Mixing the Rich and Poor: The Impact of Peers on Education and Earnings*

Elias Einiö†

December 10, 2019

Abstract

This study exploits a large-scale natural experiment in Finnish conscription to study how exposure to peers from different family backgrounds affects education, earnings, employment, and hourly wage. Our research design is based on the alphabetic rule in assigning conscripts to dorms, which induces credible exogenous variation in peer family backgrounds. Being exposed to a dormmate from a high-income family has a positive long-term effect on earnings. The effects are the largest for individuals who come from high-income families. Exposure to peers with one standard deviation higher average parent income raises their earnings at age 28-42 by 2.6%. The results suggest beneficial labor market networks as a key mechanism. Exposure to peers from high-income families has little impact on earnings and hourly wages of individuals who come from low-income families, but it increases their educational attainment in the long run. The findings imply that social stratification reinforces economic and educational inequality between rich and poor families.

JEL Classification: J24, J31

Keywords: Earnings, employment, education, wages, peer effects, family background.

*Thanks are due to Xavier Jaravel, Steve Pischke, Matti Sarvimäki, Otto Toivanen, Ulf Zölitz, and numerous seminar participants for helpful comments and suggestions. The data used in this article are confidential but the author's access is not exclusive. Financial support from the Yrjö Jahnsson Foundation is gratefully acknowledged. The permission to use Conscript Registry data is based on the decision number AM4503 of the Finnish Defense Forces.

†Valtion taloudellinen tutkimuskeskus (VATT) and Centre for Economic Performance (CEP), London School of Economics. Address: Arkadiankatu 7, 00101 Helsinki, Finland; Email: elias.einio@vatt.fi; Tel.: +358 295 519 408.

1 Introduction

What are the effects of mixing individuals from different family backgrounds? This question is of major research interest in social sciences and of great public and economic policy importance. Several scholars have argued that American and many European societies are becoming more stratified (e.g., Musterd et al., 2017; Reardon et al., 2018). This means that individuals sort to social groups that are more similar to themselves and interact less with others who are different. If social interaction with people from more affluent backgrounds improves labor market outcomes, rising stratification can reinforce economic inequality (Durlauf, 1996; Bowles et al., 2014). The unequal effects of stratification can be amplified by borrowing constraints among the poor, nonlinearities in the intergenerational transmission of human capital, and general equilibrium effects in the labor market (Fernández and Rogerson, 2001). Despite the potentially large economic and societal consequences, empirical evidence on the long-term impacts of the socioeconomic composition of social peer groups and the mechanisms through which it shapes societies is scarce.

In this paper, we shed light on these questions by examining how exposure to individuals from different family backgrounds affects economic and educational outcomes in a large-scale natural experiment generated by Finnish military conscription. A key feature of conscription that allows us to identify the impacts is that conscript dorms mix individuals from a wide range of family backgrounds. This gives us an exceptional opportunity to assess the impacts of social interaction between individuals with different socioeconomic backgrounds.

We focus on parent income as the primary measure of family background and identify causal effects by exploiting the predominant use of the alphabetic rule in assigning conscripts to dorms within squads. Due to this rule, conscripts who are next to each other in their squad's alphabetic ordering are also likely to be next to each other in their dorm's alphabetic ordering. This generates strong correlation between parent income of alphabetically nearest squadmates and parent income of alphabetically nearest dormmates. We show that, conditional on parent pre-service income and squad fixed effects that control for unobserved differences in squad composition, the variation induced by the alphabetic assignment is as good as random. This allows us to identify the causal effects of dormmates' parent income on economic and educational outcomes.

The analysis is made possible by our access to the Conscript Registry of the Finnish

Defense Forces (FDF), which includes information on squads and dorms a conscript served in. We merge this data with a rich population panel dataset and extensive data on hourly wages and hours based on a survey of employers. Our first finding is that when a conscript is assigned a dormmate with high parent income because of the alphabetic assignment rule, his earnings are higher in the long term. Our estimate implies that increasing peer parent income by one standard deviation, equal to around ten thousand euros (or nine thousand US dollars), raises earnings at age 28-42 by around 1.9%. The impacts are largest for individuals who come from high-income families, among whom we find a 2.6% increase in earnings in the long run. This estimate corresponds to around a 21 thousand euro (or 19 thousand US dollar) increase in the discounted value of wage earnings until retirement at age 65. The results suggest that a large fraction of the effect on earnings is explained by an increase in the hourly wage of individuals who come from high-income families. We find no statistically or economically significant effects on employment, annual work days, or work hours.

Conscripts in our sample are at age 18-22. At this age, individuals typically apply for post-secondary education programs. To examine the extent to which an increase in human capital investment could explain the positive long-term effects on wages and earnings, we estimate the impacts on post-secondary enrollment and years of schooling. We find that higher peer parent income increases university enrollment in the short run and years of schooling in the long run for individuals who come from low-income families, but not for individuals who come from high-income families. In the former group, increasing peer parent income by one standard deviation raises the share of individuals who are enrolled to a university one to three years after the service by around 2.2 percentage points or 7.4% and years of schooling in the long run by around two months or 1.3%.

Another potential mechanism that could explain the positive effects of peer parent income on earnings and hourly wage are labor market networks that may help individuals to find jobs with better pay or provide insider information that could help negotiate higher wages (Bayer et al., 2008; Cingano and Rosolia, 2012). We test for the existence of labor market networks by using worker-level information on company codes of employers. Using the alphabetic research design, we find that dormmates are more likely to work in the same company in the long term than conscripts who are in different dorms in the same squad. Interestingly, this effect is driven by pairs where both conscripts are from high-income families. This suggests that labor market networks among individuals from high-income families are a key mechanism generating the positive peer effects on earnings.

Our paper contributes to the small but growing literature examining the long-term impacts of family background composition of the social group an individual interacts with. Most previous research in this strand has focused on peer exposure in primary schools. In their study of peer exposure in Norwegian primary schools, Black et al. (2013) use idiosyncratic within-school variation in peer characteristics and find that schoolmates in the same grade who have very high earning fathers improve earnings of boys. Carrell et al. (2018) examine the long-term impacts of peers in elementary schools of Alachua County, finding that peers who are exposed to domestic violence reduce earnings by 3% in the long run. Our study is also closely related to literature examining peer effects within military dorms and squadrons. Carrell et al. (2009) examine peer effects on test scores in the U.S. Air Force Academy, finding positive impacts of high-achieving peers, while Carrell et al. (2013) conduct a field experiment to study optimal peer composition. In their field experiment study of the Norwegian military forces, Dahl et al. (2018) find that exposure to females in a squad causes men to adopt more egalitarian attitudes. Methodologically, our paper is closest to Rao (2019), who exploits alphabetic assignment of poor students to study groups in elite schools in Delhi. He finds that having poor classmates makes wealthy students more prosocial, generous, and egalitarian. In assessing the long-term effects of exposure to peers from different family backgrounds, our paper is also related to literature that examines the impacts of neighborhoods, as social interaction is an important potential channel through which neighborhoods may affect educational and labor market outcomes. Recent research has provided compelling evidence that exposure to a better neighborhood during childhood contributes to higher income in adulthood (e.g., Chetty et al., 2016; Chetty and Hendren, 2018).

Our study contributes to this literature in several ways. First, we provide causal evidence of the long-term impacts of peers in a large-scale natural experiment mixing individuals from all parts of the family background distribution. This allows us to assess how interaction with peers affects earnings across the family income distribution, which is particularly important for understanding the economic consequences for individuals from poor and rich families. Second, we provide novel evidence on the channels through which social interaction with people from different family backgrounds affects earnings. In particular, we show that the positive effects on earnings from being exposed to peers from high-income families are driven by an increase in hourly wage. Our finding of educational investment among the poor and job market networks among the rich as the key mechanisms of peer effects sheds new light on the question of how social spillovers shape the distribution of earnings.

The paper is organized as follows. Section 2 summarizes the institutional features of Finnish conscription. Section 3 describes the data sources and reports summary statistics. Section 4 discusses the alphabetic research design and reports tests verifying its validity. We present our main results in section 5. Section 6 examines the mechanisms and section 7 concludes.

2 Background

Finland has compulsory conscription for all men. All 18 year old men are required to register for service. Conscripts have two alternative ways to serve. They can either complete the military service or civil service. Our analysis includes individuals who choose to undertake the military service.

The first step of the conscription process is a call-up event where conscripts complete a questionnaire that assesses their cognitive and mental suitability for service. Conscripts also go through a basic health check by military medical staff and an interview with a recruitment officer, in which they can express their desired military branch and service location. Based on these tests, conscripts are allocated to their branch and service location. A small fraction of conscripts are exempt due to not being suitable for service. A conscript who has a post-primary study place may postpone the start of service. Women can serve voluntarily, but their fraction is small, around 1%. Around 30 thousand conscripts start the military service each year and they cover around 80% of the relevant male cohorts.

The military service starts with a base training period, which lasts approximately eight weeks. It is followed by four to ten months further training, the length of which depends on the conscript's branch and assignment. Some conscripts take the corporal or reserve officer school after the base training period to become junior officers, who lead teams and squads.

Military training is organized in management units *(perusyksikkö)*, which are managed by senior command personnel. Within management units, conscripts are allocated to squads and teams, which follow a hierarchical structure. Team members typically live in the same dorm and train together. In our data, the median dorm size is eight conscripts. The minimum service time is 180 days, and many conscripts spend their whole service in the same dorm. This generates intensive and long-lasting social interaction between them.

3 Data and Sample Selection

3.1 Data Sources and Variable Definitions

Our study employs several administrative data sources from Finland that we can link through unique identifiers for conscripts and their parents. Information on conscripts comes from the FDF's Conscript Registry (CR). The data include identifiers for management units, squads, and dorms a conscript served in for conscripts who started service in 1996-2006. We merge this information with longitudinal administrative registers compiled by Statistics Finland (SF), which cover the whole working aged population in Finland. The data set covers socioeconomic variables (including age, sex, earnings, income, and educational attainment) and child-parent links, which allow merging parent characteristics. Over the period 1995-2016, we can link these data to information on hourly wages from an extensive hourly wage register compiled by SF.

We next define the key variables used in our analysis.

Dorm and squad. In our empirical analysis, we examine how dormmates' family background affects post-service outcomes. A key variable is the dorm code, which identifies the last dorm a conscript served in. We observe 6,756 dorms covering around 24% of individuals in the CR. These dorms have a median size of nine conscripts. The dorm code is not recorded for all conscripts because management units are not obligated to record them in the common IT system and around a one-fourth of management units use it for administrative purposes. Squad information has full coverage in the CR, because management units are obligated to record it.

Income. Our primary measure of income includes wage earnings, income from a company, pension income, and capital income as well as taxable social security benefits. We also include housing allowances, which is a major non-taxable benefit. We impute zeros for missing values of these variables and calculate income as their sum.[1] All income is measured prior to the deduction of taxes. Parent income is the average income of parents. To reduce the influence of transitory shocks, we use the average parent income over three years preceding the service. We restrict the pre-service years to three because the capital income data are limited to the year 1993 and hence we do not observe it for more than

[1]Information on these variables is from the tax records, which have full coverage of the population. Thus, missing values for them can be treated as true zeros.

three pre-service years for the first conscript wave, which started service in 1996.

Hourly wage. We use extensive data on hourly wages from a researcher-use hourly wage database provided by Statistics Finland. The data are based on surveys to employers who are asked to send information on their employees earnings and work hours. The advantage of this data is that they are from administrative company registers, and hence are little susceptible to response bias that may introduce measurement error in individual-level wage surveys. The data are collected by employer organizations and Statistics Finland and compiled by Statistics Finland, and cover around 63% of conscripts in our baseline sample. The hourly wage variable is calculated as total wage earnings divided by total work hours (regular work time and over time hours).

Completed years of education. We construct a measure of completed years of education by using the education codes in the population panel. We estimate the average time used to complete an education programme by employing administrative data from the SF register of post-primary students, which provides information on the study programme a student attended for all Finnish students at the secondary (high school and vocational) and tertiary (universities) levels. Our primary measure of completed years of education is based on the average time used to complete a programme by 3-digit programme code. We merge this estimated completion time to the population panel by 3-digit education code for the highest degree, which is based on the same classification as the code for the study programme. For all individuals, we add nine years of schooling for compulsory primary education.[2] For programs that require high school degree, we add three years for high school studies.

3.2 Sample selection and summary statistics

In most of our analysis, we study conscripts who start the service at the age 18 to 22. This sample covers 99.8% of the conscript population and excludes conscripts who start service voluntarily at age 17 or postpone service start to age 23 or later. We further restrict the sample to conscripts for whom the dorm code is observed and who have at least one parent. We also exclude conscripts who have siblings serving in the same squad. This leaves us with a baseline sample of 50,578 conscripts.

All monetary variables are deflated to 2012 with the CPI. To reduce the effects of outliers we cap monetary variables at the upper tail to the 99.5th percentile of the distribution. Hourly wages are further limited to 50% of the national basic unemployment benefit scheme

[2]In Finland, children are obligated by law to attend the primary school for nine years at age seven to 16.

minimum wage requirement at the lower tail of the distribution.[3]

Table 1 displays summary statistics for the baseline sample. Panel A shows means and standard deviations for background characteristics, which are measured one year before service (except service start age). Conscripts start service at age 19.74, on average. 40% have been employed during the year before the service. The mean years of schooling is 10.43, with very few having a university degree, which is expected for this age group. The mean parent income is 29,980 euros. To demonstrate the representativeness of the sample in terms of family background, figure 1 displays the distribution of income for conscripts' parents and the full Finnish population at age 35-65, separately for men and women. It shows that conscripts' parents come from all parts of the income distribution.

Panel B of table 1 reports summary statistics at age 28-42.[4] This longitudinal sample includes 376,963 conscript-year observations. The mean age in this sample is around 33 years, mean wage earnings is 32,080 euros, and mean years of education is 13.77. The fraction of individuals with completed university degree is 35%. The sample includes 165,042 hourly wage observations for 32,688 conscripts. The mean hourly wage is 19.85 euros per hour.

4 Research design

4.1 Alphabetic dorm assignment rule

In general, conscripts are not randomly assigned to dorms. For causal inference, we need a research design that mimics random assignment. We exploit the alphabetic rule which is commonly used in assigning conscripts to dorms within squads. Figure 1 presents empirical evidence of the strong tendency to use the alphabetic rule when assigning conscripts to dorms within squads. It shows the fraction of conscript pairs within a squad who reside in the same dorm by the within-squad alphabetic rank distance between the pair members. Value one on the X-axis indicates pairs who are next to each other in the squad alphabetic ordering. In our data, around 42% of such pairs reside in the same dorm. As the alphabetic rank distance increases, the fraction of pairs residing in the same dorm declines rapidly, indicating a strong alphabetic pattern in dorm assignment. By the alphabetic rank distance of 30, the fraction of pair members residing in the same dorm is only around 10%.

[3] This requirement defines minimum levels of hourly compensation for work spells that are considered when determining eligibility for the benefit.

[4] The maximum age observed in the population panel ending 2016 for the conscripts in our baseline sample is 42. This is for individuals who start service at age 22 in the year 1996.

Our correspondence with the FDF staff indicated that a key reason for the use of the alphabetic rule is that it is practical in head counting during which the on-duty officer of a management unit records present conscripts in a head counting book.[5] One sheet in a head counting book covers typically one dorm. In order to make the head counting process more efficient, conscripts are commonly assigned to dorms and listed in head counting books by alphabetic order. When dorms are alphabetically assigned, the alphabetic rule is also often used to assign conscripts to bunks within the dorm, because then the order of conscripts is the same on the head counting sheet and across the bunks in the dorm, making it easier to identify missing conscripts. These assignment practices imply that conscripts who are close to each other in the squad alphabetic ordering are more likely to be assigned to the same dorm. Moreover, when alphabetically proximate conscripts are assigned to the same dorm in a squad, they are also by construction alphabetically proximate within the dorm, and more likely to reside in contiguous bunks. As a result, exposure to alphabetically nearby co-conscripts can be expected to be higher, on average, than exposure to others who are more distant in the alphabetic ranking.

4.2 Empirical specification

We use the quasi-random assignment of dormmates induced by the alphabetic rule to estimate the causal impact of peers on the outcome Y. In our main analysis, we focus on a conscript's two alphabetically nearest dormmates and estimate the following regression by two-stage least squares:

$$Y_{idst} = \gamma \overline{X}_{(i)ds} + \delta X_{ids} + \alpha_s + \alpha_t + \epsilon_{idst} \tag{1}$$

where $\overline{X}_{(i)ds}$ is the average parent income of conscript i's two alphabetically nearest dormmates, which is instrumented with the average parent income of his two alphabetically nearest squadmates. X_{ids} is the conscript's own parent income, α_s is a squad fixed effect, α_t is a fixed effect for the year of outcome measurement, and ϵ_{idst} is the error term. We also estimate specifications where we control for a large set of background variables.

Conditioning on squad fixed effects controls for average unobserved differences between squads. Thus, the inference is not affected by selection of conscripts by unobserved characteristics, such as ability, at the squad level. The IV estimate of γ is identified

[5] Also Harinen (2013) reports the use of the alphabetic rule in assigning conscripts to dorms in the Guard Jaeger Brigade. Appendix figure A1 shows a page of a head counting book.

from within-squadron variation in parent income of two alphabetically nearest squadmates across the squad alphabetic ordering. The alphabetic assignment rule induces within-squad correlation between $\overline{X}_{(i)ds}$ and the instrument because when either of a conscript's two alphabetically nearest squadmates is assigned to the same dorm with him, he will also be, by construction, his alphabetically nearest dormmates. In the case where both alphabetically nearest squadmates are assigned to the same dorm with the conscript, $\overline{X}_{(i)ds}$ and the instrument will be equivalent.

Note that our approach does not require that there is no clustering of income in specific parts of the alphabetic distribution (for instance, due to rare family names associated with high income). This is because controlling for the conscript's own parent income accounts for the level of income and for unobserved factors associated with it at the location of the conscript and his alphabetic neighbors in the alphabetic distribution. It also accounts for the potential bias due to the mechanical correlation between the mean peer parent income $\overline{X}$ and own parent income X. If this correlation is not equal to zero and own parent income X is correlated with the outcome of interest, excluding the own parent income will induce omitted-variable bias (see Angrist (2014) for a derivation and discussion of this bias).[6] It is also worth noting that our analysis is not affected by the reflection problem which arises in situations where the measure of peer characteristic is affected by peer interaction (Manski, 1993), because the peer characteristic in our regression is realized before conscripts enter the service. In order to test the validity of the research design, we next present results for similar validation tests as are typically performed in randomized settings.

4.3 Verifying quasi-randomness of the alphabetic research design

To verify that the alphabetic rule generated within-squad variation in peer parent income that is as good as random, we run a set of reduced-form regressions of a pre-service characteristic on the instrument and own parent income, and control for squad fixed effects. We use a wide set of pre-service characteristics from the population panel data set for conscripts and their parents, measured one year before the service. Table 2 shows the results. Each row shows an estimate from a regression on the outcome specified by the row label. Although own parent income is highly correlated with most of these characteristics, they are not statistically related to parent income of two alphabetically nearest squadmates: none of the coefficients on this instrument are statistically significant at the conventional

[6] Note that random assignment of peer groups does not rule out the mechanical correlation between a leave-own-out peer mean $\overline{X}$ and own characteristic X.

significance levels.

4.4 Relevance of the instrument

Panel A in figure 3 illustrates the first-stage variation that we utilize. This figure plots residuals from separate regressions of parent income of two alphabetically nearest squadmates and dormmates on own parent income and squad fixed effects. The estimation uses the baseline sample of 50,578 conscripts. To make the graph correspond to the main sample which includes observations for conscripts at age 28-42, the regression plot is weighted by the number of observations available for each conscript in it. The slope in the figure is 0.314 with a standard error of 0.010 and it is equivalent to the estimate in the first column of table 3. This result indicates that high average parent income of two alphabetically nearest squadmates strongly predicts that the average parent income of two alphabetically nearest dormmates is high. The coefficient indicates that when the average parent income of two alphabetically nearest squadmates increases by 10 thousand euros, the average parent income of two alphabetically nearest dormmates is 3,140 euros higher. This relationship arises from the prevalent use of the alphabetic rule in assigning conscripts to dorms within squads: squadmates next to each other in the squad alphabetic ordering are likely to reside in the same dorm; when they are assigned to the same dorm, they are by construction also alphabetic neighbors within the dorm. Therefore, if the parent income of a conscript's alphabetically proximate squadmates is high, also the parent income of his alphabetically proximate dormmates is likely to be high.

5 Effects of peer family background

5.1 Earnings

We begin our presentation of the main results by reporting the estimates of the impact of peer parent income on earnings. Panel B in figure 3 plots the reduced-form effect of the average parent income of two alphabetically nearest squadmates against conscript's wage earnings at age 28-42. This figure uses residuals from separate regressions of average parent income of two alphabetically nearest squadmates and wage earnings on own parent income, dummies for calendar year, and squad fixed effects. The graphical evidence suggests that when a conscript's two alphabetically nearest squadmates have high average parent income, his earnings are higher at age 28-42. The slope in the figure is 0.0193 and it is equivalent to the reduced-form estimate in the second column of table 3. The IV estimate of the impact of the average parent income of two alphabetically nearest dormmates on wage earnings in column 3 is obtained by dividing the reduced-form estimate by the first-stage estimate. It takes into account that all alphabetically proximate squadmates are not assigned to the same dorm. The IV estimate is 0.0615 and it is statistically significant at the 5% risk level. This estimate implies that when the average parent income of two alphabetically nearest dormmates increases by 10 thousand euros, which is equivalent to around one standard deviation, a conscript's annual earnings will be around 615 euros higher at age 28-42. Compared to the 32,080 euro sample mean of earnings at age 28-42, this corresponds to around a 1.9% increase, or to around a 14 thousand euro (or 12.6 thousand US dollar) increase in the discounted value of wage earnings until retirement at age 65.[7]

5.2 Internal validity

Table 2 provided support to the assumption that the alphabetic research design induces as good as random variation in parent income of co-conscripts by showing that the instrument is uncorrelated with a large set of conscript's and his parents' background characteristics. We further examine the validity of the research design in table 4. Panel B shows results when we add control variables for pre-service characteristics of the conscript and his parents and for average parent income of other dormmates, who are not the two alphabetically nearest. If the alphabetic order assigns conscripts to dorms as well as randomly, adding these variables should not affect the estimates. As expected, the estimates are little affected by the inclusion of these additional control variables.

To provide further support for the validity of the alphabetic research design, we control for the conscript's location in the alphabetic ranking by adding 100 dummies for each population alphabetic rank distribution percentile in panel C. This has very little impact on the estimates. In panel D, we exclude observations where the conscript and his two alphabetically nearest squadmates are within one population alphabetic rank percentile. This specification employs variation in parent income of a conscript's squadmates who are the two alphabetically nearest in the squad alphabetic ranking, but relatively distant in the population alphabetic ranking. This specification eliminates the potential impacts of systematic clustering of correlated unobserved characteristics at specific points of the population alphabetic distribution. Even though the sample size is smaller for this

[7] The increase in the discounted value of earnings is calculated from age 28 to age 65, using a discount rate of 3% and assuming that the effect on earnings is constant over time.

specification, it is reassuring to find that the IV estimate is larger compared to the estimates in the baseline sample. Panels E and F show that the estimates do not change appreciably when we restrict the sample to individuals at age 30-35 or exclude female conscripts.

5.3 Employment and unemployment benefits

We now turn to results for employment outcomes. Panel B in table 5 shows results for a binary employment outcome, which is equal to one if a person worked during the year and zero otherwise. We convert this binary outcome to percentages by multiplying it by 100. The IV point estimate is positive but relatively small and statistically insignificant. Panel C shows results for the number of days an individual is employed during the year, which is available in the population panel from 2005 onward. This variable is based on the number of days a person has been covered by work pension insurance, which is a mandatory social insurance for employees payed by the employer. The IV estimate for this outcome is also positive and small. These point estimates imply that when parent income of two alphabetically nearest dormmates increases by 10 thousand euros, employment increases by around 0.6% and work days by around 0.5% compared to the sample means of these variables. Panel D shows results for unemployment benefits. The effect on unemployment benefits is negative, and significant at the 10% risk level. The estimate suggests that when the parent income of two alphabetically nearest dormmates increases by 10 thousand euros, annual unemployment benefits reduce by 57 euros, or by around 5.4%. This estimate is in line with improved labor market prospects when individuals are exposed to peers with high parent income.

5.4 Wages and hours

We next turn to estimating the impact of peers on hourly wages and work hours using data from the SF wage register. We observe hourly wage at age 28-42 for around 63% of conscripts in our baseline sample. The coverage of wage data is incomplete for two reasons. First, the sample frame of the surveys does not cover the whole economy. Second, wages are observed for individuals who are working in the respondent companies, and not for unemployed. These two sources of selection may cause bias if the instrument is correlated with the incidence of wage being observed. In the first column of table 6, we test for selection in a reduced-from regression where the outcome is a binary indicator equal to one if wage is observed and zero otherwise, using the full baseline sample. The coefficient on the instrument is 0.0005 with a standard error of 0.0007. This insignificant and small estimate indicates that the average parent income of two alphabetically nearest squadmates does not affect the likelihood of being observed in the wage data. Thus, we have no reason to believe that IV estimates using the wage sample would be affected by sample selection.

Column 2 shows the first stage estimate for the wage sample, which is very similar compared to the first stage estimate for the baseline sample in column 1 of table 3. The IV estimate for the hourly wage is 0.0425 and it is highly significant (p-value<0.01). The estimate implies that when the average parent income of two alphabetically nearest dormmates increases by 10 thousand euros, hourly wage increases by around 43 cents (2.1%). The coefficient for work hours is statistically insignificant and economically small and implies a relative impact of around 0.3% compared to the sample mean.

5.5 Heterogeneity

In table 7, we explore heterogeneity of the impact of peers by own parent income. The table shows the IV estimates of the impacts of parent income of two alphabetically nearest dormmates separately for conscripts whose own parent income is below and above the median. Estimates in panel A indicate that exposure to peers from high-income families has the largest impact on earnings of individuals who also come from high-income families. For them the IV estimate of 0.0922 implies that when the average parent income of two alphabetically nearest dormmates increases by 10 thousand euros, annual wage earnings will be around 922 euros higher, or increase by around 2.6% compared to the sample mean of 34,960 euros in this sample. This effect corresponds to around a 21 thousand euro (or 19 thousand US dollar) increase in the discounted value of wage earnings until retirement at age 65.[8] The estimate for conscripts from low-income families is substantially smaller (0.0169) than the estimate for the full sample or for the sample of conscripts from high-income families, and it is not statistically significant. Results in panel E show similar heterogeneity for hourly wage. The IV estimate for conscripts from high-income families is 0.0580 and implies that when the average parent income of two alphabetically nearest dormmates increases by 10 thousand euros, their hourly wage will be around 58 cents higher, or increase by around 2.8% compared to the sample mean of 20.81 euros in this sample. Differences in the coefficients for employment and work hours between the two

[8]The discounted value of wage earnings is calculated from age 28 to age 65, using a discount rate of 3% and assuming that the effect on earnings is constant over time.

groups are relatively small.

Overall, the results so far suggest that the positive impact of peer parent income on earnings is primarily driven by increases in the hourly wage of individuals who come from high-income families. We find little evidence of better employment or increased work hours being a channel for the earnings effect. We next examine the potential mechanisms that could explain these findings.

6 Mechanisms

6.1 Human capital investment

One potential channel through which peers could affect earnings is schooling. Conscripts in our sample are at age 18-22. At this age, individuals typically apply for post-secondary education programs and sharing information on the admission process, entry examination requirements, and expected returns to education might influence their decision to apply for a particular level and field of study.

We examine the impact of peer parent income on three educational outcomes. The first is a binary indicator for whether an individual is studying in a tertiary level education programme (corresponding to U.S. colleges and universities) one to three years after the end of the service and the second is a binary indicator for whether an individual has completed a tertiary degree at age 28-42. We convert these binary outcomes to percentages by multiplying them by 100. The third outcome is a measure of completed years of schooling, which is constructed from average years used to complete a study programme (see section 3 for details).

Table 8 reports the results for the full sample and separately for conscripts who are below and above median parent income. For the baseline sample, the IV estimate for studying in a tertiary programme one to three years after the service is 0.0986 and it implies that increasing the average parent income of two alphabetically nearest dormmates by 10 thousand euros increases the likelihood of studying in a tertiary programme by around 0.1 percentage points, but this estimate is not statistically significant at the conventional significance levels. For individuals who have low-income parents, the estimate is 0.2210, and it is statistically significant at the 5% risk level. It implies that increasing the average parent income of two alphabetically nearest dormmates by 10 thousand euros increases the likelihood of studying in a tertiary programme by around 2.2 percentage points or by 7.4% from the sample mean of 29.62%. The estimate for individuals who have high-income parents is substantially smaller and implies that increasing the average parent income of two alphabetically nearest dormmates by 10 thousand euros increases the likelihood of studying in a tertiary programme by only around 0.1% from the sample mean of 51.85%.

We next turn to results for long-term educational attainment. The IV estimates for completed tertiary degree at age 28-42 have a similar pattern across the samples as the IV estimates for tertiary enrollment one to three years after the service, but the estimate for completed tertiary degree in the sample of individuals who have low-income parents is lowered to 0.1719. This estimate implies that increasing the average parent income of two alphabetically nearest dormmates by 10 thousand euros increases the likelihood of completing a tertiary degree by around 1.7 percentage points or by 6.9%, which is fairly close to the corresponding estimate for tertiary enrollment, but the estimate is insignificant at the conventional confidence levels ($p=0.12$). In an attempt to gain more statistical power, we use the years of schooling as a second outcome for educational attainment. The pattern of IV estimates across the samples is, again, very similar. The estimate for individuals who have low-income parents is 0.0164, and it is statistically significant at the 5% risk level. This estimate implies that increasing the average parent income of two alphabetically nearest dormmates by 10 thousand euros increases years of schooling by around two months (1.3%). For individuals who have high-income parents the coefficient is, again, small and insignificant.

Overall, the findings in this section suggest that investment in human capital is a key mechanism through which exposure to peers from high-income families operates among individuals from low-income families. It might be, for instance, that individuals from low-income families update their beliefs about the potential returns to education when they are exposed to peers from high-income families. Also their preferences for social status may change when they are exposed to peers from more affluent backgrounds. The absence of impacts on educational outcomes for individuals from high-income families could be due to the fact that their baseline educational attainment is high. They might also have received more information about the returns to higher education from their high-income parents, who are, on average, more educated than low-income parents.

6.2 *Labor market networks*

Another potential mechanism that could explain the positive impacts on earnings are networks that conscripts may form during the service. Previous research has provided evidence of the importance of networks for employment (e.g. Bayer et al., 2008; Cingano and Rosolia, 2012). Individuals within networks may, for instance, share valuable information on job opportunities and provide job referrals to each other. They may also share information on wage formation of employers, which may help negotiate better wages.

To test the importance of labor market networks, we examine whether conscripts are more likely to work in the same firm with their dormmates than with conscripts who are in the same squad but reside in a different dorm. We use pairwise data for all within-squad pairs of conscripts in the baseline sample and estimate a dyadic regression

$$y_{ii'} = \alpha_{s(ii')} + \phi d_{ii'} + \epsilon_{ii'} \tag{2}$$

where $y_{ii'}$ is a binary indicator equal to one if conscripts i and i' work in the same firm at age 28-42, and zero otherwise; $d_{ii'}$ is a binary indicator equal to one if conscript i' is among the two alphabetically nearest dormmates of conscript i; $\alpha_{s(ii')}$ is a fixed effect for squad s and $\epsilon_{ii'}$ is an error term. To account for the potential within-squad selection of conscripts to dorms by unobserved characteristics, we instrument $d_{ii'}$ with a binary indicator equal to one if conscript i' is among the two alphabetically nearest squadmates of conscript i. This IV approach employs the same within-squad variation across the alphabetic ordering as our baseline IV peer effect model. The instrument induces variation in who is whose alphabetically nearest dormmate because alphabetically nearest squadmates are more likely to reside in the same dorm, and when they do, they are by construction the two alphabetically nearest dormmates.

In panel A of table 9, the first-stage coefficient on the dummy for a pair being among the two alphabetically nearest squadmates is 0.262 with a standard error of 0.006. This estimate implies that the likelihood of being among the two alphabetically nearest dormmates is around 0.26 points higher among the two alphabetically nearest squadmates compared to other squadmates. The IV coefficient on the dummy for a pair being among the two alphabetically nearest dormmates is 0.0023 and it is statistically significant at the 5% risk level. This estimate implies that the fraction of conscripts who are working in the same company is 0.23 percentage points higher for conscript pairs who are among the two alphabetically nearest within their dorm than for other pairs. This is approximately a 70% increase from the sample mean of 0.33 percentage points and indicates that social networks formed during the service affect placement in the labor markets several years later.

In panels B to D, we test for the importance of networks by pair type. The IV coefficient is statistically insignificant and small for pairs where both members are from low-income family and for mixed pairs where the other pair member is from a low-income family and the other is from a high-income family. For pairs where both conscripts are from high-income families, the coefficient is 0.0087 and highly statistically significant ($p<0.001$). This estimate implies that the fraction of conscripts who are working in the same company is 0.87 percentage points higher for conscript pairs who are among the two alphabetically nearest within their dorm than for other pairs. Relative to the sample mean, this effect corresponds to almost tripling the fraction of individuals who are working in the same company. Overall, these findings suggest that beneficial labor market networks are a key mechanism generating the positive earnings effects for individuals who come from high-income families, but not for others. They also suggest that a lack of persistent networks between individuals from high- and low-income families is a potential explanation for why earnings of individuals from low-income families are little affected by peers from high-income families.

7 Conclusions

This paper provides causal evidence on the long-term effects of social sorting in a setting where we can use credibly exogenous variation in peer parent income across an exceptionally wide range of the income distribution. Our research design is based on the alphabetic rule in assigning conscripts to dorms in Finnish conscription. Due to this rule, conscripts who are close to each other in the squad alphabetic ordering are also likely to be alphabetically and physically close to each other in their dorm. We show that, conditional on squad fixed effects and measures of pre-service income, this rule generates variation in dormmate parent income that is as good as random. We find evidence of positive long-term effects of peer parent income on hourly wages and annual wage earnings which are strongest for individuals from high-income families.

Our findings imply that when individuals sort to increasingly similar socioeconomic groups, the long-term labor market prospects of individuals from high-income families improve the most. They also mean that income sorting can have persistent effects on economic inequality by widening the wage and earnings gap between individuals from the richest and poorest families. Our findings also suggest that policies that reduce social stratification in peer environments that expose individuals from different family backgrounds to social interaction can decrease economic inequality. This reduction occurs because earnings of individuals from more affluent family backgrounds decline, while earnings of individuals from less-advantaged family backgrounds are little affected.

Our findings raise the question of what are the mechanisms through which social interaction with peers from different family backgrounds shapes society? Our results indicate that these mechanisms are different for individuals from low- and high-income families. A key mechanism driving the positive earnings and wage effects for individuals from high-income families are labor market networks formed with peers who also come from high-income families. On the other hand, we find that exposure to peers from high-income families increases educational attainment of individuals from the poorest families. While we do not find that this increased human capital investment among the poor increases their earnings, it may affect the educational outcomes and labor market prospects of their children. This may amplify the unequal economic effects of social stratification. We leave the analysis of such intergenerational effects for the future, when data on the educational and economic outcomes of conscripts' children become available.

References

Angrist, J. D. (2014): "The perils of peer effects," *Labour Economics*, 30, 98–108.

Bayer, P., S. L. Ross, and G. Topa (2008): "Place of work and place of residence: Informal hiring networks and labor market outcomes," *Journal of Political Economy*, 116, 1150–1196.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2013): "Under pressure? The effect of peers on outcomes of young adults," *Journal of Labor Economics*, 31, 119–153.

Bowles, S., G. C. Loury, and R. Sethi (2014): "Group inequality," *Journal of the European Economic Association*, 12, 129–152.

Carrell, S. E., R. L. Fullerton, and J. E. West (2009): "Does your cohort matter? Measuring peer effects in college achievement," *Journal of Labor Economics*, 27, 439–464.

Carrell, S. E., M. Hoekstra, and E. Kuka (2018): "The long-run effects of disruptive peers," *American Economic Review*, 108, 3377–3415.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013): "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 81, 855–882.

Chetty, R. and N. Hendren (2018): "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects," *Quarterly Journal of Economics*, 133, 1107–1162.

Chetty, R., N. Hendren, and L. F. Katz (2016): "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment," *American Economic Review*, 106, 855–902.

Cingano, F. and A. Rosolia (2012): "People I know: job search and social networks," *Journal of Labor Economics*, 30, 291–332.

Dahl, G. B., A. Kotsadam, and D.-O. Rooth (2018): "Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams," NBER Working Paper 24351.

Durlauf, S. N. (1996): "A theory of persistent income inequality," *Journal of Economic Growth*, 1, 75–93.

Fernández, R. and R. Rogerson (2001): "Sorting and long-run inequality," *The Quarterly Journal of Economics*, 116, 1305–1341.

Harinen, O. (2013): "Varusmiesten maanpuolustustahto, kansallinen identiteetti ja kokemukset palvelusajasta: Tuloksia puolustusvoimissa vuosina 2005-2009 tehdyistä varusmieshaastatteluista," Käyttäytymistieteiden laitos 1/2013, Maanpuolustuskorkeakoulu, Helsinki.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 60, 531–542.

Musterd, S., S. Marcińczak, M. Van Ham, and T. Tammaru (2017): "Socioeconomic segregation in European capital cities: Increasing separation between poor and rich," *Urban Geography*, 38, 1062–1083.

Rao, G. (2019): "Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools," *American Economic Review*, 109, 774–809.

Reardon, S. F., K. Bischoff, A. Owens, and J. B. Townsend (2018): "Has income segregation really increased? Bias and bias correction in sample-based segregation estimates," *Demography*, 55, 2129–2160.

Figure 1: Income distributions, conscripts' parents and full population at age 35-65

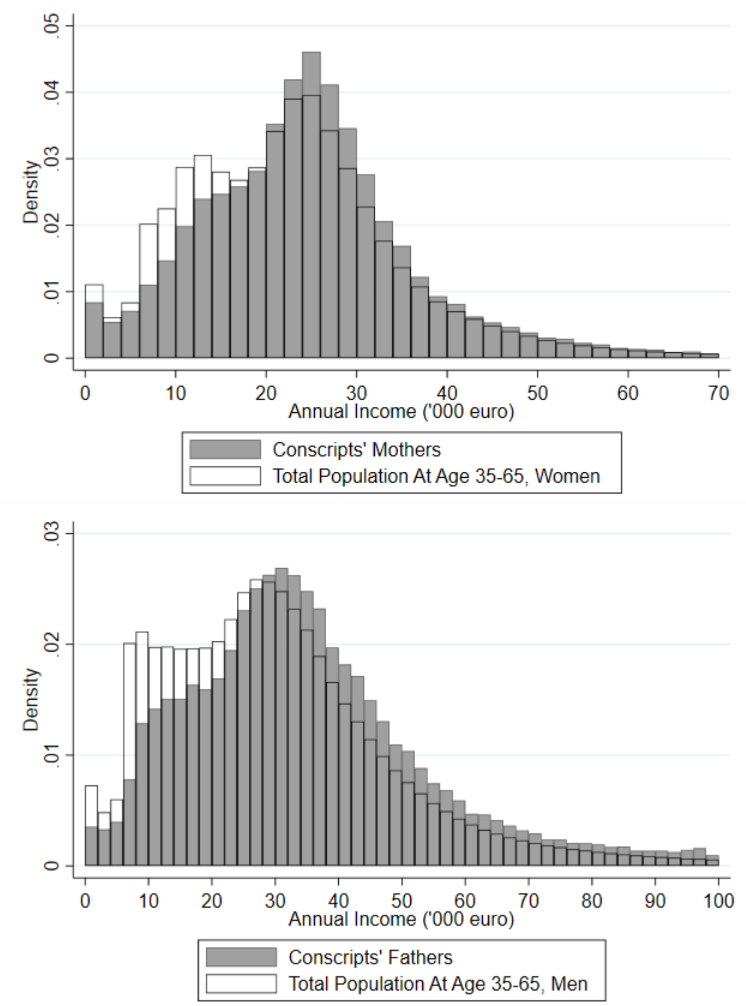


Figure 2: Within-squad alphabetic rank distance and the likelihood of residing in same dorm

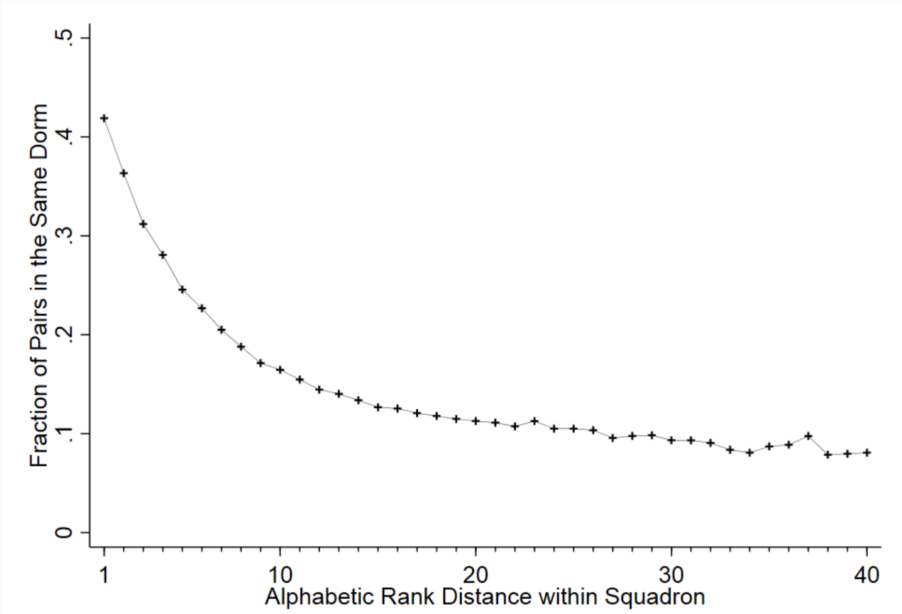
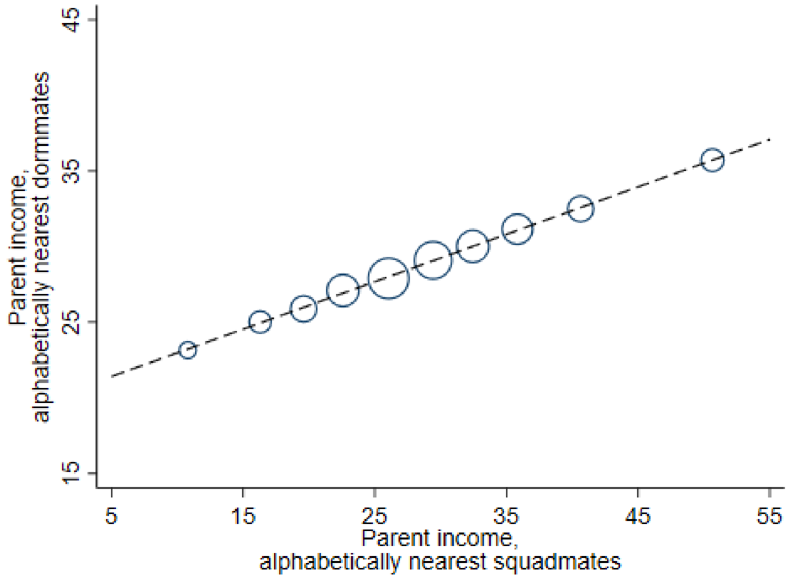
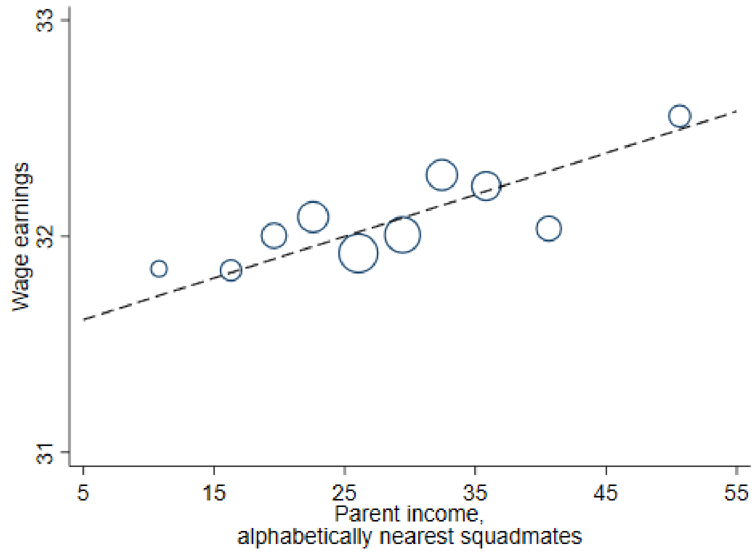


Figure 3: Effect of nearest squadmates' parent income on nearest dormmates' parent income (first stage) and conscript's earnings at age 28-42 (reduced form)

A. First stage



B. Reduced form



Notes: The figure displays the effect of average parent income of two alphabetically nearest squadmates on average parent income of two alphabetically nearest dormmates (panel A) and wage earnings at age 28-42 (panel B). Data include 376,963 conscript-year observations ($50,578$ conscripts). Income and earnings are in thousand euro. We plot the residuals from separate regressions of wage earnings, nearest squadmates' parent income, and nearest dormmates' parent income on own parent income, dummies for calendar year, and squad fixed effects. The size of the circle represents the number of observations within each bin. The slope is 0.314 in panel A and 0.0193 in Panel B.

Table 1:
Summary statistics

	Mean	Std. Dev.
A. Background characteristics		
Age (at start of service)	19.74	0.67
Wage earnings	5.00	5.79
Employed	0.40	0.49
Years of schooling	10.43	1.52
University degree	0.00	0.01
Married	0.00	0.05
Foreign	0.00	0.01
Primary language Finnish	0.96	0.20
Parent income	29,980	15,200
Parent income, two alphabetically nearest dormmates	28,940	10,860
Parent income, two alphabetically nearest squadmates	29,030	10,850
B. Age 28-42		
Age	32.85	2.96
Wage earnings	32,080	20,600
Hourly wage	19.85	6.58
Employed	0.84	0.37
Employment days	315.09	112.08
Years of schooling	13.77	3.23
University degree	0.35	0.48

Notes: Data for 50,578 conscripts. Unless otherwise stated, all conscript and parent characteristics in panel A are measured the year before the start of the service. Means and standard errors in panel B are calculated for 376,963 conscript-year observations, except for hourly wage, for which they are calculated for 165,042 conscript-year observations. Nominal values are deflated to 2012.

Table 2:
Validation regressions

	Independent variable				
	Own parent income		Parent income, two alphabetically nearest squadmates		Dependent
Dependent variable	coeff.	s.e.	coeff.	s.e.	mean
Wage earnings	0.0041*	(0.0023)	-0.0020	(0.0025)	5.00
Employed (%)	0.0419**	(0.0175)	-0.0180	(0.0220)	39.58
Years of schooling	0.0054***	(0.0005)	0.0009	(0.0006)	10.43
Married (%)	-0.0044***	(0.0017)	0.0035	(0.0022)	0.27
Foreign (%)	-0.0002	(0.0002)	0.0003	(0.0002)	0.00
Primary language Finnish (%)	0.0256***	(0.0056)	-0.0045	(0.0073)	95.78
Unemployment benefits	-0.0045***	(0.0003)	0.0002	(0.0003)	0.28
General housing allowance	-0.0086***	(0.0003)	-0.0002	(0.0004)	0.26
Parent wage earnings	0.8079***	(0.0071)	-0.0011	(0.0050)	23.74
Parent employed (%)	0.8180***	(0.0153)	-0.0132	(0.0132)	79.12
Parent years of schooling	0.0947***	(0.0011)	0.0007	(0.0011)	12.30
Parent pension income	-0.0210***	(0.0013)	-0.0001	(0.0015)	1.11
Parent unemployment benefits	-0.0299***	(0.0007)	0.0010	(0.0009)	0.94
Parent general housing allowance	-0.0080***	(0.0003)	-0.0002	(0.0003)	0.20

Notes: N=50,578. Each row presents coefficients from a separate regression on the dependent variable indicated by the row label. Outcomes are measured one year before the service. Income, earnings, benefits, allowances, and pensions are in thousand euro. All regressions include squad fixed effects. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 3:
Impact of peer parent income on earnings

	First stage	Reduced form	IV
Parent income:			
of nearest squadmates	0.314***	0.0193**	
	(0.010)	(0.0093)	
of nearest dormmates			0.0615**
			(0.0297)
Dependent mean	29.00	32.08	32.08
Observations		376,963	
Conscripts		50,578	
Dorms		6,756	

Notes: The table displays the estimates of the impact of average parent income of two alphabetically nearest dormmates on wage earnings at age 28-42, using average parent income of two alphabetically nearest squadmates as the instrument. All regressions include linear terms for own parent pre-service income, dummies for calendar year, and squad fixed effects. Income and earnings are in thousand euro. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 4:
Specification checks

	First stage	Reduced form	IV	N
A. Baseline specification	0.3136***	0.0193**	0.0615**	376,963
	(0.0104)	(0.0093)	(0.0297)	
B. Additional controls	0.3097***	0.0194**	0.0628**	376,963
	(0.0104)	(0.0092)	(0.0297)	
C. Include 100 alphabetic rank percentile dummies	0.3096***	0.0195**	0.0628**	376,963
	(0.0104)	(0.0092)	(0.0297)	
D. Exclude if nearest squadmate within 1 percentile	0.2885***	0.0270**	0.0935**	204,781
in the population alphabetic ranking	(0.0119)	(0.0136)	(0.0471)	
E. Age 30-35	0.3141***	0.0203**	0.0647**	224,726
	(0.0109)	(0.0099)	(0.0314)	
F. Exclude women	0.3157***	0.0199**	0.0630**	374,560
	(0.0104)	(0.0093)	(0.0296)	

Notes: The table displays the estimates of the impact of average parent income of two alphabetically nearest dormmates on wage earnings at age 28-42, using average parent income of two alphabetically nearest squadmates as the instrument. All regressions include linear terms for own parent pre-service income, dummies for calendar year, and squad fixed effects. Additional controls are measured one year before the service and are wage earnings, dummy for employment, years of schooling, dummy for married, dummy for foreign, dummy for primary language Finnish, unemployment benefits, general housing allowances, parent wage earnings, dummy for parent employed, parent years of schooling, parent pension income, parent unemployment benefits, parent general housing allowances, dummies for age of outcome measurement, dummies for service start year, and average parent income of other dormmates, who are not the two alphabetically nearest. Parent income and earnings are in thousand euro. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 5:
Impact of peer parent income on employment and unemployment benefits

Dependent variable	Reduced form	IV	Dependent mean
A. Earnings	0.0193**	0.0615**	32.08
	(0.0093)	(0.0297)	
B. Employed (%)	0.0168	0.0535	83.71
	(0.0135)	(0.0431)	
C. Annual work days	0.0494	0.1577	315.09
	(0.0456)	(0.1458)	
D. Unemployment benefits	-0.0018*	-0.0057*	1.04
	(0.0009)	(0.0030)	

Notes: Baseline sample including 50,578 conscripts and 376,963 conscript-year observations. All regressions include linear terms for own parent pre-service income, dummies for calendar year, and squad fixed effects. Income, earnings, and benefits are in thousand euro. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 6:
Impact of peer parent income on hourly wage and work hours

	Hourly wage observed	Wage sample				
			Hourly wage		*Work hours*	
	Reduced form	First stage	Reduced form	IV	Reduced form	IV
Parent income:						
of nearest squadmates	0.0005	0.3170***				
	(0.0007)	(0.0122)				
of nearest dormmates			0.0135***	0.0425***	0.0141	0.0446
			(0.0039)	(0.0124)	(0.0125)	(0.0394)
Dependent mean	0.44	32.08	19.87	19.87	163.42	163.42
Observations	376,963			165,042		
Conscripts	50,578			32,688		
Dorms	6,756			6,693		

Notes: The table displays the estimates of the impact of average parent income of two alphabetically nearest dormmates on hourly wage and work hours at age 28-42, using average parent income of two alphabetically nearest squadmates as the instrument. All regressions include linear terms for own parent pre-service income, dummies for calendar year, and squad fixed effects. Parent income is in thousand euro. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 7:
Heterogeneity of peer effects by own parent income

Dependent variable	All	Low parent income	High parent income
A. Wage earnings	0.0615**	0.0169	0.0922**
	(0.0297)	(0.0410)	(0.0451)
B. Employment	0.0535	0.0542	0.0404
	(0.0431)	(0.0687)	(0.0590)
C. Employment days	0.1577	0.1179	0.1834
	(0.1458)	(0.2312)	(0.2016)
D. Unemployment benefits	-0.0057*	-0.0060	-0.0063*
	(0.0030)	(0.0050)	(0.0038)
E. Hourly wage	0.0425***	0.0096	0.0580***
	(0.0124)	(0.0169)	(0.0190)
F. Work hours	0.0446	0.0622	0.0611
	(0.0394)	(0.0570)	(0.0623)

Notes: The table displays the estimates of the impact of average parent income of two alphabetically nearest dormmates on outcomes indicated by the row labels, using average parent income of two alphabetically nearest squadmates as the instrument. All regressions include linear terms for own parent income, dummies for calendar year, and squad fixed effects. Each estimate is from a separate regression. Low (high) parent income samples include individuals whose own parent income is below (above) the median. The number of observations in panels A to D is 376,963 in column 1, 183,991 in column 2, and 192,965 in column 3. The number of observations in panels E and F is 165,042 in column 1, 73,193 in column 2, and 91,638 in column 3. Income, earnings, and benefits are in thousand euro. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 8:
Impact of peer parent income on education

	Studies in tertiary programme 1-3 years after service (%)			*Tertiary degree at age 28-42 (%)*			*Years of schooling at age 28-42*		
	IV	N	Dep. mean	IV	N	Dep. mean	IV	N	Dep. mean
A. All	0.0986	148,739	41.06	0.0863	376,963	34.70	0.0077	376,963	13.77
	(0.0696)			(0.0709)			(0.0048)		
B. Low parent income	0.2210**	72,182	29.62	0.1719	183,991	24.80	0.0164**	183,991	13.05
	(0.1081)			(0.1107)			(0.0074)		
C. High parent income	0.0065	76,553	51.85	0.0065	192,965	44.15	-0.0017	192,965	14.45
	(0.1038)			(0.1122)			(0.0075)		

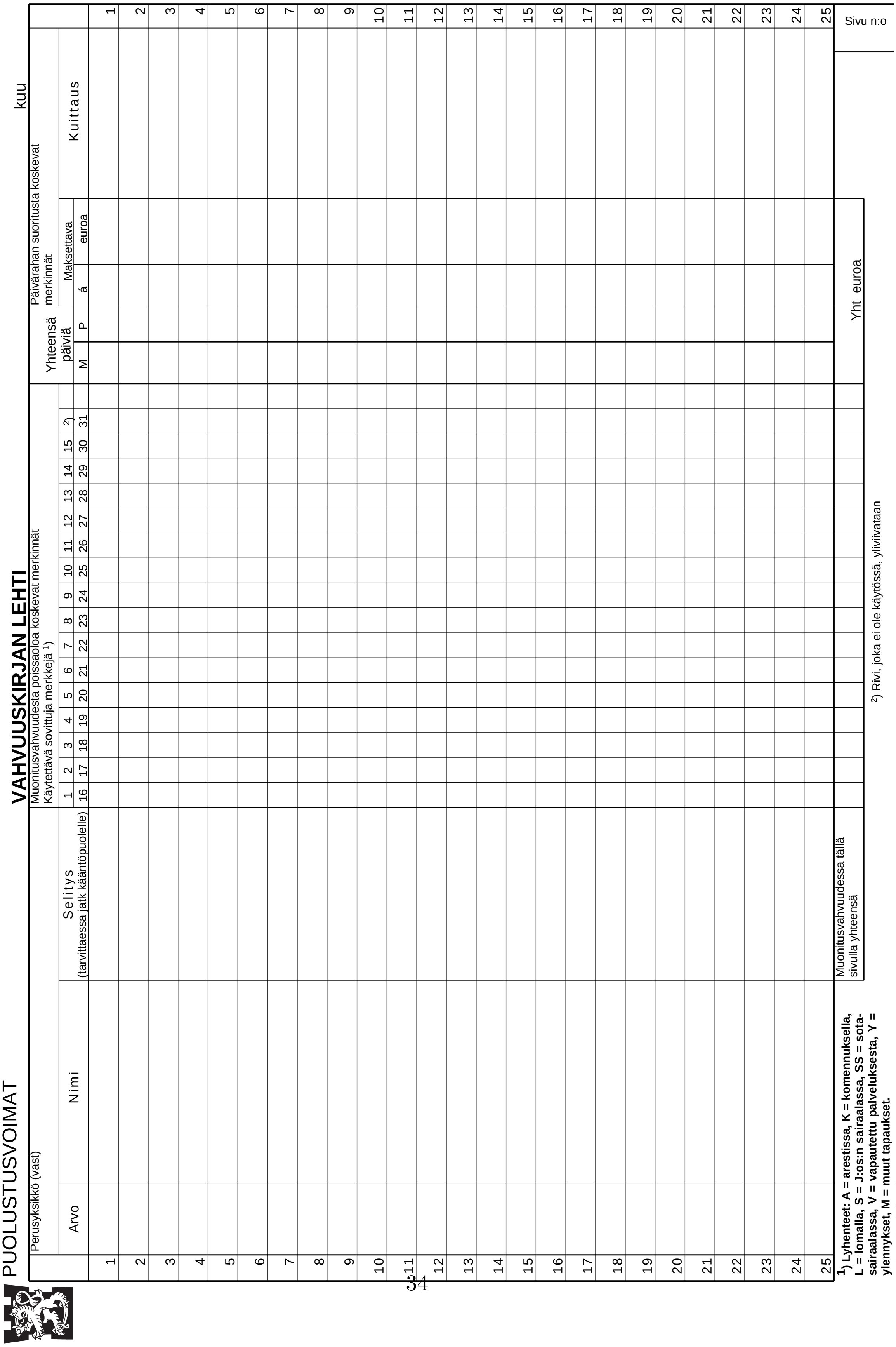
Notes: The table displays the estimates of the impact of average parent income of two alphabetically nearest dormmates on outcomes indicated by the column labels, using average parent income of two alphabetically nearest squadmates as the instrument. All regressions include linear terms for own parent income, dummies for calendar year, and squad fixed effects. Parent income is in thousand euro. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

Table 9:
Impact of residing in the same dorm on the likelihood of working in the same company, pairwise IV regressions

	First stage	Reduced form	IV
		A. All	
Pair among two alphabetically nearest squadmates	0.2612***	0.0006**	
	(0.0060)	(0.0003)	
Pair among two alphabetically nearest dormmates			0.0023**
			(0.0011)
Dependent mean		0.0033	
Pairwise observations		883,549	
	B. Both have low parent income		
Pair among two alphabetically nearest squadmates	0.2675***		
	(0.0082)		
Pair among two alphabetically nearest dormmates		0.00018	0.0004
		(0.00057)	(0.0021)
Dependent mean		0.0025	
Pairwise observations		240,329	
	C. Mixed pairs (high and low parent income)		
Pair among two alphabetically nearest squadmates	0.2590***		
	(0.0065)		
Pair among two alphabetically nearest dormmates		-0.0001	-0.0006
		(0.0004)	(0.0016)
Dependent mean		0.0031	
Pairwise observations		420,657	
	D. Both have high parent income		
Pair among two alphabetically nearest squadmates	0.2600***		
	(0.0073)		
Pair among two alphabetically nearest dormmates		0.0023***	0.0087***
		(0.0007)	(0.0028)
Dependent mean		0.0046	
Pairwise observations		221,952	

Notes: Pairwise data for all within-squad pairs of the conscripts in the baseline sample. The table displays the estimates of the impact of a binary indicator equal to one if a pair is among the two alphabetically nearest dormmates on a binary indicator equal to one if the pair is working in the same company at age 28-42, using a binary indicator equal to one if the pair is among the two alphabetically nearest squadmates as the instrument. All regressions include squad fixed effects. Standard errors allowing for clustering at the level of management unit are in parentheses. The 90%, 95%, and 99% confidence levels are denoted by *, **, and ***, respectively.

A Additional Tables and Figures


PUOLUSTUSVOIMAT

VAHVUUSKIRJAN LEHTI

kuu

Perusyksikkö (vast)

Arvo	Nimi	Selitys (tarvittaessa jatk kääntöpuolelle)	Muonitusvahvuudesta poissaoloa koskevat merkinnät Käytettävä sovittuja merkkejä 1)	Yhteensä päiviä		Päivärahan suoritusta koskevat merkinnät		Kuittaus	
			1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 2)			Maksettava			
			16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	M	P	á	euroa		
1									1
2									2
3									3
4									4
5									5
6									6
7									7
8									8
9									9
10									10
11									11
12									12
13									13
14									14
15									15
16									16
17									17
18									18
19									19
20									20
21									21
22									22
23									23
24									24
25									25
		Muonitusvahvuudessa tällä sivulla yhteensä				Yht euroa			Sivu n:o

1) Lyhenteet: A = arestissa, K = komennuksella, L = lomalla, S = J:os:n sairaalassa, SS = sota-sairaalassa, V = vapautettu palveluksesta, Y = ylennykset, M = muut tapaukset.

2) Rivi, joka ei ole käytössä, yliviivataan

7550-428-2654/Esikuntakoulu 5.99

Figure A1: Head counting sheet