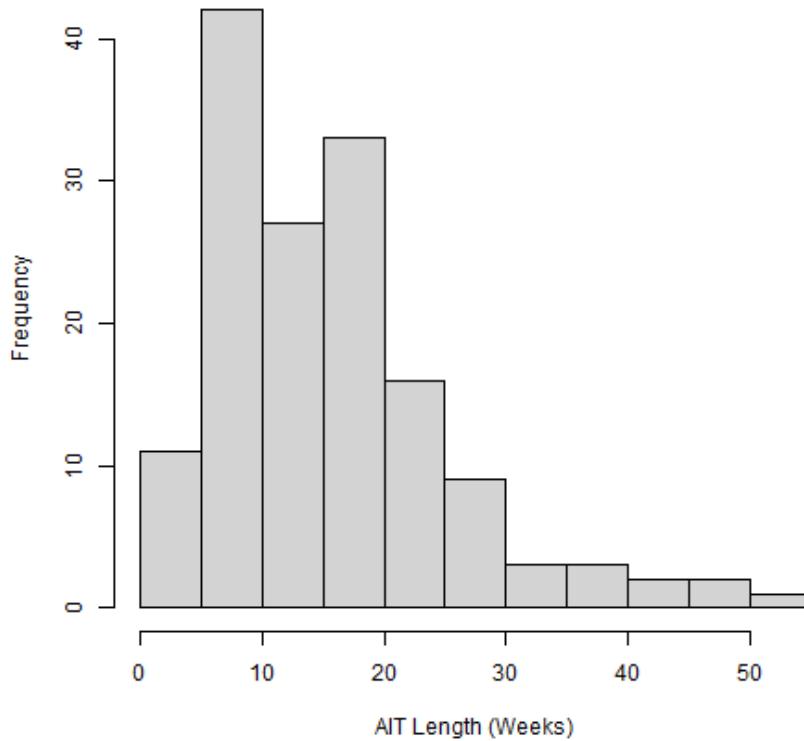
Appendix Figures

Figure F.1: Years in Army



Notes: This figure plots the share of soldiers still in the Army by years since entry to the Army. It conditions on those that enlist with a valid entry date and partitions the data into soldiers with initial entry terms of 2–4 years in the left panel and 5–6 years in the right panel. Within each panel, the share of soldiers still in the Army is calculated separately for enlistees from 1992–2001 (cyan) and from 2004–2014 (orange). All means are shown for the unbalanced set of soldiers for whom the relevant years since enlistment is prior to the end of our data in February 2024.

Figure F.2: Distribution of AIT Training Length Across MOS

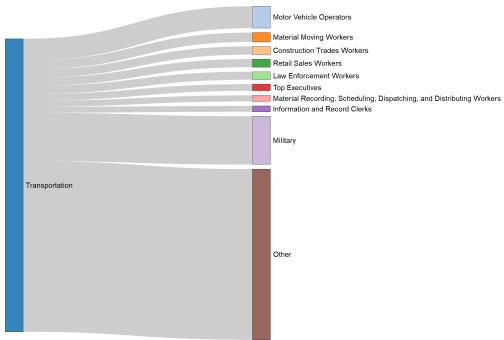


Notes: This histogram shows the distribution of AIT training requirements across all MOS in our sample, measured in weeks. For combat arms, in which new recruits move on to OSUT training after Basic Training, we have counted OSUT similarly to AIT training. This measure also includes extended training requirements that occur within the first two years of a new Soldier's experience. See Appendix Section B for detailed information on how these are constructed.

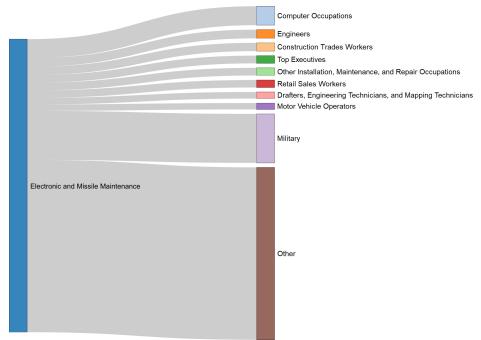
Figure F.3: Sankey Plots: Most Common Occupations (SOC Codes) 11-20 Years out by CMF



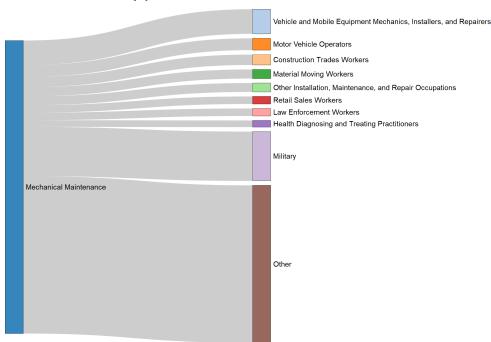
(g) Transportation



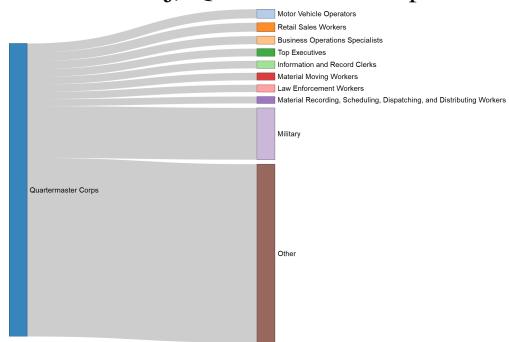
(h) Electronic and Missile Maintenance



(i) Mechanical Maintenance

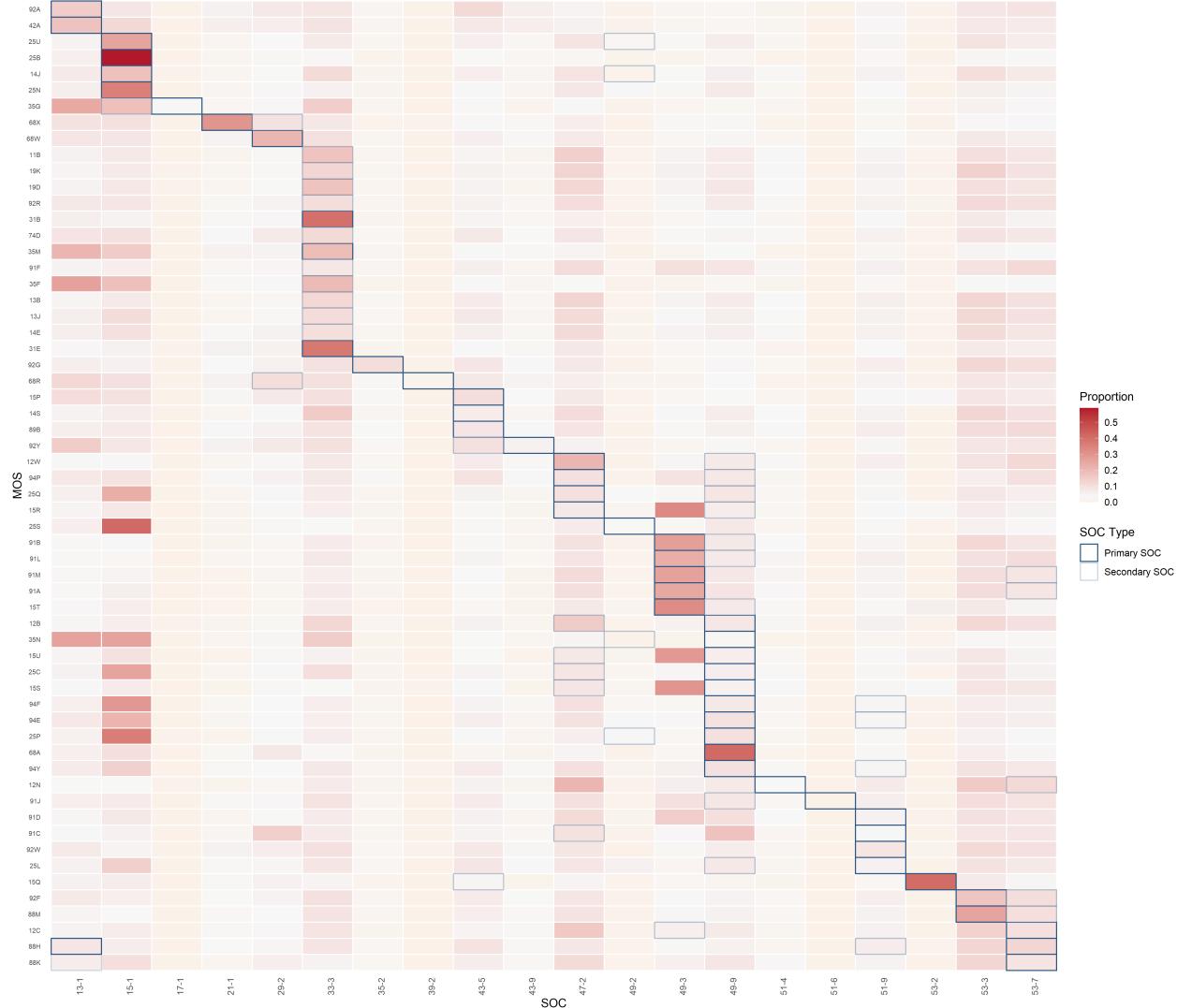


(j) Quartermaster Corps



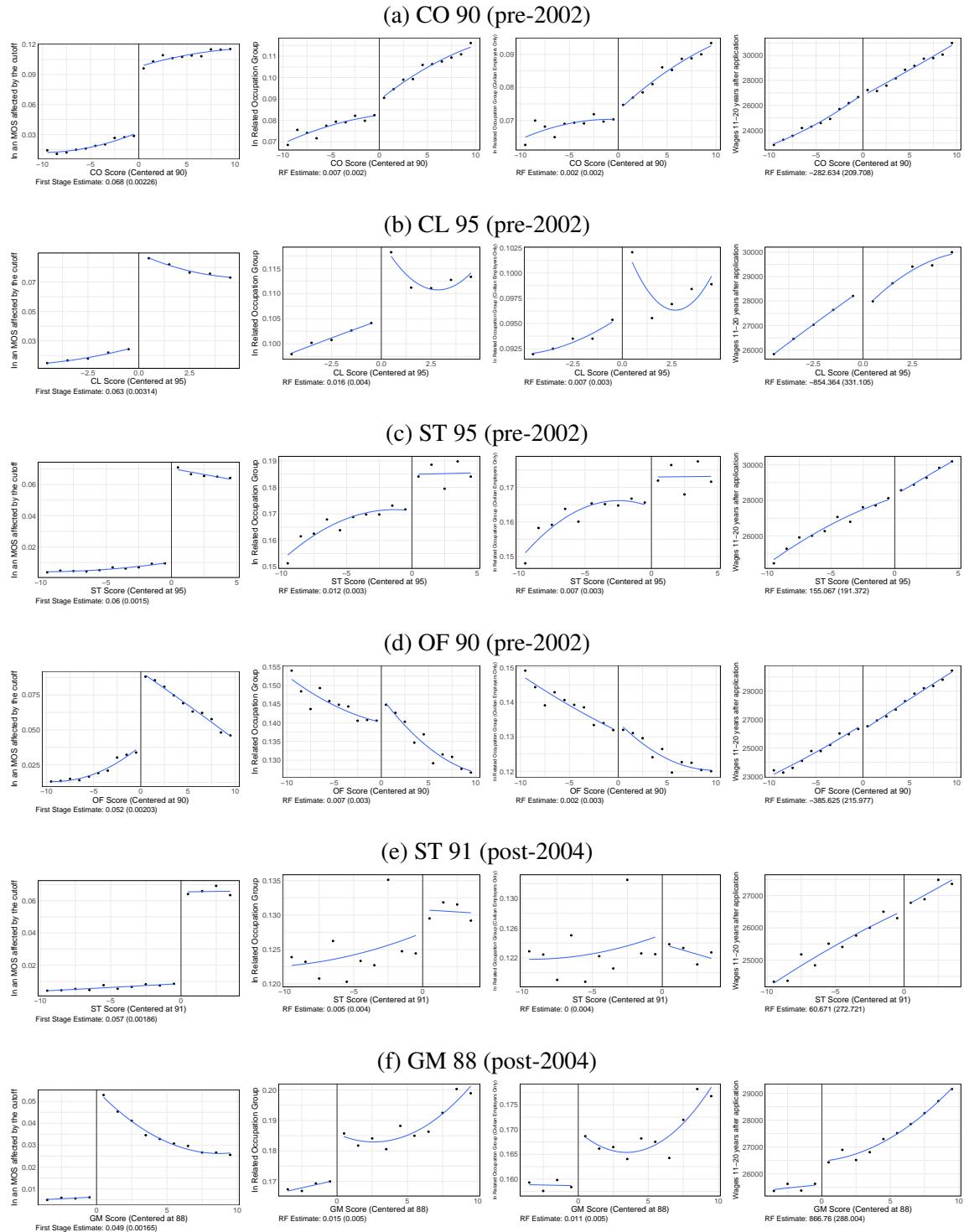
Notes: These plots demonstrate the distribution of occupations 11-20 years after enlistment for the 10 largest Career Management Fields (CMF) in our sample, combining the traditional combat arms CMFs (infantry, field artillery, combat engineers, aviation, and air defense) into one group. Specifically, the outcome is the average share of years 11-20 an individual is in the stated SOC, with missings and weights handled as in our baseline occupational path dependence measure. The top 10 most common long-run SOC codes for each CMF are shown, while all other SOC codes outside the top 10 are grouped together under “Other.” Occupations are listed from largest to smallest, with Military and Other always grouped at the bottom.

Figure F.4: Distribution of Occupations 11-20 Years Post-ASVAB by Entry MOS

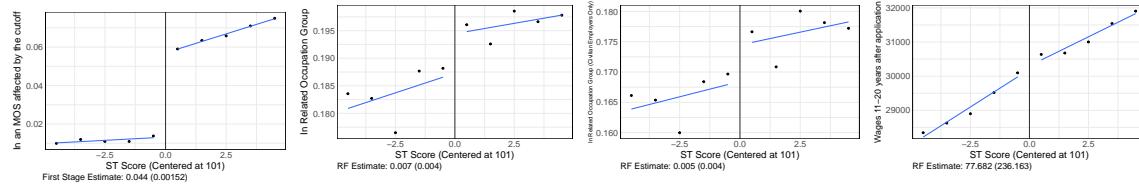


Notes: Each cell represents the proportion of Soldiers within that row's MOS who are in the corresponding SOC code 11-20 years after the ASVAB. Specifically, the outcome is the average share of years 11-20 an individual is in the stated SOC code, with missings and weights handled as in our baseline occupational path dependence measure. Cells outlined in dark blue indicate the primary SOC codes, and those in light blue indicate the secondary SOC codes we have crosswalked to the corresponding MOS. In the interest of keeping the figure down to a reasonable dimension, we limit the set of MOS and SOC codes considered. We have included only SOC codes that are considered primary codes for at least one of the MOS in the sample. We also omit military SOC codes, and we only include MOS that are attached to a cutoff in our baseline analysis. All proportions are normalized to add to one within each row, though we note that the figure looks very similar even without this normalization. The MOS and SOC codes are ordered such that the primary SOC code falls roughly along the diagonal. This is done by sorting SOC codes then sorting MOS by their corresponding primary SOC code.

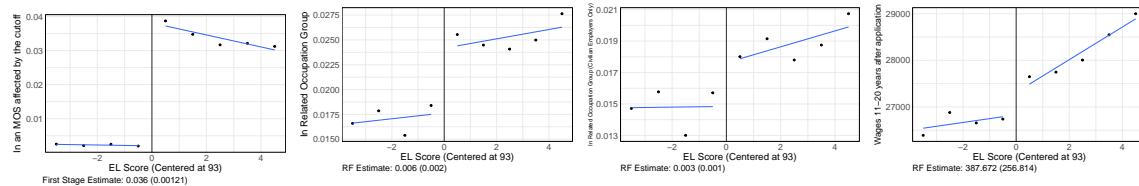
Figure F.5: Cutoff-by-cutoff First Stage and Reduced Form Plots



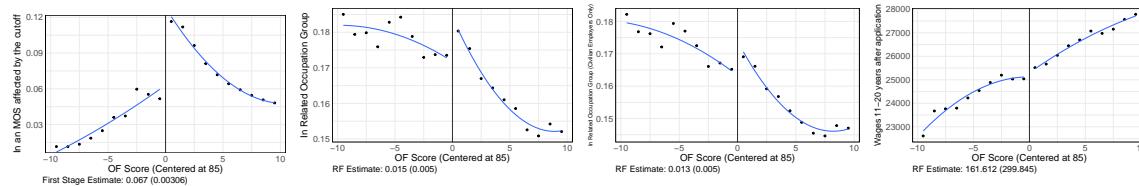
(g) ST 101 (post-2004)



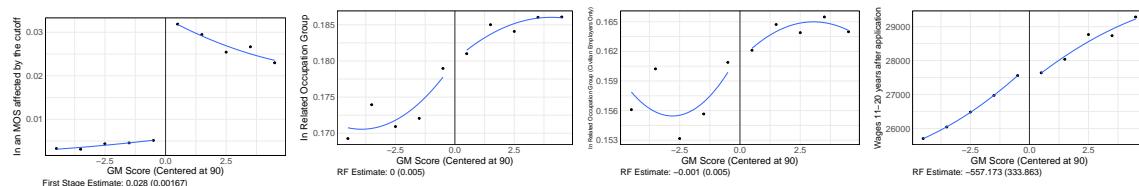
(h) EL 93 (post-2004)



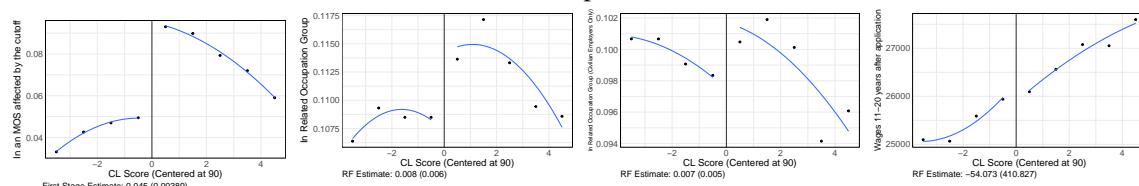
(i) OF 85 (post-2004)



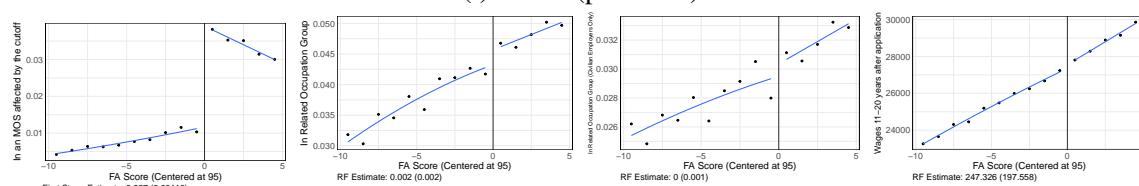
(j) GM 90 (pre-2002)



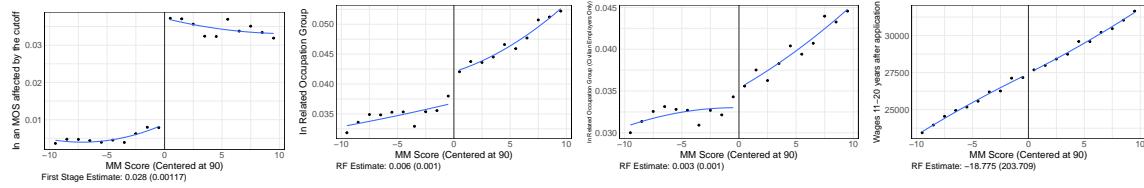
(k) CL 90 (post-2004)



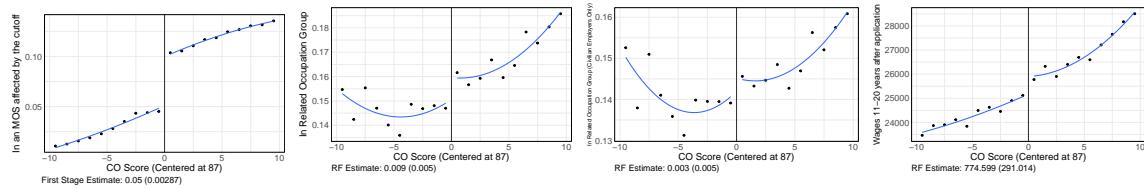
(l) FA 95 (pre-2002)



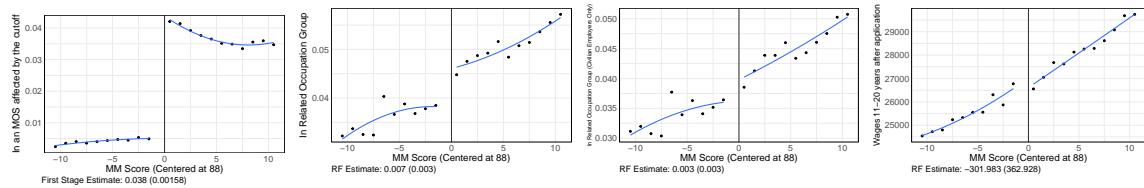
(m) MM 90 (pre-2002)



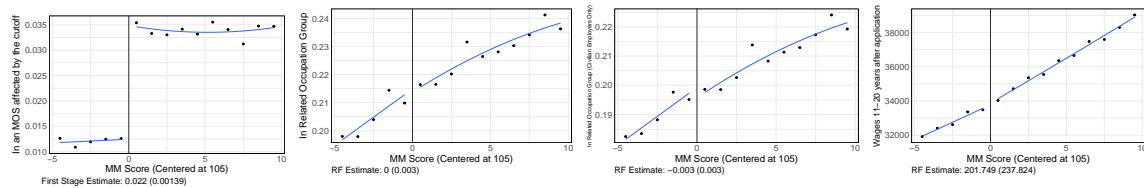
(n) CO 87 (post-2004)



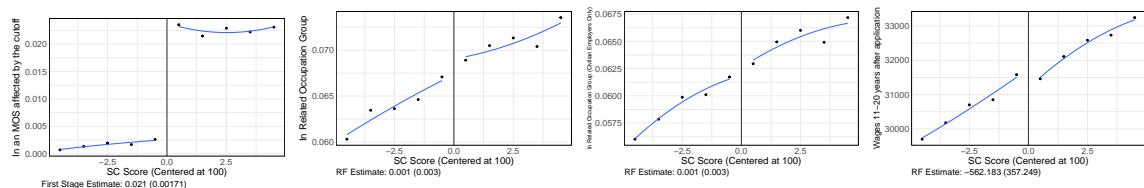
(o) MM 88 (post-2004)



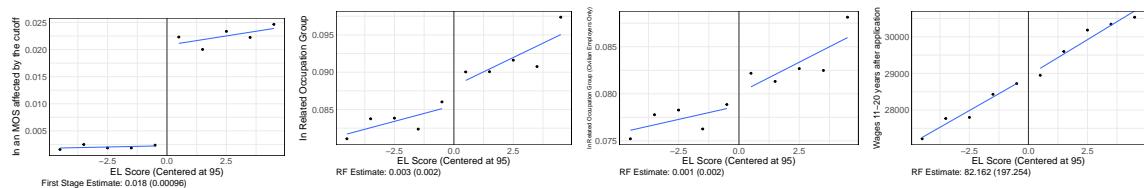
(p) MM 105 (pre-2002)



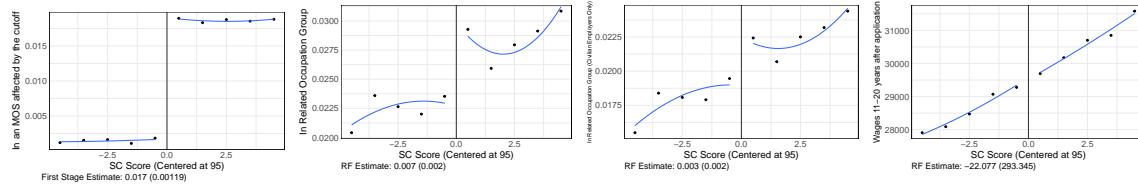
(q) SC 100 (pre-2002)



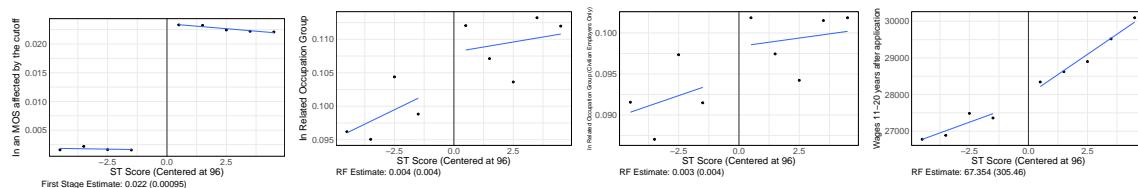
(r) EL 95 (pre-2002)



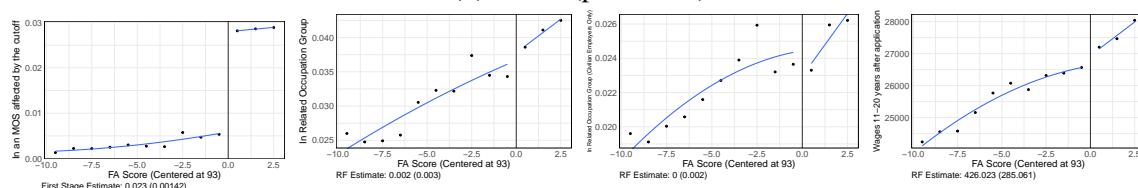
(s) SC 95 (pre-2002)



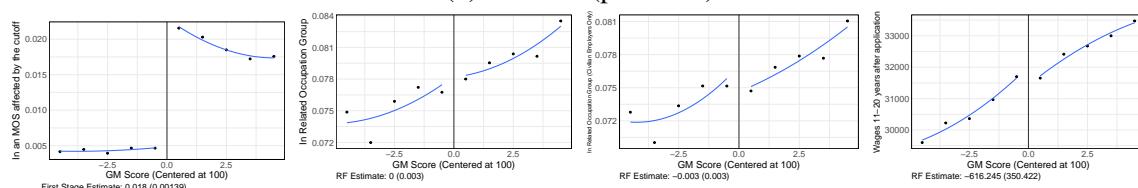
(t) ST 96 (post-2004)



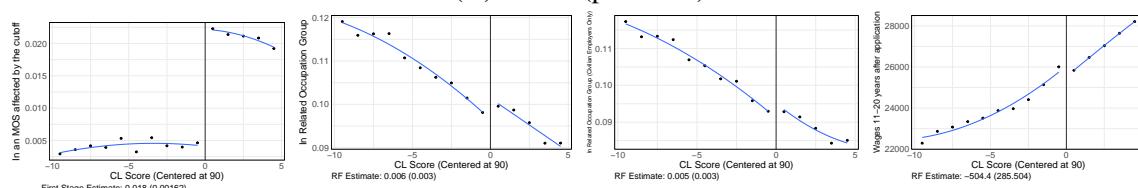
(u) FA 93 (post-2004)



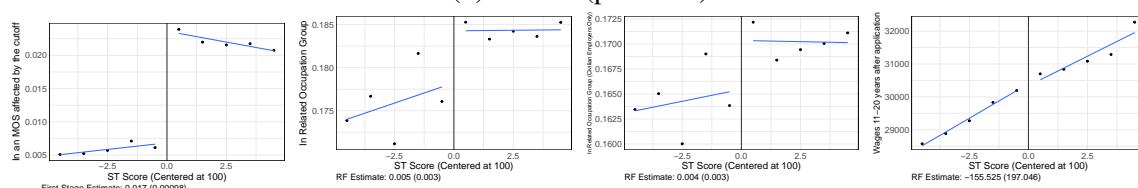
(v) GM 100 (pre-2002)



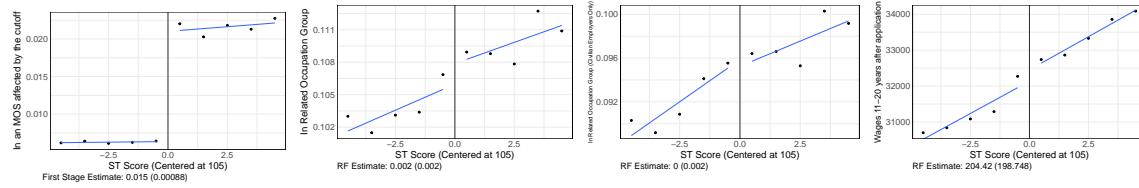
(w) CL 90 (pre-2002)



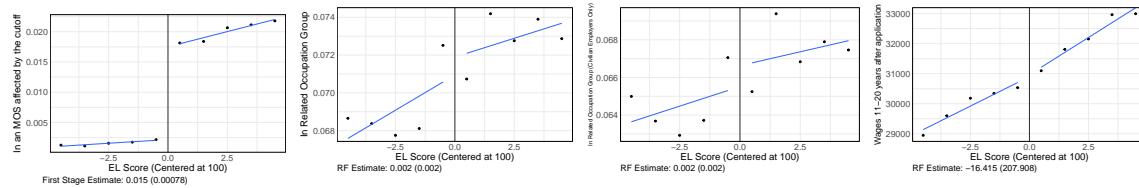
(x) ST 100 (pre-2002)



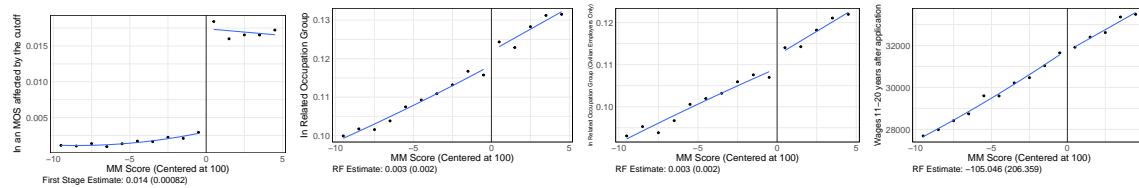
(y) ST 105 (pre-2002)



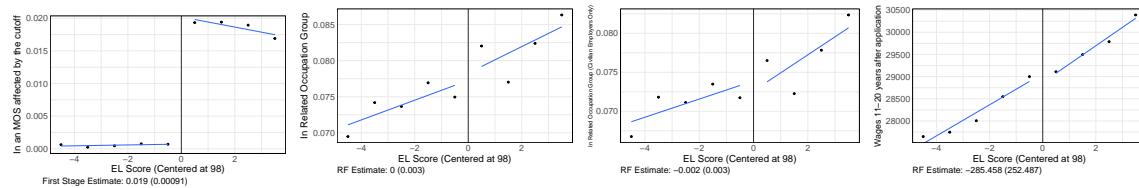
(z) EL 100 (pre-2002)



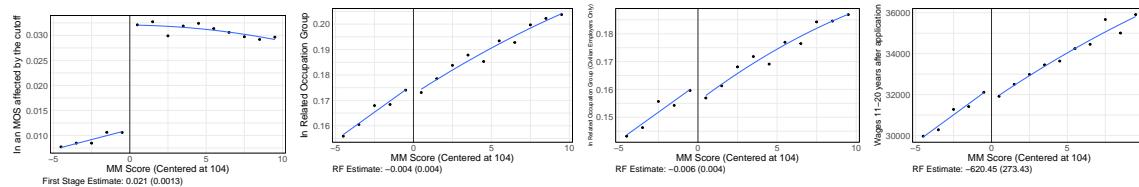
(aa) MM 100 (pre-2002)



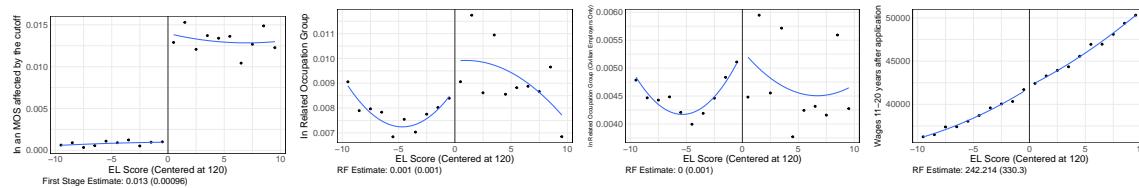
(ab) EL 98 (post-2004)



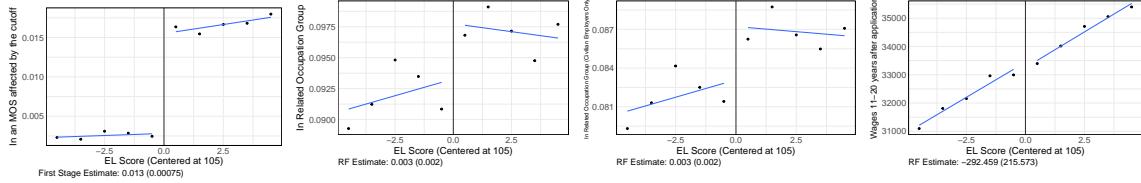
(ac) MM 104 (post-2004)



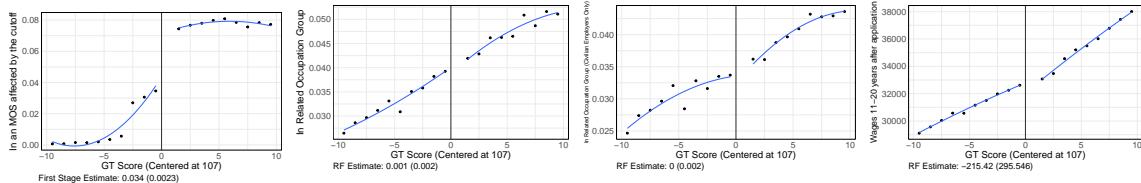
(ad) EL 120 (pre-2002)



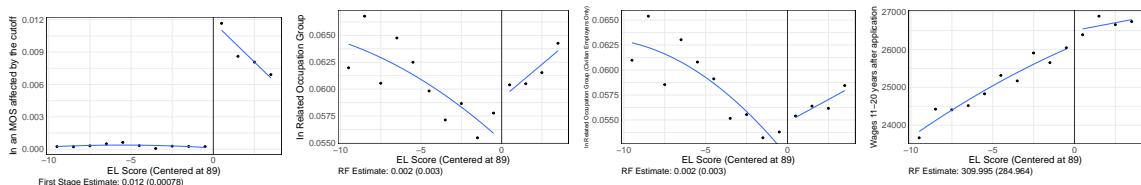
(ae) EL 105 (pre-2002)



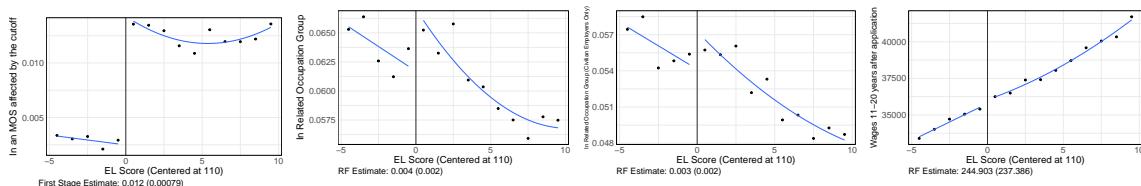
(af) GT 107 (post-2004)



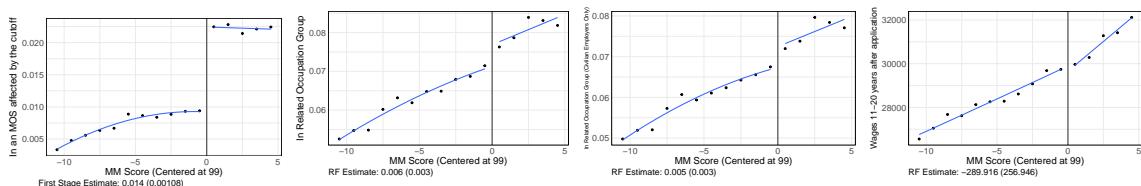
(ag) EL 89 (post-2004)



(ah) EL 110 (pre-2002)

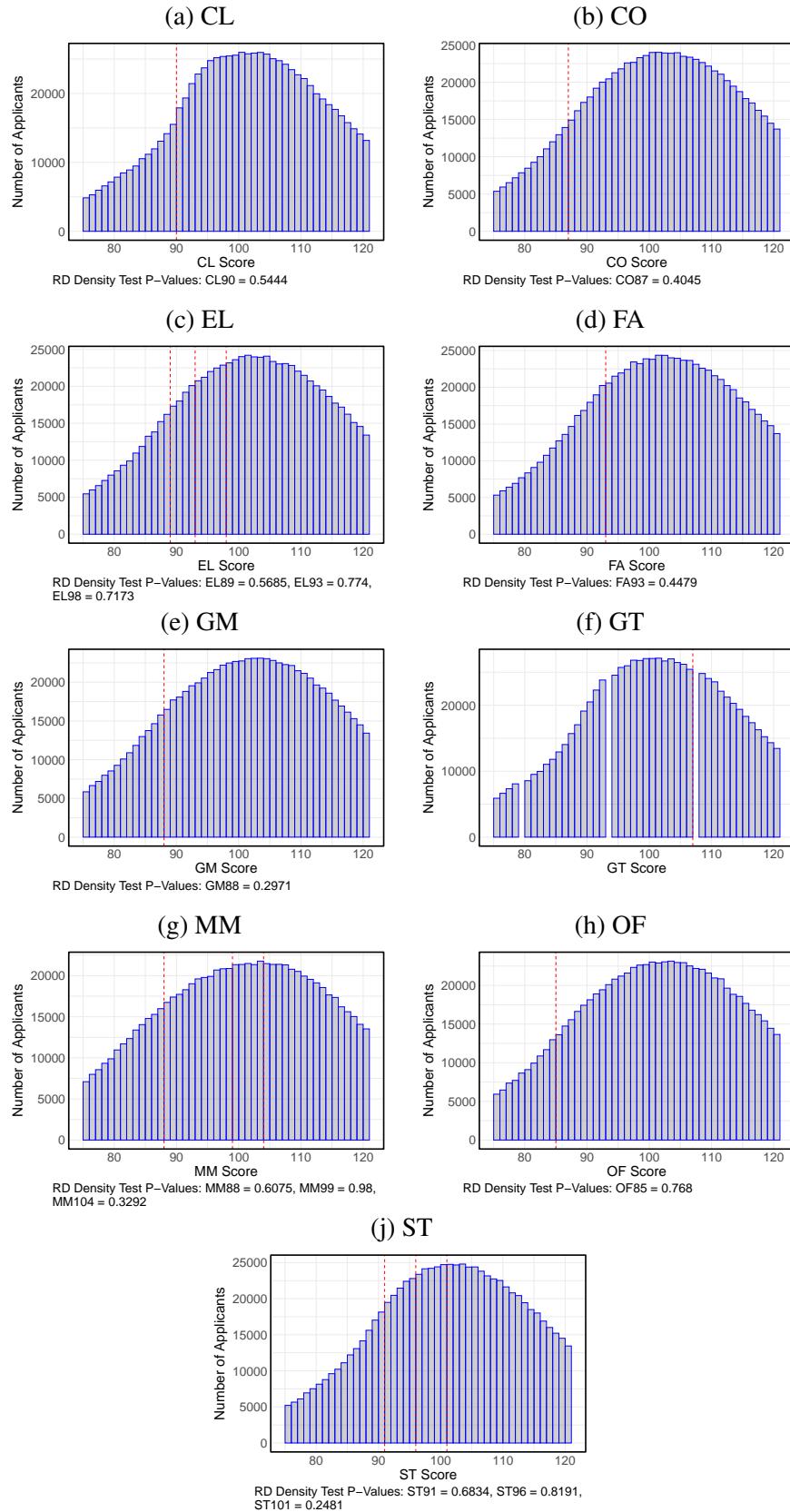


(ai) MM 99 (post-2004)



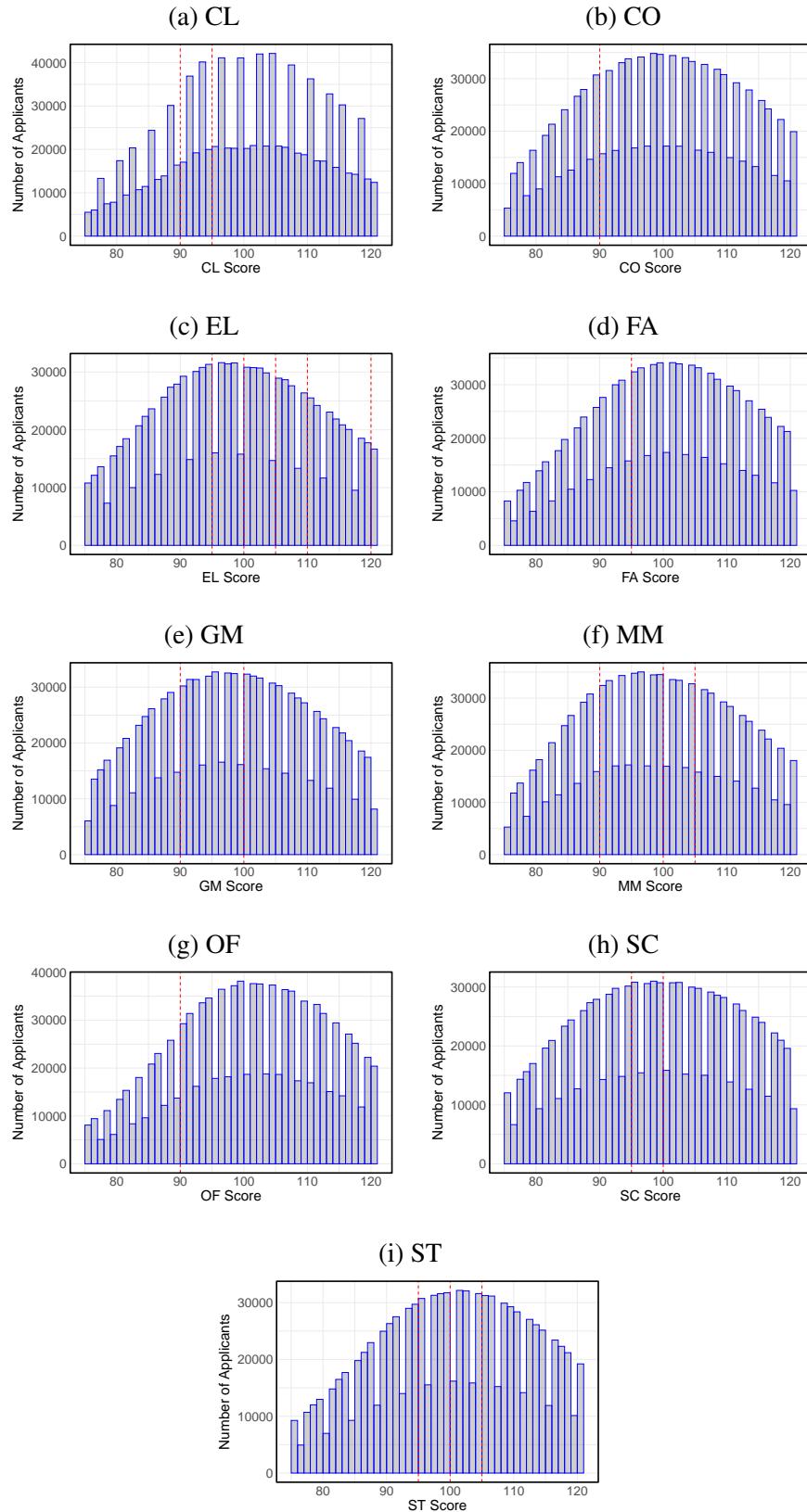
Notes: This figure displays cutoff-specific first stage and reduced form regression discontinuity plots for our key outcomes. Each panel denotes a cutoff. The left-most plot within each panel is the first stage. The next plot displays the reduced form effect of that cutoff on the share of time, 11-20 years after the ASVAB date, that an individual spent in an occupation unlocked by that cutoff. The next plot is identical, except that it only includes time spent in the relevant occupation at a non-Army employer. The final, rightmost plot displays the reduced form impacts on earnings. In each plot, black dots denote binned averages at points in the line-score distribution. In all plots, the support of the running variable and the polynomial used to fit the curve (in blue) are identical within-panel, and reflect the bandwidth and polynomial used in our preferred model. The corresponding first-stage and reduced form point estimates are displayed beneath each figure, along with standard errors that are robust to heteroskedasticity.

Figure F.6: 2004-2014 Density Plots



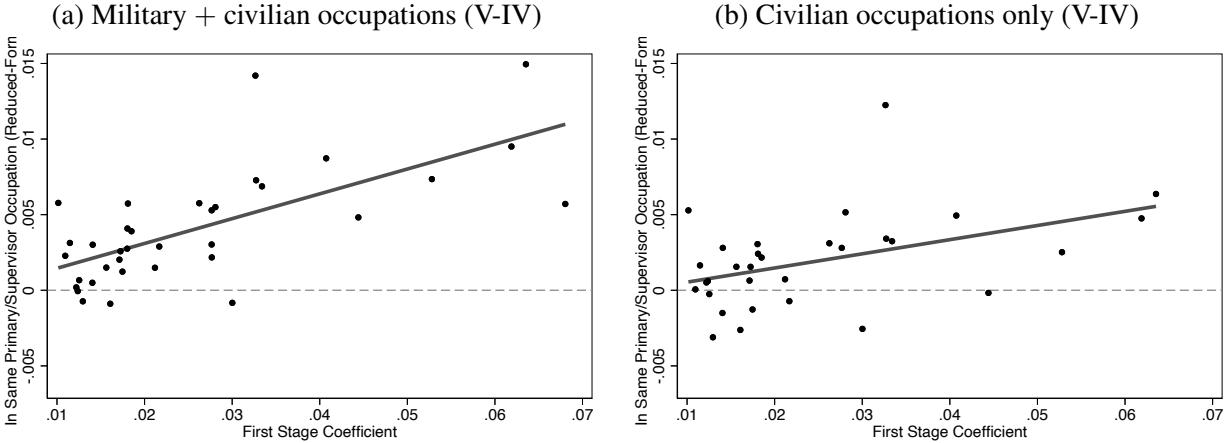
Notes: These figures show density plots for the 2004-2014 line scores. Red lines correspond to a cutoff used in our paper. For the densities that have a non-discrete, non-lumpy running variable, we report the p-values from the Cattaneo et al. (2018) density test using their optimal bandwidth and a local linear functional form.

Figure F.7: 1992-2001 Density Plots



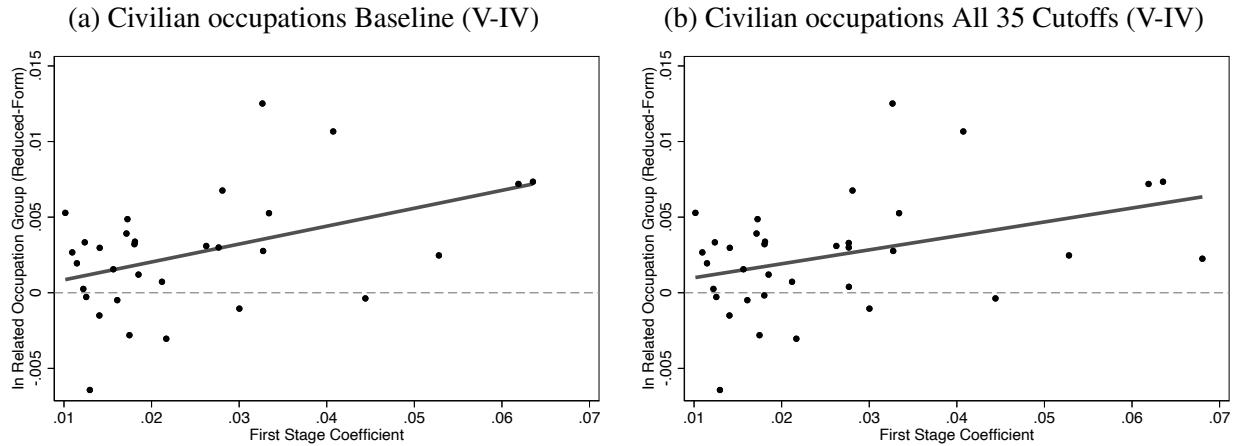
Notes: These figures show density plots for the 1992-2001 line scores. Red lines correspond to a cutoff. Densities for this period are discrete. The lumpy appearance is due to the way that multiple raw test scores are converted, via two conversion tables, to final line scores. Effectively, some support points in the line-score distribution encompass more points in the support of the underlying raw scores, which makes them more likely and results in the increased density.

Figure F.8: Occupational Persistence - Primary and Supervisors Only



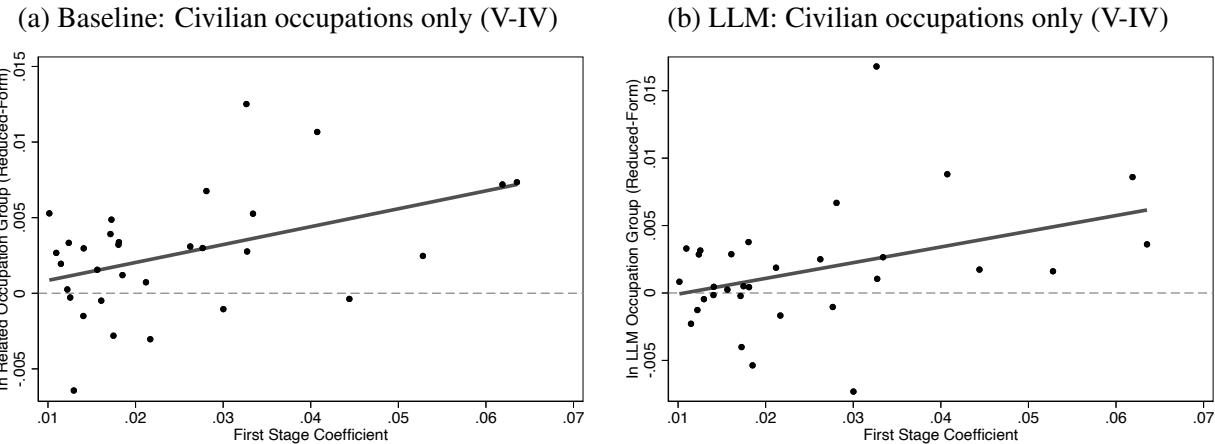
Notes: These figures are analogous to those in Figure 3 except that we now construct the outcome — in the focal occupation group — using only the primary three digit SOC (and corresponding supervisor SOC) in our crosswalk (as opposed to any of the primary, secondary, or corresponding supervisory SOCs). The slope coefficient in panel (a) is 0.164 ($se = 0.026$), while the slope coefficient in panel (b) is 0.094 ($se = 0.031$). Standard errors are calculated using the delta method from the full covariance-matrix of reduced form and first stage estimates.

Figure F.9: Civilian Occupational Persistence - Including Combat and Field Artillery Cutoffs



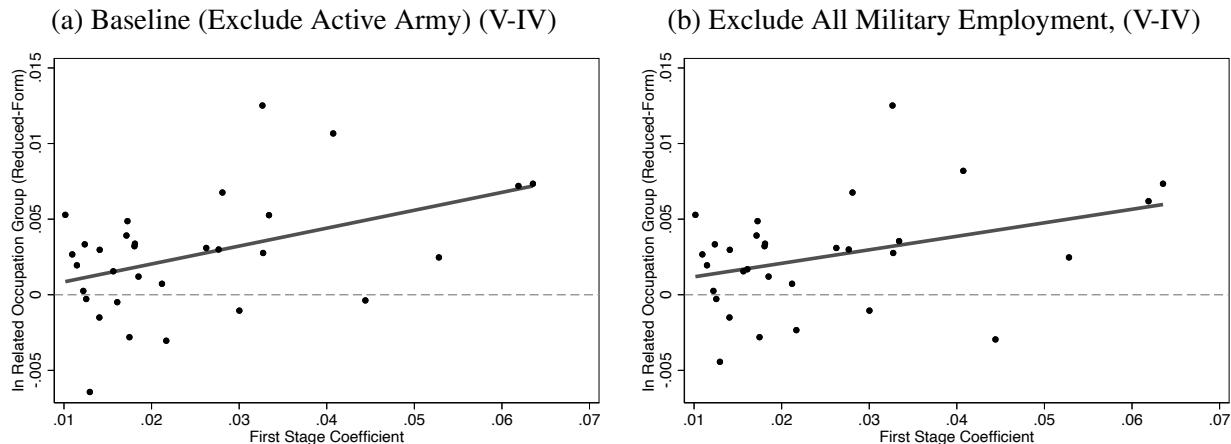
Notes: The left panel reproduces Figure 3 (b). The right panel is similar except we also retain the 4 cutoffs linked to combat and field artillery. For these cutoffs, civilian persistence is primarily based on persistence in secondary jobs since most of the MOS at these cutoff's are assigned a military-specific primary SOC. The slope coefficient in panel (b) is 0.092 ($se = 0.032$). Standard errors are calculated using the delta method from the full covariance-matrix of reduced form and first stage estimates.

Figure F.10: Civilian Occupational Persistence - LLM



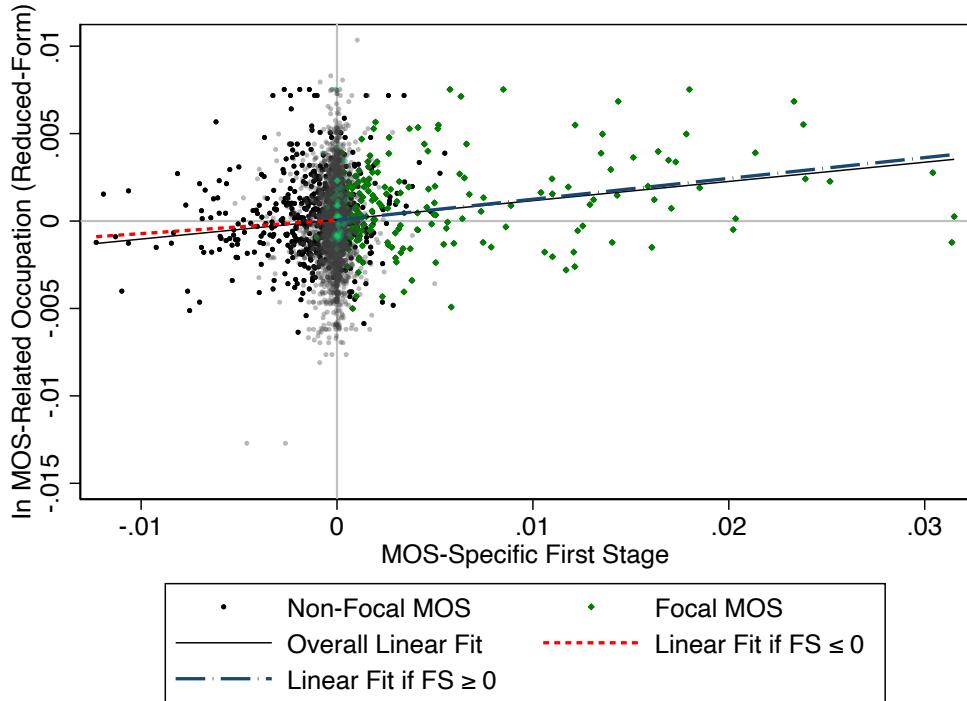
Notes: For ease of comparison, panel (a) reproduces panel (b) of Figure 3; i.e. the cutoff-specific reduced form effects on our preferred measure of occupational path dependence against the first stage effects at that same cutoff for occupations held at non-Army employers. Panel (b) shows the cutoff-specific reduced form effects on being in any occupation linked by the LLM to the corresponding Army occupations affected by the cutoff (held at non-Army employers). The slope in panel (a) is 0.118 and the slope in panel (b) is 0.117.

Figure F.11: Civilian Occupational Persistence - Alternative Definition



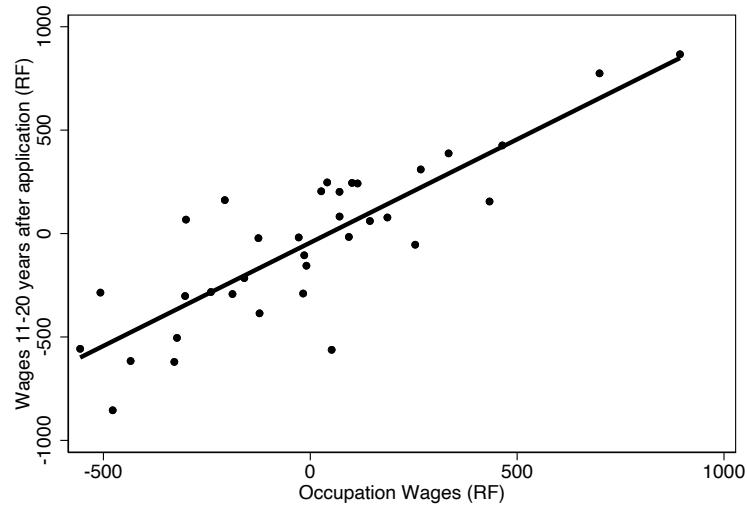
Notes: For ease of comparison, panel (a) reproduces panel (b) of Figure 3; i.e. the cutoff-specific reduced form effects on our preferred measure of occupational path dependence against the first stage effects at that same cutoff for occupations held at non-Army employers. Panel (b) is like (a) but we further define the path dependence outcome as 0 if the person reports any military-linked SOC, so as to also exclude people who moved to other active-duty services or whose main job is in the reserves. The slope in panel (a) is 0.118 (or 0.108 without a constant) and the slope in panel (b) is 0.089 (or 0.098, without a constant).

Figure F.12: MOS-level Civilian Occupational Persistence



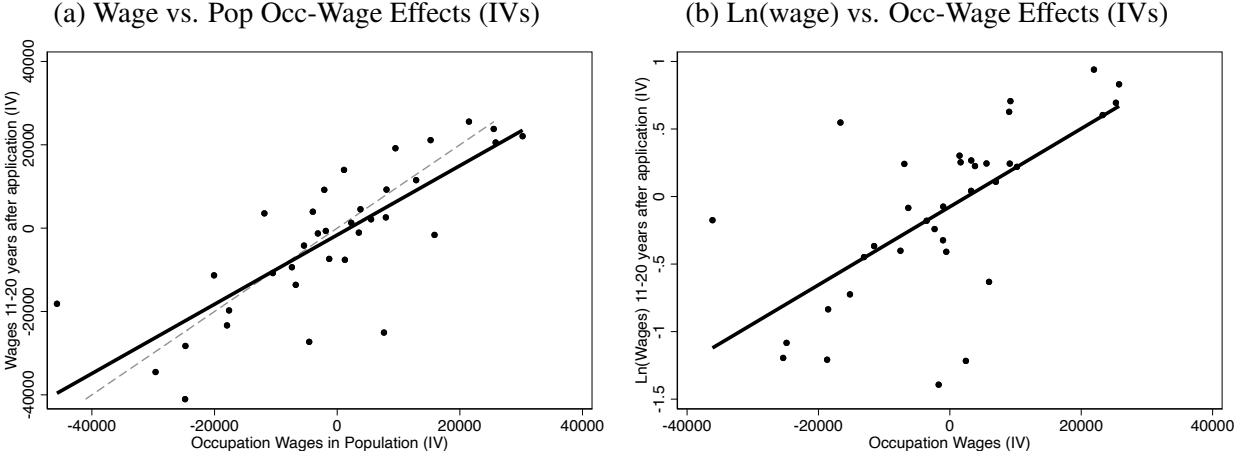
Notes: For every MOS at each of our 35 cutoffs, we estimate the MOS-specific “pseudo” first stage (i.e. the change in probability of being in that MOS that results from crossing the relevant threshold). We also estimate the reduced form effect on being in a related non-Army occupation 11-20 years out (i.e the reduced-form effect on whether you are in any of the MOS-specific cross-walked occupations – primary, secondary or their supervisory counterparts – with those still in the Army receiving 0s). We then scatter all of these reduced form estimates against the corresponding first stages. Green diamonds correspond to MOS that are focal at (i.e. unlocked by) the given cutoff and black dots to the MOS that are not focal at the given cutoff. For both colors, the opaque points have a significant MOS-specific first stage (p -value <0.05) while the lighter/less opaque points have insignificant first stages. The raw overall line of best fit is shown in solid black, the line of best fit on the subset of the sample with a negative first stage is shown in red, and the line of best fit on the subset of the sample with a positive first stage is shown in blue. The corresponding regression estimates are reported in Table F.9.

Figure F.13: Realized Earnings and Long-run Shifts in Average Occupational Earnings (Reduced form vs. Reduced form)



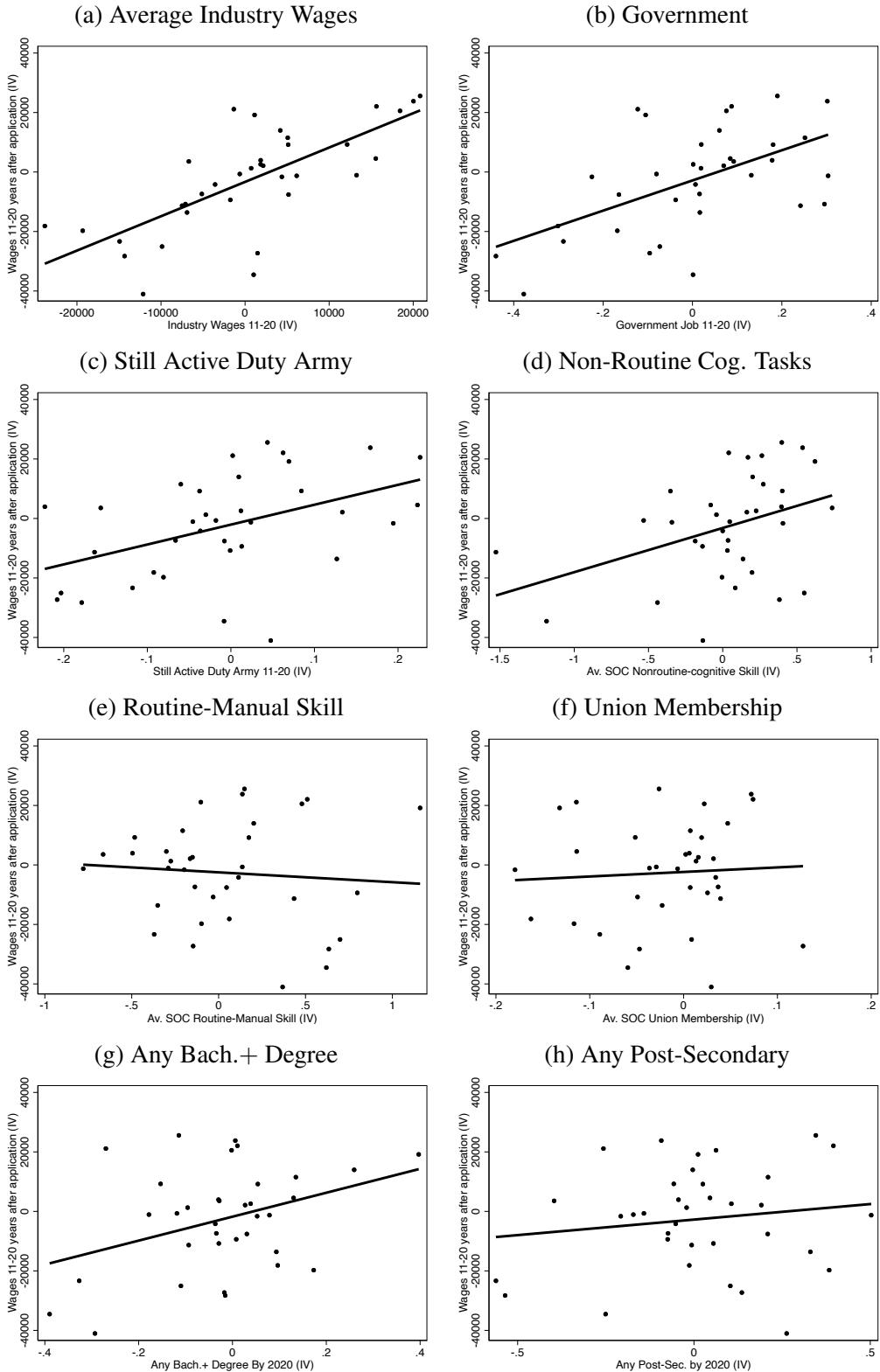
Notes: This figure plots reduced form causal effects on realized earnings against reduced form causal effects on occupation wage premia. Each point is a cutoff. The Y-axis denotes the 11-20 years reduced-form effect of that cutoff on average own wages. The X-axis denotes the 11-20 years reduced form effect of that cutoff on leave-out average occupation wages. Thus, the X-axis directly measures which cutoffs tend to move individuals out of low-paying occupations and into high-paying occupations (as well as the converse) over the long-run. The occupation variable used to construct the causal effects on the X-axes reflect full sample, leave-out average wages based on three-digit SOC codes (treating all military as its own SOC code). The line of best fit is shown in solid black. We also plot the 45-degree line as a dashed, grey line; however, it is not visible since the best fit line falls on top of it. The slope of the best fit line is 0.999 with a heteroskedasticity robust standard error (calculated using the full variance covariance matrix of reduced form estimates) of 0.155 and adjusted R^2 of 0.702. As additional robustness, Appendix Table F.7 also provides corresponding estimates of the slope estimated directly from the micro-data.

Figure F.14: Robustness: Wages vs. Occ-wages



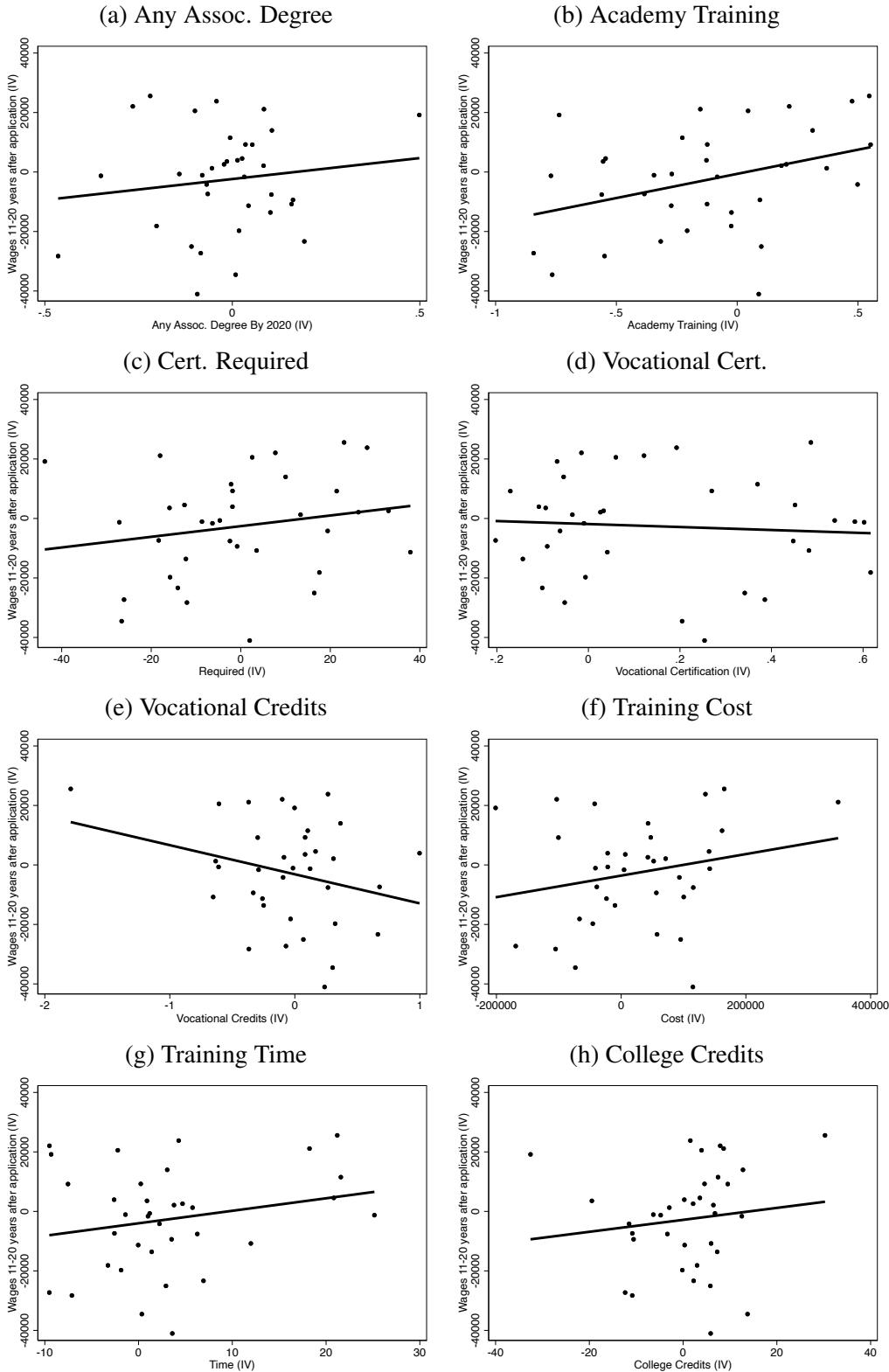
Notes: These figures are similar to panel (a) of figure 5 in the main text, except with some key differences meant to demonstrate robustness. In panel (a) of this figure, the key difference relative to the main text is that the X-axis denotes the year 11-20 IV effect of each cutoff on occupation wages calculated using a 50% random sample pulled from the universe of tax records for those aged 31-44, whereas our baseline uses the in-sample, leave-out average. The slope continues to be indistinguishable from one, and the R^2 continues to be high (0.617). This demonstrates that our finding is robust to using other sensible measures of occupational wage premia. In panel (b) of this figure, the key difference relative to the main text is that the Y-axis gives the causal effect on log wages, whereas our baseline just uses wages. This shows that our results continue to hold even after accounting for extensive margin effects on employment. Since the logarithm is a non-linear transformation, the 45-degree line is no longer a useful benchmark and hence we remove it from the figure in panel (b).

Figure F.15: Relationship between Long-run Earnings and the Covariates in Figure 6



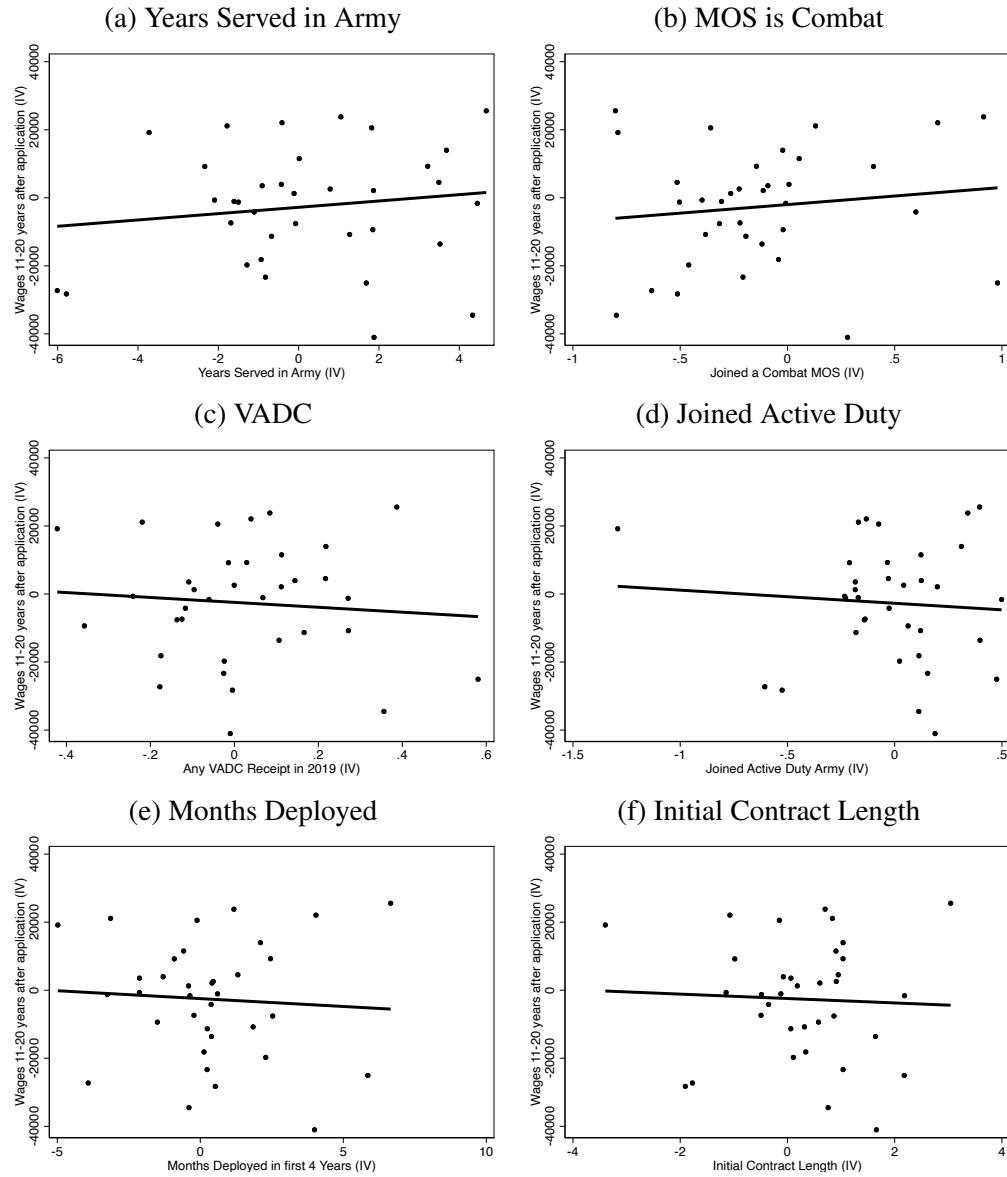
Notes: This figure plots the underlying relationship between the IV effects on 11-20 earnings and the IV effects on each of the covariates in Figure 6. Variable definitions can be found in Appendix B.

Figure F.16: Relationship between Long-run Earnings and the Covariates in Figure 6



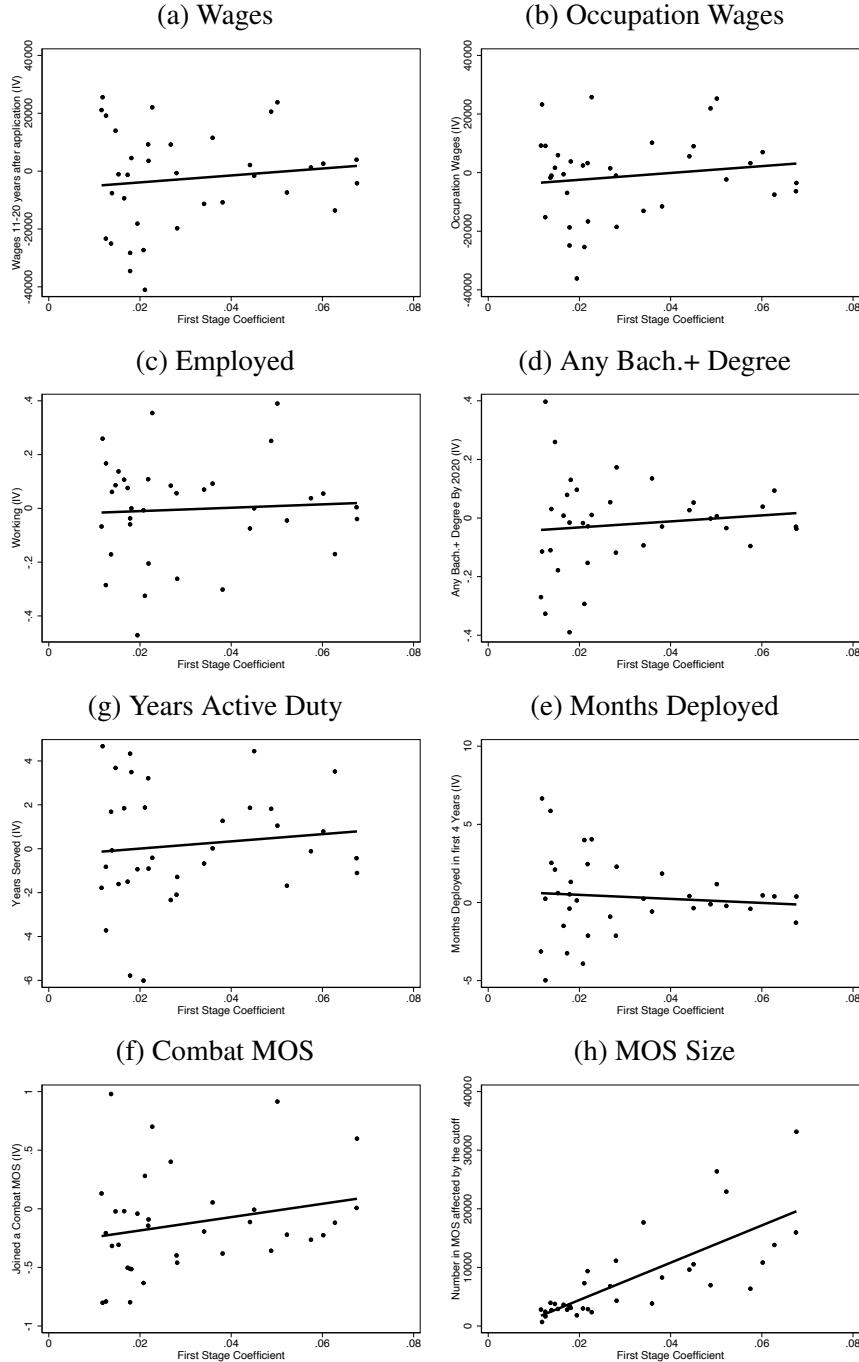
Notes: This figure plots the underlying relationship between the IV effects on 11-20 earnings and the IV effects on each of the covariates in Figure 6. Variable definitions can be found in Appendix B.

Figure F.17: Relationship between Long-run Earnings and the Covariates in Figure 6



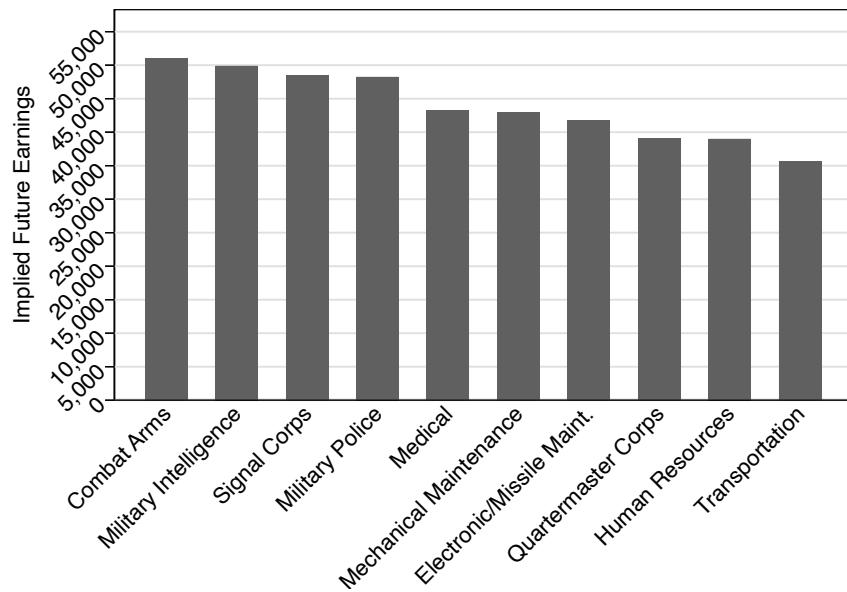
Notes: This figure plots the underlying relationship between the IV effects on 11-20 earnings and the IV effects on each of the covariates in Figure 6. Variable definitions can be found in Appendix B.

Figure F.18: What predicts the magnitude of the first stage?



Notes: This figure plots the causal effect on select long-run outcomes and select military characteristics against the size of the first stage. More precisely, the Y-axis of panels (a) through (d) plot the same cutoff level IV estimates on the indicated outcome used to generate figure 6. The X-axis of these panels gives the corresponding first stage estimate. In all cases, the line is flat and not statistically significant. Thus, it does not appear that compliance is systematically higher when the causal effect on expected earnings is larger (e.g. panel a) or lower when the causal effect on the likelihood of deploying to a dangerous warzone is higher (e.g. panel e). The one covariate that does appear to be systematically correlated with the magnitude of the first stage is the MOS size (panel h): compliance is higher for cutoffs tied to larger military occupations. This is consistent with the institutional details in section 3: applicants can only select an occupation that is unlocked by their performance on the line score if the Army forecasts a vacancy and therefore has a seat in a training unit available at the time of the decision. Larger, more common occupations have more churn and hence have more open training slots during any given period. Thus, this figure is consistent with the idea that variation in the magnitude of the first stage is largely related to availability at the time of the applicant's MEPs appointment.

Figure F.19: MOS Implied Future Earnings Aggregated to the CMF Level



Notes: This figure plots CMF-level weighted averages of our MOS-specific ‘implied future earnings’ measure. We do so for the top 15 CMFs in Table F.1, combining the traditional combat arms CMFs (infantry, field artillery, combat engineers, aviation, and air defense) into one group. Our MOS-specific ‘implied future earnings’ is constructed as follows: first we map each MOS to up to 4 SOC codes (primary, secondary, and up to two supervisors) using our crosswalk. Each SOC is assigned our usual average 11-20 occupation earnings value for that SOC, and then for each MOS we take the simple mean of occupation earnings across all mapped SOCs. This results in an ‘implied future earnings’ value for each MOS. In addition, for those who do not join the Army, we assign predicted occupation wages based on a simple regression of 11-20 occupation wages on AFQT. For this figure, to visually convey patterns more clearly, we collapse the MOS-specific ‘implied future earnings’ to the career-management field (CMF) level (weighting by the number of people in each MOS).

Table F.1: CMF Summary

CMF Number	CMF Title	Number of Soldiers	Percentage of Soldiers
11	Infantry	163,991	18.41
92	Quartermaster Corps	98,833	11.10
91	Mechanical Maintenance	86,123	9.67
25	Signal Corps	68,732	7.72
68	Medical	66,427	7.46
13	Field Artillery	59,041	6.63
19	Armor	46,982	5.27
12	Corps of Engineers	43,395	4.87
35	Military Intelligence	43,265	4.86
15	Aviation	35,123	3.94
88	Transportation	34,858	3.91
31	Military Police	32,368	3.63
42	Adjutant General Corps (Human Resources)	25,376	2.85
14	Air Defense Artillery	21,331	2.39
94	Electronic and Missile Maintenance	14,049	1.58
89	Ammunition	12,529	1.41
74	Chemical Corps	12,006	1.35
36	Finance and Comptroller	3,695	0.41
27	JAG	2,700	0.30
18	Special Forces	2,400	0.27

Notes: This table reports the share of soldiers in the top 20 military CMFs for all those who enlist in our sample, sorted from largest to smallest CMFs. Percentages are expressed as a total of our entire sample, not the top 20 shown in this table.

Table F.2: ASVAB Standard Scores and Descriptions

Abbreviation	Full Name	Description
AR	Arithmetic Reasoning	Measures the ability to solve arithmetic word problems
AI	Auto Information	Assesses knowledge of automobile technology, tools, and terminology
SI	Shop Information	Evaluates knowledge of tools, shop practices, and building materials
EI	Electronic Information	Assesses knowledge of electrical current, circuits, devices, and electronic systems
GS	General Science	Measures knowledge of physical and biological sciences
MC	Mechanical Comprehension	Assesses understanding of mechanical principles and ability to apply them
MK	Mathematics Knowledge	Evaluates knowledge of high school mathematics principles
PC	Paragraph Comprehension	Measures the ability to obtain information from written passages
WK	Word Knowledge	Assesses the ability to understand the meaning of words and how they are used

Notes: The ASVAB is a timed multi-aptitude test developed and maintained by the Department of Defense. The test currently consists of 10 subtests, including the 9 standard scores listed above and the Assembling Objects (AO) subtest, which does not contribute to any of these standard scores. The entire test takes approximately 2-3 hours to complete. In the past, there were additional subtests and scores, such as Numerical Operations (NO), which measured the speed and accuracy of simple arithmetic computations, and Coding Speed (CS), which measured the ability to quickly and accurately assign code numbers to words based on a key. These subtests were removed from the ASVAB in 2002. Additionally, before 2002, the Auto Information (AI) and Shop Information (SI) scores were combined into a single score called Auto & Shop Information (AS). In 2002, the AS score was split into the separate AI and SI scores.

Table F.3: Army ASVAB Line Scores

Abbreviation	Full Name	Description
CL	Clerical	Assesses abilities in administrative tasks, data management, and attention to detail; related to roles such as human resources specialists and administrative assistants
CO	Combat Operations	Evaluates skills needed for combat roles, such as situational awareness, decision-making under pressure, and physical fitness
EL	Electronics	Measures knowledge of electronic systems, circuits, and equipment maintenance, necessary for jobs such as avionics technicians and computer network specialists
FA	Field Artillery	Assesses abilities relevant to field artillery roles, including mathematical reasoning, spatial visualization, and mechanical comprehension
GM	General Maintenance	Evaluates skills in troubleshooting, repairing, and maintaining various equipment, which are necessary for roles such as general purpose vehicle mechanics and construction equipment repairers
GT	General Technical	Measures overall technical aptitude, problem-solving skills, and the ability to learn and apply new information; related to many technical jobs
MM	Mechanical Maintenance	Assesses knowledge of mechanical systems, tools, and equipment maintenance, important for roles such as wheeled vehicle mechanics
OF	Operators and Food	Evaluates skills relevant to equipment operation and food service roles, such as attention to detail, multi-tasking, and following procedures
SC	Surveillance and Communications	Measures abilities related to gathering and communicating information, necessary for jobs like intelligence analysts and signal support systems specialists
ST	Skilled Technical	Assesses overall technical knowledge and skills needed for various specialized technical roles, such as medical laboratory specialists and geospatial intelligence imagery analysts

Notes: This table provides descriptions of the different line scores used to determine Army occupational eligibility. These line scores are built from sub-tests contained in the ASVAB (see Appendix Table F.2). Separately, the AFQT score, which is intended to measure general cognitive ability, is also constructed from 4 of the underlying ASVAB sub-tests. However, the AFQT score is used to determine eligibility to join the Army and does not otherwise impact eligibility for any specific occupation.

Table F.4: Summary of Additional Variables and Sources

Variable Group	Variables	Years	Primary Data Sources
Tasks Characteristics	Non-routine tasks	2015	O*Net 20.1, ACS 2015 (IPUMS)
	Routine tasks		
Union Membership	Union member indicator	2015	CPS ASEC 2015
Education	Bachelors+	1992–2020	National Student Clearinghouse (NSC)
	Bachelors		
	College attendance		
	Associates		
Military Service	Years served	1992–2024	Army Personnel records
	MOS type		
	Active duty status		
	Deployment months		
	Contract length		
VA Disability	VADC indicator	1992–2020	Army Personnel, VA records
Training	AIT courses	2022	Army cost data, Personnel records
	Training weeks		
	Cost estimates		
Academic Credits	College credits	2023	American Council on Education (ACE) credit recommendations
	Vocational credits		
Certification Requirements	Certification	2022-2024	Primarily BLS, Supplemented with data
	Vocational training		and reports from SOLI, Institute for
	Academy training		Justice, and NCSL
	Higher ed requirements		

Notes: This table provides an overview of the additional data sources used in Figure 6. More detailed information on each data source is provided in Appendix B.2.

Table F.5: Balance Table - Time of Application Covariates

	Age	Male	White	Black	Hisp.	In HS	<HS	HS Dip.	Some College+
co90_pre	-0.018 (0.029)	-0.002 (0.004)	-0.004 (0.004)	0.000 (0.004)	0.001 (0.003)	-0.001 (0.004)	-0.002 (0.003)	0.005 (0.005)	-0.002 (0.002)
of85_post	0.023 (0.039)	0.007 (0.006)	-0.005 (0.006)	0.006 (0.006)	0.000 (0.005)	-0.004 (0.005)	0.002 (0.004)	0.008 (0.006)	-0.006 (0.002)**
cl95_pre	0.028 (0.046)	-0.008 (0.006)	-0.002 (0.007)	0.005 (0.007)	-0.002 (0.005)	-0.009 (0.006)	-0.008 (0.005)	0.013 (0.005)	0.004 (0.003)
st95_pre	-0.016 (0.026)	-0.001 (0.004)	-0.002 (0.004)	0.003 (0.004)	-0.002 (0.003)	0.001 (0.004)	0.000 (0.003)	-0.003 (0.004)	0.002 (0.002)
st91_post	-0.005 (0.034)	0.003 (0.005)	0.000 (0.005)	0.001 (0.005)	-0.005 (0.004)	0.002 (0.004)	0.005 (0.004)	-0.007 (0.005)	0.000 (0.002)
of90_pre	-0.066 (0.030)**	-0.007 (0.005)	0.006 (0.005)	-0.002 (0.005)	-0.004 (0.003)	0.005 (0.004)	-0.003 (0.003)	0.002 (0.005)	-0.003 (0.002)*
co87_post	-0.095 (0.038)**	0.008 (0.005)	-0.002 (0.005)	0.001 (0.005)	0.006 (0.004)	0.001 (0.005)	-0.002 (0.004)	0.003 (0.006)	-0.002 (0.002)
gm88_post	-0.035 (0.037)	0.004 (0.005)	0.003 (0.005)	0.001 (0.005)	-0.001 (0.004)	0.008 (0.004)*	-0.006 (0.004)	-0.005 (0.005)	0.003 (0.002)
cl90_post	0.112 (0.053)**	-0.004 (0.007)	-0.006 (0.008)	0.007 (0.007)	0.007 (0.006)	-0.001 (0.006)	-0.002 (0.006)	0.001 (0.008)	0.002 (0.003)
st101_post	-0.065 (0.028)**	-0.002 (0.003)	0.002 (0.004)	-0.004 (0.003)	0.003 (0.003)	0.005 (0.003)	-0.005 (0.003)	-0.001 (0.004)	0.001 (0.002)
mm88_post	0.022 (0.045)	-0.005 (0.006)	-0.010 (0.006)	0.008 (0.006)	0.000 (0.005)	0.004 (0.006)	0.002 (0.005)	-0.004 (0.007)	-0.001 (0.003)
el93_post	0.005 (0.032)	0.000 (0.004)	0.002 (0.005)	-0.006 (0.004)	0.002 (0.004)	0.002 (0.004)	0.002 (0.004)	-0.002 (0.005)	-0.001 (0.002)
gt107_post	0.003 (0.034)	0.005 (0.004)	0.011 (0.005)**	0.000 (0.003)	-0.003 (0.003)	0.005 (0.003)	-0.004 (0.004)	-0.001 (0.005)	0.000 (0.003)
gm90_pre	-0.008 (0.045)	-0.014 (0.007)**	-0.020 (0.007)***	0.013 (0.007)*	0.004 (0.005)	0.002 (0.007)	0.001 (0.005)	-0.001 (0.007)	-0.002 (0.003)
mm90_pre	0.010 (0.028)	-0.002 (0.004)	-0.005 (0.004)	0.002 (0.004)	0.003 (0.003)	-0.001 (0.004)	0.000 (0.003)	-0.003 (0.004)	0.004 (0.002)**
fa95_pre	-0.008 (0.028)	0.004 (0.004)	0.005 (0.004)	0.000 (0.004)	-0.002 (0.003)	-0.005 (0.004)	-0.003 (0.003)	0.008 (0.004)*	0.000 (0.002)
fa93_post	0.000 (0.036)	-0.003 (0.005)	0.000 (0.005)	-0.001 (0.005)	-0.003 (0.004)	-0.002 (0.004)	-0.004 (0.004)	0.008 (0.005)	-0.002 (0.002)
st96_post	0.018 (0.037)	0.001 (0.005)	0.007 (0.006)	-0.001 (0.005)	-0.007 (0.004)*	0.000 (0.005)	0.000 (0.004)	-0.002 (0.006)	0.002 (0.002)
mm105_pre	0.022 (0.028)	-0.004 (0.003)	-0.011 (0.004)**	0.010 (0.003)	0.001 (0.002)	-0.003 (0.004)	0.000 (0.003)	0.002 (0.004)	0.000 (0.002)
mm104_post	-0.028 (0.031)	0.003 (0.003)	0.002 (0.004)	0.002 (0.003)	0.000 (0.003)	0.003 (0.004)	-0.001 (0.003)	0.000 (0.004)	-0.002 (0.002)
sc100_pre	0.041 (0.044)	-0.010 (0.006)*	0.000 (0.007)	0.001 (0.006)	0.002 (0.004)	0.005 (0.006)	0.002 (0.005)	-0.007 (0.007)	0.000 (0.003)
el98_post	-0.043 (0.030)	0.004 (0.004)	0.003 (0.004)	0.000 (0.004)	0.000 (0.003)	0.000 (0.004)	0.005 (0.003)	-0.007 (0.004)	0.002 (0.002)
el95_pre	0.027 (0.026)	-0.008 (0.004)**	-0.002 (0.004)	0.007 (0.004)*	-0.002 (0.003)	-0.002 (0.004)	-0.004 (0.003)	0.005 (0.004)	0.001 (0.002)
cl90_pre	-0.028 (0.042)	0.002 (0.006)	-0.007 (0.006)	0.003 (0.006)	0.004 (0.004)	0.010 (0.006)	-0.003 (0.005)	-0.007 (0.006)	0.000 (0.002)
gm100_pre	0.024 (0.043)	0.002 (0.006)	-0.013 (0.007)*	0.006 (0.006)	0.003 (0.004)	0.007 (0.006)	-0.006 (0.005)	0.002 (0.007)	-0.004 (0.003)
sc95_pre	-0.033 (0.038)	0.003 (0.006)	0.009 (0.006)	-0.009 (0.005)*	0.002 (0.004)	0.010 (0.006)*	-0.005 (0.004)	-0.008 (0.006)	0.002 (0.002)
st100_pre	-0.016 (0.025)	-0.002 (0.004)	0.004 (0.004)	-0.007 (0.003)*	0.002 (0.002)	0.002 (0.004)	-0.001 (0.003)	-0.007 (0.004)*	-0.004 (0.002)**
el100_pre	0.027 (0.026)	-0.003 (0.004)	-0.004 (0.004)	0.004 (0.004)	0.001 (0.003)	0.006 (0.004)	0.000 (0.003)	-0.007 (0.004)*	0.001 (0.002)
st105_pre	-0.041 (0.024)*	0.001 (0.003)	0.006 (0.004)	-0.003 (0.003)	-0.001 (0.002)	0.003 (0.004)	-0.002 (0.003)	-0.001 (0.004)	-0.001 (0.002)
mm100_pre	-0.018 (0.025)	-0.007 (0.003)**	0.004 (0.004)	0.001 (0.003)	-0.006 (0.002)	0.001 (0.004)	0.000 (0.003)	-0.002 (0.004)	0.001 (0.002)
mm99_post	0.000 (0.030)	0.000 (0.004)	-0.001 (0.004)	-0.004 (0.003)	0.002 (0.003)	-0.004 (0.004)	0.001 (0.003)	0.004 (0.004)	-0.001 (0.002)
el120_pre	0.051 (0.036)	-0.002 (0.004)	0.001 (0.004)	0.000 (0.003)	0.001 (0.002)	0.001 (0.005)	-0.001 (0.003)	-0.008 (0.005)	0.009 (0.003)**
el105_pre	-0.014 (0.026)	0.000 (0.003)	-0.006 (0.004)*	0.000 (0.003)	0.004 (0.002)	0.008 (0.004)**	-0.001 (0.003)	-0.008 (0.004)**	0.002 (0.002)
el89_post	0.016 (0.036)	0.003 (0.005)	0.001 (0.005)	0.000 (0.005)	-0.003 (0.004)	0.006 (0.004)	-0.001 (0.004)	-0.006 (0.005)	0.001 (0.002)
el110_pre	0.077 (0.027)***	-0.001 (0.003)	0.000 (0.004)	0.000 (0.003)	-0.002 (0.002)	-0.008 (0.004)**	0.000 (0.003)	0.006 (0.004)	0.002 (0.002)

P-value on test that all covariates \times cutoffs are 0: 0.173

Notes: This table reports the covariate balance (reduced-form) results for our 35 cutoffs \times 9 balance covariates, measured at the time of application, from the Army applicant data. Table F.6 reports the covariate balance (reduced-form) results for our 35 cutoffs \times 5 balance covariates measured in the year prior to application from the IRS data. We also report the p-value on a test that all 35 cutoffs \times 14 balance covariates across both tables is 0 ($P = 0.173$).

* $p < .10$, ** $p < .05$, *** $p < .01$

Table F.6: Balance Table - Year Prior to Application Covariates

	Filed Taxes	Earnings	Married	Post-Secondary Attendance	Employment
co90_pre	-0.006 (0.006)	-23.434 (168.646)	-0.003 (0.004)	-0.006 (0.008)	-0.001 (0.008)
of85_post	0.004 (0.006)	66.054 (93.258)	0.002 (0.005)	-0.004 (0.005)	-0.001 (0.005)
cl95_pre	-0.012 (0.010)	217.962 (288.740)	-0.008 (0.007)	0.003 (0.011)	-0.007 (0.012)
st95_pre	0.004 (0.006)	-251.166 (159.577)	-0.001 (0.004)	-0.001 (0.006)	0.000 (0.007)
st91_post	-0.004 (0.005)	-140.306 (85.282)*	0.000 (0.004)	0.001 (0.004)	-0.006 (0.005)
of90_pre	0.005 (0.007)	29.533 (169.622)	-0.013 (0.005)***	0.002 (0.008)	0.008 (0.008)
co87_post	-0.014 (0.006)**	-77.588 (91.942)	-0.003 (0.005)	-0.004 (0.004)	-0.005 (0.005)
gm88_post	-0.003 (0.005)	-155.864 (88.262)*	-0.005 (0.004)	0.000 (0.004)	-0.012 (0.005)**
cl90_post	0.006 (0.008)	-2.345 (132.343)	0.002 (0.007)	0.001 (0.006)	0.002 (0.007)
st101_post	-0.002 (0.004)	-15.235 (73.403)	-0.003 (0.003)	0.000 (0.003)	-0.002 (0.003)
mm88_post	-0.006 (0.007)	130.964 (106.451)	-0.003 (0.006)	0.000 (0.005)	0.003 (0.006)
el93_post	-0.005 (0.005)	-20.410 (78.174)	0.001 (0.004)	-0.001 (0.004)	-0.001 (0.004)
gt107_post	0.003 (0.005)	-70.345 (93.826)	-0.001 (0.004)	0.006 (0.004)	-0.001 (0.004)
gm90_pre	0.007 (0.010)	-8.817 (274.183)	-0.010 (0.007)	0.000 (0.012)	0.000 (0.012)
mm90_pre	-0.005 (0.006)	-283.746 (162.703)*	-0.001 (0.004)	-0.017 (0.007)***	-0.016 (0.008)**
fa95_pre	0.009 (0.006)	-14.819 (174.647)	-0.005 (0.004)	-0.003 (0.007)	0.004 (0.007)
fa93_post	0.008 (0.005)	-22.856 (88.494)	0.007 (0.004)	-0.001 (0.004)	-0.004 (0.005)
st96_post	0.006 (0.006)	26.691 (93.880)	-0.002 (0.005)	0.004 (0.004)	-0.008 (0.005)
mm105_pre	-0.005 (0.006)	-45.102 (189.024)	0.001 (0.004)	-0.004 (0.008)	-0.008 (0.007)
mm104_post	0.000 (0.004)	-193.924 (81.647)**	-0.007 (0.003)**	-0.001 (0.004)	-0.001 (0.004)
sc100_pre	-0.028 (0.009)***	-116.621 (279.753)	0.009 (0.007)	-0.003 (0.012)	0.000 (0.011)
el98_post	-0.005 (0.004)	-62.481 (77.096)	-0.002 (0.004)	-0.001 (0.003)	-0.005 (0.004)
el95_pre	0.002 (0.006)	-228.868 (165.364)	-0.003 (0.004)	-0.001 (0.007)	-0.006 (0.007)
cl90_pre	0.010 (0.009)	17.793 (246.231)	0.010 (0.007)	-0.020 (0.009)***	-0.002 (0.011)
gm100_pre	-0.019 (0.009)**	-263.830 (275.690)	0.013 (0.007)**	-0.006 (0.012)	-0.015 (0.011)
sc95_pre	0.001 (0.008)	-472.526 (227.669)**	-0.005 (0.006)	0.008 (0.010)	-0.007 (0.010)
st100_pre	0.006 (0.005)	-44.998 (164.649)	0.006 (0.004)	0.001 (0.007)	-0.007 (0.007)
el100_pre	-0.004 (0.006)	-19.111 (173.343)	0.005 (0.004)	-0.007 (0.007)	-0.005 (0.007)
st105_pre	-0.011 (0.005)**	121.362 (161.313)	0.005 (0.004)	0.002 (0.007)	-0.010 (0.006)*
mm100_pre	-0.004 (0.005)	-76.618 (161.953)	-0.003 (0.004)	0.008 (0.007)	-0.003 (0.007)
mm99_post	0.003 (0.004)	-48.061 (76.557)	0.001 (0.003)	-0.002 (0.004)	-0.001 (0.004)
el120_pre	-0.008 (0.007)	-48.436 (274.991)	0.002 (0.005)	0.001 (0.012)	0.003 (0.009)
el105_pre	-0.001 (0.006)	-12.251 (175.743)	0.001 (0.004)	-0.004 (0.007)	-0.001 (0.007)
el89_post	-0.008 (0.005)	-98.920 (87.800)	0.000 (0.004)	-0.006 (0.004)	-0.011 (0.005)**
el110_pre	0.001 (0.006)	198.030 (191.689)	0.001 (0.004)	-0.001 (0.008)	-0.003 (0.007)

P-value on test that all covariates \times cutoffs are 0: 0.173

Notes: This table reports the covariate balance (reduced-form) results for our 35 cutoffs \times 5 balance covariates measured in the year prior to application from the IRS data. Table F.5 reports the covariate balance (reduced-form) results for our 35 cutoffs \times 9 balance covariates measured at the time of application, from the Army applicant data. We also report the p-value on a test that all 35 cutoffs \times 14 balance covariates across both tables is 0 (P = 0.173).

* $p < .10$, ** $p < .05$, *** $p < .01$

Table F.7: Robustness: Micro-data Regressions

	Path Dependence (1)	Path Dependence (Civilian) (2)	Wages and Occ-wages (3)
<i>Panel (a): Baseline cutoffs (V-IV)</i>			
Estimate	0.186*** (0.034)	0.118*** (0.040)	0.969*** (0.208)
<i>Panel (b): Baseline cutoffs (Micro-data)</i>			
Estimate	0.163*** (0.018)	0.095*** (0.021)	1.263*** (0.141)
Overid P-value	0.306	0.264	0.625
<i>Panel (c): Expanded cutoffs (Micro-data)</i>			
Estimate	0.165*** (0.018)	0.095*** (0.021)	1.165*** (0.107)
Overid P-value	0.369	0.529	0.888

Notes: This table shows robustness to (A) measurement error; and (B) the restriction in our preferred model to focus on cutoffs with a first stage in excess of 0.01. In this table, column (1) denotes estimates of path dependence. Column (2) denotes estimates of path dependence restricted to non-Army employers only. Column (3) denotes estimates of the relationship between causal effects on wages and causal effects on occupation wages. Panel (a) gives our baseline results, repeating the key coefficients and standard errors from table 2 in the main text. Panel (b) replicates our results using a micro-data regression that is equivalent (up to a weighting) to the slope of the V-IV and Wages-Occupation Wages lines. Specifically, we stack the data for all 35 regression discontinuities and instrument a single endogenous variable (treatment take-up, in the case of the V-IV, or the occupation wage variable, in the case of wages versus occ-wages) with the 35 indicators for crossing the relevant line score threshold, all conditional on the cutoff specific polynomials (including cutoff specific fixed effects) implied by our baseline RD models. For standard errors, we cluster at the person level. See appendix D for detail on the micro-regression and proof of equivalence. Since none of the variables used in the micro-data regression are estimated quantities, the results in this panel cannot be driven by classical or non-classical measurement error. Panel (c) also presents results estimated using the micro-data regression, but expanding the number of cutoffs to include all those with an F-stat in excess of 10. Thus, our results are not an artifact of the smaller set of cutoffs that are the focus of the preferred model.

Table F.8: Civilian Only Occupational Persistence: Alternative Specifications

	Baseline standardized (1)	LLM standardized (2)	Baseline weighted, standardized (3)	LLM weighted, standardized (4)
V-IV Coefficient	0.294** (0.111)	0.250*** (0.084)	0.405*** (0.125)	0.392*** (0.109)

Notes: This Table reports different standardized, V-IV slopes for occupational persistence 11-20 years out for civilian only persistence. Column (1) reports a specification like our baseline civilian only V-IV, where the only change is that the reduced-form coefficients on being in a linked occupation are standardized at each cutoff to have mean 0 and standard deviation 1, before regressing these reduced form estimates on the usual first stages across cutoffs. We do this so that all estimates in this table can be put on a similar scale. Column (2) is as in column (1), but using indicators for being in any LLM-linked occupation (as opposed to our baseline cross-walked occupations). Column (3) is as in column (1), but now we use information contained in the MOS-specific ‘first stages’ to upweight SOC-codes linked to MOS’s that are more heavily shifted by our cutoffs as described in Appendix C. We then standardize at each cutoff (to put estimates on the same scale across cutoffs) and re-estimate the relevant V-IV slope. Column (4) is as in column (3), but we use LLM-linked occupations instead of cross-walked linked occupations. In column 4, the weighting procedure is similar to that of column (3), but weights are constructed using both MOS-first stage weights and LLM probability weights for each SOC at each MOS (see Appendix C). We continue to standardize the corresponding reduced-form estimates at each cutoff afterwards and re-estimate the relevant V-IV slope. Heteroskedasticity robust standard errors, calculated directly from the V-IV regression, are reported in parentheses.

Table F.9: MOS-level Occupational Persistence

	(1) All Baseline	(2) FS \geq 0 Baseline	(3) FS \leq 0 Baseline	(4) All Cutoff FEs	(5) FS \geq 0 Cutoff FEs	(6) FS \leq 0 Cutoff FEs	(7) All FEs and Weights	(8) FS \geq 0 FEs and Weights	(9) FS \leq 0 FEs and Weights
VIV Coefficient	0.110*** (0.019)	0.119*** (0.024)	0.077* (0.040)	0.112*** (0.020)	0.123*** (0.024)	0.103** (0.043)	0.088*** (0.030)	0.104*** (0.030)	0.105** (0.051)
Observations	4799	2164	2635	4799	2164	2635	4799	2164	2635
Cutoff Fixed Effects	N	N	N	Y	Y	Y	Y	Y	Y
Weighted by MOS size	N	N	N	N	N	N	Y	Y	Y

Notes: This table accompanies Figure F.12 and reports regressions of the MOS-specific reduced form effects on civilian persistence against the MOS-specific first stages across all cutoffs. Columns (1)-(3) report simple regressions with no further adjustments, columns (4)-(6) include cutoff fixed effects for each of the 35 cutoffs, and columns (7)-(10) additionally weight by MOS size (in the full sample). Within each of these sets, the first column reports the regression on the whole, relevant sample, the second does so on the subset of MOS's with a positive first stage, and the third does so on the subset of MOS's with a negative first stage. Heteroskedasticity robust standard errors are reported in parentheses.

Table F.10: Occupational Persistence (11-15) By Race

	(1) Baseline	(2) Baseline	(3) Black	(4) Black	(5) White	(6) White
	Military & Civ.	Civilian Only	Military & Civ.	Civilian Only	Military & Civ.	Civilian Only
V-IV:						
Slope Coef.	0.186 (0.034)	0.118 (0.040)	0.147 (0.061)	0.036 (0.055)	0.167 (0.063)	0.139 (0.091)
Slope Coef. (no constant)	0.186 (0.020)	0.108 (0.022)	0.135 (0.029)	0.035 (0.029)	0.199 (0.036)	0.166 (0.044)

Notes: This table reports heterogeneity on the path dependence results in table 2. Columns (1) and (2) restate our baseline for comparison purposes. Columns (3) and (4) report results when we restrict the sample to only Black applicants. Columns (5) and (6) restrict to White applicants. For each, we report all occupational persistence results from table 2 panel (a). Standard errors are robust to heteroskedasticity (constructed using the delta method and the full variance-covariance matrix).

Table F.11: Stability of Correlation between Wages and MOS-Implied Wages

	Estimate (SE)	Adj R^2	Adj R^2 of cov. alone
<i>Panel A: Baseline (MOS-Implied Wages)</i>			
No controls	0.779*** (0.335)	0.173	
<i>Panel B: Tasks and Union Features</i>			
Occupation Non-routine Tasks	0.982*** (0.346)	0.171	-0.007
Occupation Routine Tasks	0.759*** (0.236)	0.150	-0.008
Occupation Union Membership	0.842*** (0.273)	0.148	0.062
<i>Panel C: Education Levels</i>			
Bachelors+	0.778*** (0.260)	0.147	-0.016
College Attendance	0.962*** (0.286)	0.194	-0.030
Associates Degree	0.823*** (0.240)	0.179	-0.019
<i>Panel D: Certification Requirements</i>			
Academy Training	0.722** (0.314)	0.148	0.110
Any Certification Required	1.007*** (0.275)	0.168	0.010
Vocational Certification	0.784*** (0.227)	0.156	-0.025
<i>Panel E: Training Features</i>			
Vocational Credits	0.702** (0.303)	0.156	0.047
Total Cost	0.769** (0.292)	0.147	0.023
Time (Weeks)	0.746*** (0.270)	0.149	0.018
College Credits	0.922*** (0.278)	0.165	-0.013
<i>Panel F: Military Features</i>			
Years Served in Army	1.353*** (0.212)	0.345	-0.024
MOS is Combat	0.853*** (0.172)	0.154	-0.013
VADC	1.287*** (0.211)	0.348	-0.016
Joined Active Duty	1.276*** (0.209)	0.320	-0.025
Months Deployed	1.128*** (0.208)	0.268	-0.026
Initial Contract Length	1.293*** (0.245)	0.318	-0.026

Notes: This table presents results from bivariate regression models of the form $Y_c = \alpha + \beta Z_c + \pi X_c + e_c$, where Y_c is the 2SLS estimate (model 2) of the causal effect on realized earnings generated by the occupations unlocked at cutoff c ; Z_c is the corresponding estimate of the causal effect on MOS-level implied future earnings; and X_c is a control variable that corresponds to the estimated causal impact on another MOS-level outcome (e.g. total training weeks – see appendix B for more detail on how these MOS level variables are constructed). Panel (A) replicates the unconditional results from figure 7 as a baseline. Panels (B) through (F) give results conditional on corresponding changes in other MOS-level characteristics. The control variables are all constructed as MOS-level leave-out averages so as to mirror the construction of MOS-level implied future earnings. For example, years served is average, leave-out years served in the MOS you join (instead of your own years served). Column (2) reports the adjusted R^2 from these bivariate regressions, while Column (3) reports the adjusted R^2 from regressing the causal effect on realized earnings against the control variable alone (without including MOS-level implied future earnings). Heteroskedasticity robust standard errors, calculated directly from the regression, are reported in parentheses.

Table F.12: First Stages for the Multiple Endogenous Variable 2SLS

Rel. Line Score:	Combat / Construction	Computers / Electronics	Utilities / Equipment Management	Warehouse Operations / HR	Transport / Food Prep	Health / Police / Intelligence / Other Specialists	Mechanics / Maintenance	Other Army Job
	co fa (1)	el (2)	gm (3)	cl (4)	of (5)	st (6)	mm (7)	(8)
co inst	0.046*** (0.002)	-0.002** (0.001)	-0.004*** (0.001)	-0.003 (0.002)	-0.016*** (0.002)	-0.005*** (0.002)	-0.009*** (0.001)	-0.003*** (0.001)
fa inst	0.016*** (0.002)	-0.002** (0.001)	-0.004*** (0.001)	-0.004** (0.002)	-0.006*** (0.002)	-0.002 (0.001)	-0.002 (0.001)	0.001 (0.001)
el inst 1	-0.007*** (0.002)	0.022*** (0.001)	-0.003*** (0.001)	-0.001 (0.002)	0.000 (0.001)	-0.004** (0.002)	-0.004*** (0.001)	0.001 (0.001)
el inst 2	0.000 (0.002)	0.013*** (0.001)	-0.003** (0.001)	-0.002 (0.001)	-0.005*** (0.001)	-0.006*** (0.001)	-0.002 (0.002)	0.002* (0.001)
el inst 3	0.000 (0.002)	0.007*** (0.001)	0.000 (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.002 (0.001)	0.000 (0.002)	-0.002* (0.001)
gm inst	-0.012*** (0.003)	0.000 (0.002)	0.029*** (0.002)	-0.006*** (0.002)	-0.005*** (0.002)	-0.003 (0.002)	-0.004* (0.002)	0.000 (0.002)
cl inst	-0.001 (0.004)	-0.001 (0.001)	-0.005*** (0.002)	0.054*** (0.003)	-0.016*** (0.003)	0.000 (0.002)	-0.006** (0.002)	-0.004** (0.002)
of inst	-0.015*** (0.002)	0.000 (0.001)	-0.005*** (0.001)	-0.015*** (0.002)	0.049*** (0.002)	-0.004*** (0.001)	-0.006*** (0.001)	-0.002** (0.001)
st inst 1	-0.012*** (0.002)	-0.001 (0.001)	-0.006*** (0.001)	-0.017*** (0.002)	-0.013*** (0.002)	0.056*** (0.001)	-0.006*** (0.001)	-0.004*** (0.001)
st inst 2	0.001 (0.002)	-0.003*** (0.001)	-0.002** (0.001)	-0.004*** (0.002)	-0.002 (0.001)	0.015*** (0.002)	-0.001 (0.002)	-0.001 (0.001)
st inst 3	-0.001 (0.002)	-0.005*** (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.003*** (0.001)	0.011*** (0.002)	0.000 (0.001)	-0.003*** (0.001)
mm inst 1	-0.007*** (0.002)	-0.001* (0.001)	-0.005*** (0.001)	-0.005*** (0.002)	-0.013*** (0.002)	0.002 (0.002)	0.028*** (0.001)	-0.001 (0.001)
mm inst 2	-0.001 (0.002)	-0.002* (0.001)	-0.003*** (0.001)	0.000 (0.001)	0.000 (0.001)	-0.003** (0.001)	0.012*** (0.001)	-0.001 (0.001)
mm inst 3	-0.002 (0.002)	-0.003* (0.001)	-0.003*** (0.001)	0.001 (0.001)	-0.002 (0.001)	-0.003* (0.002)	0.015*** (0.002)	-0.002 (0.001)
firstafqt 31	0.027*** (0.002)	-0.001 (0.001)	0.011*** (0.001)	0.006*** (0.001)	0.025*** (0.001)	0.001 (0.001)	0.018*** (0.001)	0.005*** (0.001)
firstafqt 50	0.014*** (0.002)	-0.003*** (0.001)	0.002** (0.001)	0.024*** (0.001)	0.008*** (0.001)	0.006*** (0.002)	0.011*** (0.001)	0.003*** (0.001)
Observations	2,106,817	2,106,817	2,106,817	2,106,817	2,106,817	2,106,817	2,106,817	2,106,817

Notes: This table reports estimates of Equation 6. Each column is a separate first stage regression, with column titles indicating the outcome (i.e. the eventual endogenous variable in Equation 5 that indicates whether or not you joined an MOS in the relevant group.) The gray-shaded cells indicate the threshold-crossing instruments for which we expect positive effects (i.e. they correspond to a threshold that unlocks the MOS in the indicated job group). The first stages are weighted by the same weights as income 11-20.

Table F.13: Multiple Endogenous Variable 2SLS Results

	(1)	(2)
Combat / Construction	8418*** (3049)	9202*** (2893)
Computers / Electronics	10679 (7450)	12833* (6914)
Utilities / Equipment Management	6797 (7293)	9455 (6438)
Warehouse Operations / HR	-3610 (3733)	-3049 (3670)
Transport / Food Prep	5450* (2961)	6182** (2819)
Health / Police / Intelligence / Other Specialists	8160*** (3031)	7789*** (2992)
Mechanics / Maintenance	4661 (4445)	5948 (4080)
Other Army Job	18713 (23454)	
Observations	2,106,817	2,106,817
Dep. Var Mean	\$32,785	\$32,785
Dep. Var Mean (Didn't Join Army)	\$29,727	\$29,727
Over-id p-value	0.144	0.157
P-value for all $\beta_c = 0$	0.013	< 0.001

Notes: This table reports estimates of Equation 5. It uses 16 instruments for crossing a specific threshold in either the pre-2002 or post-2004 period. There are 28 (14 in each period) occupation cutoffs and 4 (2 in each period) first AFQT cutoffs. There are 8 endogenous variables spanning 9 mutually exclusive categories, with the omitted category being civilian. The 8 endogenous variables correspond to different sets of army occupations, with the last of these being ‘other Army job’ (i.e. an Army MOS that is not shifted by any of the occupation cutoffs), which the AFQT cutoffs help identify. All specifications include the relevant running variable controls, defined separately in the pre-2002 and post-2004 period and only non-zero within each relevant RD window, and a dummy for the post-2004 period. Since some occupation groups are shifted by multiple instruments (e.g. the mechanical maintenance and repair group is identified from 3 separate MM instruments – above 88post/90pre, above 99post/100pre, and above 104post/105pre), we also report results from an over-identification test. Column (2) is as in column (1) but we drop the endogenous variable for ‘other army job’ effectively making the omitted category ‘civilian or other Army job’, but we note that effectively this is still mostly civilians as other Army jobs make up only 4% of the sample.

Table F.14: MOS-to-SOC Crosswalk

cutoff	MOS-FS	Counts	MOS	Pri	Sup	Sec	Sup2	cutoff	MOS-FS	Counts	MOS	Pri	Sup	Sec	Sup2	cutoff	MOS-FS	Counts	MOS	Pri	Sup	Sec	Sup2
co90,pre	0.031	1	11B	55-0	NA	33-3	33-1	gm90,pre	0.007	1	12N	51-4	51-1	53-7	53-1	st105,pre	0.006	1	35F	55-0	NA	33-3	33-1
co90,pre	0.014	1	19K	55-0	NA	33-3	33-1	gm90,pre	0.006	1	88H	53-7	53-1	51-9	51-1	st105,pre	0.003	1	35N	49-9	49-1	49-2	49-1
co90,pre	0.010	1	12B	49-9	49-1	47-2	47-1	gm90,pre	0.004	1	92W	51-9	51-1	51-8	51-1	st105,pre	0.001	0	68X	21-1	NA	29-2	NA
co90,pre	0.007	1	19D	55-0	NA	33-3	33-1	gm90,pre	0.003	1	12W	47-2	47-1	49-9	49-1	st105,pre	0.001	0	35L	33-3	33-1	33-9	NA
co90,pre	0.003	0	92R	55-0	NA	33-3	33-1	gm90,pre	0.003	1	92R	55-0	NA	33-3	33-1	st105,pre	0.001	0	68S	19-4	NA	51-8	NA
co90,pre	0.001	0	12C	53-7	53-1	49-3	49-1	gm90,pre	0.001	0	91F	55-0	NA	33-3	33-1	st105,pre	0.001	0	35G	17-1	NA	15-1	NA
co90,pre	0.001	0	11C	55-0	NA	49-9	NA	gm90,pre	0.001	0	92M	39-4	39-1	11-9	NA	st105,pre	0.001	0	15Q	53-2	53-1	43-5	43-1
cl95,pre	0.024	1	92A	13-1	11-9	11-9	NA	gm90,pre	0.001	0	12K	47-2	47-1	49-9	NA	st105,pre	0.000	0	12T	17-3	NA	17-1	NA
cl95,pre	0.019	1	92Y	43-9	43-1	43-5	43-1	gm90,pre	0.001	0	12V	47-2	47-1	47-4	NA	st105,pre	0.000	0	37F	27-3	NA	11-2	NA
cl95,pre	0.017	1	42A	13-1	NA	43-4	43-1	gm90,pre	0.001	0	12M	33-2	33-1	43-5	NA	st105,pre	0.000	0	68F	29-1	NA	29-1	NA
cl95,pre	0.002	0	56M	21-1	NA	21-2	NA	cl90,post	0.016	1	92A	13-1	11-9	11-3	NA	st105,pre	0.000	0	68V	29-1	NA	29-2	NA
cl95,pre	0.001	0	68J	11-3	NA	43-5	NA	cl90,post	0.014	1	92Y	43-9	43-1	43-5	NA	st105,pre	0.000	0	68Y	29-2	NA	29-1	NA
cl95,pre	0.001	0	68G	29-2	NA	29-9	NA	cl90,post	0.008	1	42A	13-1	NA	43-4	43-1	st105,pre	0.000	0	68U	29-2	NA	29-1	NA
st95,pre	0.021	1	68W	29-2	NA	29-1	NA	cl90,post	0.003	0	56M	21-1	NA	21-2	NA	st105,pre	0.000	0	68C	29-1	NA	29-2	NA
st95,pre	0.015	1	31B	33-3	33-1	33-9	33-1	cl90,post	0.003	0	68J	11-3	NA	43-5	NA	st105,pre	-0.000	0	68B	31-2	NA	51-9	NA
st95,pre	0.014	1	74D	55-0	NA	33-3	33-1	cl90,post	0.002	0	68G	29-2	NA	29-9	NA	st105,pre	-0.000	0	68N	29-1	NA	29-2	NA
st95,pre	0.004	1	15P	43-5	43-1	43-4	43-1	fa95,pre	0.020	1	13B	55-0	NA	33-3	33-1	st105,pre	-0.000	0	35S	49-2	49-1	15-1	NA
st95,pre	0.001	0	68E	51-9	51-1	31-9	NA	fa95,pre	0.006	1	13J	55-0	NA	33-3	33-1	st105,pre	-0.000	0	35M	33-3	33-1	33-9	33-1
st95,pre	0.001	0	68D	29-2	NA	31-1	NA	mm90,pre	0.016	1	91B	49-3	49-1	49-9	49-1	el100,pre	0.011	1	25G	47-2	47-1	49-9	49-1
st95,pre	0.001	0	35P	27-3	NA	13-1	NA	mm90,pre	0.005	1	91L	49-3	49-1	49-9	49-1	el100,pre	0.003	1	25C	49-9	49-1	47-2	47-1
st95,pre	0.001	0	92L	51-8	51-1	19-4	NA	mm90,pre	0.005	1	91J	51-6	51-1	49-9	49-1	el100,pre	0.001	0	94T	49-2	49-1	17-3	NA
st95,pre	0.001	0	68Q	29-2	NA	31-9	NA	mm90,pre	0.002	0	91H	49-3	49-1	53-3	NA	el100,pre	0.001	0	15Y	49-3	49-1	49-2	NA
st95,pre	0.000	0	25V	27-4	NA	25-4	NA	mm90,pre	0.000	0	91S	49-3	49-1	49-9	NA	mm100,pre	0.005	1	91A	49-3	49-1	53-7	53-1
st95,pre	0.000	0	25M	27-1	NA	27-4	NA	co87,post	0.018	1	11B	55-0	NA	33-3	33-1	mm100,pre	0.004	1	15R	47-2	47-1	49-9	49-1
st95,pre	0.000	0	68T	29-2	NA	31-9	NA	co87,post	0.012	1	12B	49-9	49-1	47-2	47-1	mm100,pre	0.002	1	15S	49-9	49-1	47-2	47-1
st95,pre	0.000	0	31K	33-9	33-1	39-2	NA	co87,post	0.008	1	19D	55-0	NA	33-3	33-1	mm100,pre	0.002	1	88K	53-7	53-1	53-5	NA
of90,pre	0.016	1	92F	53-3	53-1	53-7	NA	co87,post	0.006	1	19K	55-0	NA	33-3	33-1	mm100,pre	0.001	1	14J	15-1	NA	49-2	NA
of90,pre	0.014	1	92G	35-2	35-1	11-9	NA	co87,post	0.005	1	12C	53-7	53-1	49-3	49-1	e198,post	0.016	1	25Q	47-2	47-1	49-9	49-1
of90,pre	0.013	1	88M	53-3	53-1	53-7	NA	co87,post	0.002	0	11C	55-0	NA	49-9	NA	e198,post	0.003	1	25C	49-9	49-1	47-2	47-1
of90,pre	0.009	1	14S	43-5	NA	17-3	NA	co87,post	-0.000	0	92R	55-0	NA	33-3	33-1	mm100,pre	0.005	1	94T	49-2	49-1	17-3	NA
st91,post	0.020	1	31B	33-3	33-1	33-9	33-1	mm88,post	0.024	1	91B	49-3	49-1	49-9	49-1	e198,post	0.000	0	15Y	49-3	49-1	49-2	NA
st91,post	0.012	1	89B	43-5	NA	43-4	NA	mm88,post	0.005	1	91J	51-6	51-1	49-9	49-1	mm104,post	0.006	1	15T	49-3	49-1	49-9	49-1
st91,post	0.011	1	74D	55-0	NA	33-3	33-1	mm88,post	0.002	0	91M	49-3	49-1	53-7	53-1	mm104,post	0.004	1	15U	49-9	49-1	47-2	47-1
st91,post	0.006	1	15P	43-5	43-1	43-4	43-1	mm88,post	0.002	0	91H	49-3	49-1	53-3	NA	mm104,post	0.003	1	14E	55-0	NA	33-3	33-1
st91,post	0.004	1	35M	33-3	33-1	33-9	33-1	mm88,post	0.002	0	91L	49-3	49-1	49-9	49-1	mm104,post	0.002	0	15G	51-4	51-1	49-3	NA
st91,post	0.002	0	68E	51-9	51-1	31-9	NA	mm88,post	0.002	0	91A	49-3	49-1	53-7	53-1	mm104,post	0.002	0	15J	49-2	49-1	49-9	NA
st91,post	0.001	0	68T	29-2	NA	31-9	NA	mm88,post	0.001	0	91S	49-3	49-1	49-9	NA	mm104,post	0.001	0	15D	49-3	49-1	53-6	NA
st91,post	0.001	0	92L	51-8	51-1	19-4	NA	mm88,post	0.001	0	91P	49-3	49-1	49-9	NA	mm104,post	0.001	0	15B	49-3	49-1	51-2	NA
st91,post	0.000	0	68D	29-2	NA	31-1	NA	mm105,pre	0.005	1	15T	49-3	49-1	49-9	49-1	mm104,post	0.001	0	15E	49-2	49-1	49-3	NA
st91,post	0.000	0	35P	27-3	NA	13-1	NA	mm105,pre	0.003	1	14E	55-0	NA	33-3	33-1	mm104,post	0.001	0	15F	49-2	49-1	47-2	NA
st91,post	0.000	0	68Q	29-2	NA	31-9	NA	mm105,pre	0.003	1	91M	49-3	49-1	53-7	53-1	mm104,post	0.001	0	15H	49-3	49-1	49-9	NA
st91,post	0.000	0	31K	33-9	33-1	39-2	NA	mm105,pre	0.003	1	15U	49-9	49-1	47-2	47-1	mm104,post	0.000	0	15M	49-3	49-1	49-2	NA
st91,post	-0.000	0	25V	27-4	NA	25-4	NA	mm105,pre	0.002	0	15G	51-4	51-1	49-3	NA	mm104,post	-0.001	0	15Y	49-3	49-1	49-2	NA
gm88,post	0.012	1	91D	51-9	51-1	51-8	NA	mm105,pre	0.001	0	15Y	49-3	49-1	49-2	NA	el105,pre	0.007	1	25N	15-1	NA	11-3	NA
gm88,post	0.007	1	92R	55-0	NA	33-3	33-1	mm105,pre	0.001	0	15D	49-3	49-1	53-6	NA	el105,pre	0.002	1	94F	49-9	49-1	51-9	51-1
gm88,post	0.007	1	88H	53-7	53-1	51-9	NA	mm105,pre	0.001	0	15B	49-3	49-1	51-2	NA	el105,pre	0.002	1	94E	49-9	49-1	51-9	51-1
gm88,post	0.006	1	91C	51-9	51-1	47-2	47-1	mm105,pre	0.001	0	15H	49-3	49-1	49-9	NA	el105,pre	0.001	0	94D	53-2	53-1	49-2	NA
gm88,post	0.005	1	92W	51-9	51-1	51-8	NA	mm105,pre	0.000	0	91P	49-3	49-1	49-9	NA	el105,pre	0.000	0	94A	49-9	49-1	51-9	51-1
gm88,post	0.004	1	91F	55-0	NA	33-3	33-1	mm105,pre	0.000	0	88L	49-3	49-1	53-6	NA	gt107,post	0.032	1	68W	29-2	NA	29-1	NA
gm88,post	0.003	0	12W	47-2	47-1	49-9	49-1	mm105,pre	0.000	0	15M	49-3	49-1	49-2	NA	gt107,post	0.001	0	68C	29-1	NA	29-2	NA
gm88,post	0.002	0	91E	51-4	51-1	47-2	NA	mm105,pre	0.000	0	15E	49-2	49-1	49-3	NA	gt107,post	0.000	0	18X	55-0	NA	33-3	NA
gm88,post	0.001	0	12K	47-2	47-1	49-9	NA	mm105,pre	-0.000	0	15J	49-2	49-1	49-9	NA	gt107,post	0.000	0	18B	55-0	NA	33-3	NA
gm88,post	0.001	0	92M	39-4	39-1	11-9	NA																

Table F.15: MOS in each occupation category for the joint 2SLS model

Notes: This table lists the Army Occupations (MOS) that are counted in each occupation category in the multiple endogenous variable 2SLS regression (Equation 5). Unlisted MOS are counted in the “Other Army Job” category. ‘MOS Perc.’ reports the percentage of each occupation category represented by the stated MOS, calculated using the full sample MOS counts. ‘Relevant Cutoffs’ reports the cutoffs that primarily unlock MOS in the stated occupation category.