Regional Science and Urban Economics 105 (2024) 103989

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Educational and gender heterogeneity of the rural-urban earnings premium: New evidence from Norway

George C. Galster [a], Liv Osland [b,*]

[a] Department of Urban Studies and Planning, Wayne State University, 656 W. Kirby St. FAB Room 3198, Detroit, MI 48202, USA

[b] Western Norway University of Applied Sciences. Haugesund, Postboks 7030 5020 Bergen, Norway and Nova Oslo Metropolitan University, Oslo, Norway

ARTICLE INFO

JEL classification:
J24
J31
J61
R10
R23
R31

Keywords:
Earnings
Agglomeration
Internal migration
Gender gaps
Urbanization
Education differentials

ABSTRACT

We explore urban earnings premiums for young, native, rural-to-urban movers in Norway. Using an augmented difference-in-differences estimator (DiD-TR) on microdata we challenge previous claims about urban earnings premium's size and sources. Conventional econometric estimators understate the static premium and overstate dynamic premiums. We find that migrants exhibit lower mean but faster pre-move earnings growth than non-migrants. Post-move, the static earnings premium dominates. The observed trajectory is related to frequent pre-move changes of industrial sector, presumably to obtain better job-worker matches. Post-move, these changes occur less frequently. Highly educated females exhibit largest static premiums (34%), less-educated females least (24%), males an intermediate amount. Our findings suggest that cities primarily generate earnings premiums through agglomeration-based efficiencies and superior job-worker matches varying heterogeneously by education and gender.

1. Introduction

Rising inequality of individual economic outcomes is a widespread and longstanding international phenomenon (Piketty and Saez, 2014) that appears especially prevalent in major metropolitan areas (Baum-Snow and Pavan, 2013; Moretti, 2013; Behrens and Robert-Nicoud, 2014**)**. In Norway, income inequality has been low for many decades, yet in line with the international situation it has been increasing recently, though with distinct spatial, gender, and educational dimensions.

In the largest Norwegian cities and their travel-to-work areas, income level and inequality typically are both above the national average and have been rising faster there, though with significant spatial variations among municipalities of the same size (Statistics Norway, 2019). It is also well-known that there is a gender dimension to inequality: income inequality is higher among females than males, and males have higher average incomes. When it comes to education, however, the traditional gender gap is now reversed (Borgonovi et al., 2018), and especially in some rural areas. In this paper, we focus on the role that internal migration from rural to urban areas plays in generating rising inequality in Norway, considering both gender and educational dimensions. According to Slettebakk (2021), the impact of internal migration on inequality is an understudied area.

A large body of international scholarship has emerged over the last quarter century—the "urban wage premium" literature—that suggests that urbanization intensifies inequality both within cities and between urban and rural areas. It finds consistently that men who move from rural to urban areas can expect an immediate boost in their inflation-adjusted earnings and a faster growth in their earnings over time, especially if they are highly skilled and the cities are larger, compared to those with similar observable characteristics who remain in rural areas (Wheeler, 2001; Baum-Snow and Pavan, 2012; Combes and Gobillon, 2015; Baum-Snow et al., 2018).

Two important issues related to measuring returns from rural-urban migration remain unresolved, one methodological and the other topical. Methodologically, empirically identifying the causal effect of urbanization on earnings is challenging primarily because urban areas tend to attract the most promising, motivated, and highly skilled workers from rural areas. The conventional approach for avoiding this selection bias arising from failure to control for unobserved characteristics of workers

* Corresponding author.
E-mail addresses: george.galster@wayne.edu (G.C. Galster), liv.osland@hvl.no (L. Osland).

https://doi.org/10.1016/j.regsciurbeco.2024.103989
Received 2 December 2022; Received in revised form 1 February 2024; Accepted 11 February 2024
Available online 28 February 2024

affecting both migration to cities and their economic performance has been to specify individual fixed effects (FE) or difference-in-differences (DiD) estimators in a longitudinal analysis. Unfortunately, as we amplify below, theory and our new evidence suggests that both static and dynamic urban earnings premiums (i.e., measured at the point of migration and as experience in the city grows) will be biased with such estimators. Topically, the urban wage premium literature has paid insufficient attention to how these premiums vary jointly by education and gender. In sum, there remain uncertainties about the degree to which rural-to-urban migration intensifies inequality of earnings not only because these two types of places treat their workers differently as a whole, but also because they treat subsets of workers differently. Clarifying this issue by employing a rarely used empirical approach is the fundamental motivation and contribution of this paper.

Specifically, we examine the annual profile of earnings after formal education is completed continuing through ages 37 to 39 by those born in Norway in 1979, 1980 and 1981 and raised in rural areas, and compare profiles of those who move to urban areas to those who stay. We study Norway because it has experienced an unusually high rate of rural-to-urban migration over the last half century, between 2.2 and 3.3 percent annually since 1969 (Steskal, 2015). Moreover, Norway has unusually rich annual panel data on all residents that have been collected in government administrative registers over an extended period. Our analytical strategy for identifying causal (as distinct from non-random selection) effects of rural-urban migration on earnings employs an augmentation of the standard DiD that not only measures pre- and post-migration differences between movers and stayers in their *levels* of earnings but also in the *trend* in their earnings over time (DiD-TR).

Our study makes both methodological and substantive contributions. We demonstrate that DiD-TR offers a more persuasive and valid means of identification than the traditional FE and DiD specifications employed in urban wage premium research. In this realm, we are the first to estimate DiD-TR earnings models jointly stratified by gender and education to test explicitly for heterogeneous effects of rural-urban migration. We find that DiD-TR produces much different estimates than FE or DiD. Conventional specifications understate the static premium because they fail to account for migrants' exceptionally low starting point but high growth in rural earnings; the latter we attribute to their more frequent changes of rural industrial sector. Moreover, these conventional specifications erroneously exaggerate the premiums gained by less-educated females. Our results not only provide nuanced insights into the substantial heterogeneity in the rural-urban earnings premiums across jointly specified gender and educational strata, but challenge on methodological grounds longstanding claims about the size and sources of the urban earnings premium.

2. The earnings returns from rural-to-urban migration

2.1. Theory

2.1.1. The reasons for an urban wage premium

There are four distinct sets of explanations why moving from a rural to an urban area could yield an increase in earnings: "compensation," "selection," "efficiency," and "productivity" (Yankow, 2006). The first set suggests that higher earnings in urban areas are needed to *compensate* workers for undesirable aspects of cities, such as greater crime, congestion, pollution and housing costs.[1] In this framing, internal migration flows and wages adjust across cities until quality of life (perceived utility) is equalized across urban and rural residences for a given worker skill level (Rosen, 1979; Roback, 1982; Hwang et al., 1992).

The second set posits that urban workers earn more because they represent a non-randomly *selected*, more-productive (currently and/or in the future) subset of all workers (Fuchs, 1967; Combes et al., 2008; Matano and Naticchioni, 2012). Those with bundles of attributes that can earn them a greater prospective rate of return (net of moving) in the city will selectively move from rural areas (Borjas et al., 1992; Fielding, 1992; Champion et al., 2014; Gordon, 2015; Gordon et al., 2015). Implicit in this framing is that adults' fundamental intellect, attitudes, aptitudes and skills are, to a large extent, fixed once formal education or training has been completed. Urban employment draws on and rewards these skills but does not substantially enhance them.

The third set argues that urban workers are paid more because firms there are more *efficient* due to "agglomeration economies" (Kim, 1987; Ciccone and Hall, 1996; Glaeser, 1998; Rosenthal and Strange, 2004; Puga, 2010). These economies associated with higher densities are viewed as increasing proximity to customers, suppliers, and workers, thereby reducing costs of various types. Density may also enhance technological and intellectual spillovers among firms, specialization and competition among firms, and the ability to share expensive facilities (Duranton and Puga, 2004; Wixe, 2015).

The final set claims that cities make workers more *productive*, hence they command higher earnings. These labor productivity gains might arise from externalities in learning through social networks and human capital production on-the-job (Rauch, 1993; Glaeser, 1999; Moretti, 2004)[2] and in the broader urban milieu (Gordon et al., 2015). They also could arise because urban areas provide superior opportunities for firm-worker coordination—matching people with jobs that complement their skill sets (Kim, 1990; Helsey and Strange, 1990; Sato, 2001)—and denser social networks through which employment information is transmitted (Granovetter, 1995). Finally, workers in a larger labor market could benefit from shared complementary resources (Rosenthal and Strange, 2004; Duranton and Puga, 2004).

The efficiency and productivity theoretical framings provide the foundation for the prediction that urbanization increases rural-urban earnings differentials, but with important differences. The efficiency view predicts that workers would receive an immediate, persistent boost in earnings (the "static premium") when moving to the city, whereas the productivity view predicts that their real earnings not only would be higher at the point of migration but also would grow the more urban experience they accumulated (the "dynamic premium") (Duranton and Puga, 2004; Glaeser and Maré, 2001; Yankow, 2006).

2.1.2. Heterogeneity in urban wage premiums

The efficiency and productivity framings predict larger static and/or dynamic urban wage premiums on the bases of higher skill or education levels of urban migrants (Duranton and Puga, 2004; Rosenthal and Strange, 2004; Bacolod et al. 2009). Davis and Dingel (2019), e.g., posit reductions in the costs of exchanging ideas as the prime agglomeration force and develop a spatial equilibrium theory of why static skill premiums are higher in larger cities. Wheeler's (2006) formulation suggested that better productivity via urban firm-worker matching disproportionately benefitted those with more specialized skills by pairing them with the more efficient firms, with resulting increases in inequality of earnings among urban workers both at the time of hiring and over time.

What is less clear in this body of theory is how returns from rural-urban migration differ on the basis of gender, independent of education and specialized skills. There are at least three plausible reasons why females might evince smaller initial premiums and/or lower growth in premiums over time. First, males and females on average may place

[1] Of course, this compensation may move wages in the opposite direction in the context of an amenity-rich city. Lee (2010), for example, shows that if variety of consumption activities is a luxury good that highly skilled urban residents might actually incur an urban wage penalty.

[2] Although these externalities may vary across education or even university specializations (Liu, 2017).

different relative weights on family proximity, educational credentials, occupational aspirations, rural amenities, etc. compared to pure economic gains and thus manifest distinctive patterns of selection into migration from rural areas. For instance, some women may sublimate their own career aspirations to follow their partners' urban job opportunities (Mulder and van Ham, 2005; Nisic, 2017). This implies that females moving from the countryside may exhibit less selection on highly remunerated observed and unobserved attributes compared to males. Second, females are more likely to have interruptions in their work careers related to child-bearing (Altonji and Blank, 1999), and thus may be reap smaller advantages from the learning and coordination aspects of urbanized labor markets than male counterparts (Phimister, 2005). Third, females may exhibit less-developed urban social networks through which employment and other business information is conveyed (Rosenthal and Strange, 2012; Bacolod, 2017).

However, there are also at least three compelling reasons why females' static and/or dynamic premiums from urbanization might be greater than males'. First, females and males likely exhibit different distributions of physical, intellectual and interactive skills that may differentially affect their productivity through matching, learning, and input-sharing after they migrate to cities (Duranton and Puga, 2004; Rosenthal and Strange, 2004). If urban employers systematically prefer intellectual and interactive skills over physical skills because they enhance these productivity-enhancing processes, and female migrants generally exhibit a comparative advantage in these attributes, differences in premiums favoring females could result (Bacolod, 2017). Second, females may benefit more than males from enhanced job matching associated with thicker urban labor markets, given their more spatially confined job search patterns and heavier reliance on finding proximate child care as a prerequisite to working (Phimister, 2005). Third, rural employers may exhibit less efficiency in matching females with available jobs than males, insofar as they downgrade female applicants' observable and unobservable skills due to their gender stereotypes about occupational roles or sexist biases (Bacolod, 2017). Such rural-urban differences in employer behaviors could result in distinctive trajectories of promotion, retention and wage raises on the basis of gender that yield superior static and dynamic premiums for females when they migrate.

Finally, the compensating differences perspective on rural-urban wage differentials offers another potential framework for predicting gendered differences, though again with no predicted direction. If rural males and females differ systematically on their preferences for salient rural and urban amenities and disamenities, they will require different degrees of compensation to induce a move to the city. For example, if rural females were to disproportionately value quiet, uncongested, low-pollution, safe environments compared to rural men, they would require a larger urban wage premium before being convinced to move to the city.

There has been little theorizing about whether the interaction between education and gender modifies the urban earnings premium. Frank (1978) suggests that thin labor markets in rural areas produces more mismatched workers who are "over-educated" compared to the requirements (and earnings) of the job, and that women are more likely to hold such positions since they traditionally have been assigned to secondary breadwinner status (especially in rural areas). In such circumstances, more highly-educated women should evince larger static and dynamic returns from rural-urban migration than less-educated ones, and perhaps than highly-educated males (who were less mismatched in rural areas than their female counterparts).

In sum, there are strong theoretical reasons from the efficiency and productivity perspectives to posit that initial level and growth of earnings arising from migration from rural to urban areas will be larger for those with higher educational credentials. No clear predictions can be made *a priori* in this regard, however, in the case of gender. Females may exhibit smaller premiums than comparably educated males if their: (1) traits encouraging selection into cities are not as highly remunerated; (2) on-the-job learning is more disrupted by child-bearing; and/or (3) urban job-finding networks are less developed. On the other hand, females may exhibit larger premiums than comparably educated males if urban labor markets: (1) favor intellectual and interactive skills over physical ones; (2) are thicker in job and child care options; and/or (3) are less gender-biased in hiring, occupational assignment, promotion and retention than rural labor markets. Finally, undetermined female-male differentials in premiums could arise from gendered differences in preferences for the amenity packages associated with rural and urban life.

2.2. Evidence

2.2.1. The sources and magnitudes of urban wage premiums

In overview, one finds support in the econometric literature for all of the aforementioned four sources of urban earnings premiums; the debate revolves around the relative importance of particular sources. The literature consensually concludes that an urban earnings premium persists even after cost-of-living differences have been taken into account, though uncertainties in the degree of amenity compensation in wages remain given the endogeneity of selective migration by skills and preferences, housing prices, and wages (Gyourko et al., 2013). Many studies further indicate that the predominate share of the cost-of-living-adjusted urban earnings premium can be explained by more productive individuals selecting to migrate to cities (Glaeser and Maré, 2001; Gould, 2007; Combes et al., 2008; Mion and Naticchioni, 2009; Andersson et al., 2014; Eeckhout et al., 2014; Carlsen et al., 2016; Korpi and Clark, 2019).[3] After accounting statistically for non-random selection into cities based on time-invariant individual characteristics, the literature indicates that from a 2%–6% premium can generally be attributed to efficiency and/or productivity gains in urban areas, depending on how categories of urban areas are defined and the national context (Korpi and Clark, 2019). There is also widespread agreement that migrants from rural to urban areas receive both immediate, static gains and longer term, dynamic gains, though their relative magnitudes are subject to dispute (Glaeser and Maré, 2001; Wheeler, 2006; Gould, 2007; Baum-Snow and Pavan, 2012; De la Roca and Puga, 2017; D'Costa and Overman, 2014; Steskal, 2015; Matano and Naticchioni, 2016; Wang, 2016; Carlsen et al., 2016; Korpi and Clark, 2019).

Several efforts have attempted to parse the contributions of "learning" vs. "coordination/matching" mechanisms for enhancing worker productively in cities. Analyses of how wages grow over time in cities favor learning as the preferred explanation (Glaeser and Maré, 2001; Baum-Snow and Pavan, 2012; D'Costa and Overman, 2014; Carlsen et al., 2016; De La Roca and Puga, 2017). Models examining changes in wages between and within jobs tend to support the importance of better coordinated labor-market matching (Yankow, 2006; Wheeler, 2006; Andersson et al., 2007; Bleakley and Lin, 2012), especially among better-educated workers (Matano and Naticchioni, 2016; Carlsen et al., 2016; Korpi and Clark, 2019).

2.2.2. Heterogeneity in urban wage premiums

Fewer studies have investigated the heterogeneity of returns from urban migration, with most focused on educational attainments or skills.[4] There is general agreement[5] that there are larger earnings premiums for better-educated individuals (Möller and Haas, 2003; Wheeler, 2001, 2006; Gould, 2007; Glaeser and Resseger, 2010; Baum-Snow and Pavan, 2012; Gordon, 2015; Carlsen et al., 2016) and those who have superior non-routine, cognitive and interactive ("people") skills

[3] Although Phimister (2005), Baum-Snow and Pavan (2012), Combes et al. (2015), De la Roca and Puga (2017) and Wessel and Magnusson Turner (2020) find that selection plays only a modest role.

[4] An exception is Ananat et al. (2018), who identify lower dynamic urban wage premiums for Blacks than Whites in US metro areas.

[5] Though see the dissenting work of Korpi and Clark (2019).

(Bacolod et al., 2009; Andersson et al., 2014; Bacolod, 2017). This pattern of selective dynamic returns leads to greater earnings inequality overall within cities (Carlsen et al., 2016).

Issues of gender figure less prominently in studies of urban–rural earnings premiums, and results and attributed causal factors are mixed.[6] Phimister (2005) uses British Household Panel Survey data to estimate longitudinal earnings models jointly stratified by gender and rural/-urban residence.[7] He finds that females' static urban earnings premium is roughly twice than of men's, although returns to urban experience (the dynamic premium) are lower for women. He concludes that improved job matching in denser urban areas of Britain plays an especially important role for women compared to learning spillover effects. Combes et al. (2017) employ data from household income surveys to estimate a cross-sectional model of earnings in major Chinese cities. They also find that female workers evince slightly higher hourly wages and earnings in higher-density, larger-land area cities—a result they attribute to better job matches—though they cannot distinguish static and dynamic premiums.[8] Bacolod (2017) investigates gender wage gaps associated with agglomeration differences across U.S. metropolitan areas using aggregate census data combined with skill requirements associated with different occupations. Unlike the prior authors, she finds that individual females reap *inferior* earnings payoffs from agglomeration compared to males with the same skills, a result she attributes to weaker networks, gendered differences in household division of labor, and/or labor market discrimination. De La Roca and Puga (2017) similarly observe in their fixed-effects modeling of geographically specific experience-earnings profiles in Spain that females' elasticity of wage gains associated with city size was less than half that of males, though education was not controlled. Wessel and Magnusson Turner (2020) specify a structural equation model of internal migration, education, job change, and earning ranks, operationalized with rich, administrative register data on a cohort of Norwegian individuals. Their gender-stratified, fixed-effects estimators show that gender differences in payoffs from migration between less-to more-urbanized areas are more nuanced: they depend on temporal and geographic context. Moving from rural areas before completing education yields significantly larger relative earnings gains for women than men—not primarily because they are paid more for doing the same job, but through increased work hours and changed industrial sector. Migration after completing education, on the other hand, benefits males slightly more. Migration from rural areas to modest-scale, tier 2 cities benefits males more than females, which the authors attribute to the booming petroleum industries there that traditionally favor male, high-wage employment. Relative advantages for females are especially large in tier 1 Oslo, however, which the authors attribute to its quintessential post-industrial structure favoring cognitive and interactive skills and a well-developed transport system that facilitates better female labor force participation and job matching. In all cases, most of the gendered differences emerge immediately after relocation (the static premium), as opposed to different earnings growth patterns subsequently.

We could identify only one study that investigated heterogeneity jointly by gender and education: a recent working paper by Bennett et al. (2022) that used individual-level Norwegian data from 1967–2017.[9] Unfortunately, they only explored joint heterogeneity for the raw, unconditional urban premium, finding that over the period of analysis the gender gap switched from favoring males to favoring females, though this was due primarily to the larger declines in raw premiums for rural, low-education males compared to their female counterparts since 2000. By contrast, their stratified fixed-effects estimates of the selection-adjusted static urban wage premium only compared males and females, and found different gender differentials depending on when migration occurred: before 1995 females' premiums were greater than males'; this reversed during 1995–2009; and they were the same from 2010 to 2107.

2.3. Challenges of measuring the earnings returns from rural-urban migration

A critical evaluation of the aforementioned empirical literature necessitates discussion of a core methodological issue that undergirds it. Following the seminal work of Glaeser and Maré (2001), the common strategy for identifying the urban wage premium has been to use panel data on individuals and specify worker fixed-effects (FE) to control for time-invariant, unobserved characteristics that may affect both migration and earnings (Yankow, 2006; Baum-Snow and Pavan, 2012; D'Costa and Overman, 2014; Andersson et al., 2014; Steskal, 2015; Matano and Naticchioni, 2016; Wang, 2016; Carlsen et al., 2016; De la Roca and Puga (2017); Korpi and Clark, 2019; Wessel and Magnusson Turner, 2020).[10] The Difference-in-Differences (DiD) estimator has also been used (Steskal, 2015). Years of employment experience is also controlled in most models, typically differentiated by rural and urban experience so the dynamic premium can be identified.

The FE and DiD estimators produce consistent, unbiased estimates under the assumptions that unobserved worker characteristics do not influence intertemporal earnings profiles, ability is equally valued at the margin by employers in rural and urban locales, and mobility is random conditioned on the fixed effect (Verbeek and Nijman, 1992). These assumptions seem especially tenuous, given what we know about migration, work, and agglomeration (Yankow, 2006; Baum-Snow and Pavan, 2012). For example, it is likely that rural individuals' decision to migrate will be influenced by not only the level at which they are being paid but also how much their experience is being valued. In the case of females,

[6] On the contrary, a common practice is to exclude all but white males from analysis to reduce sample heterogeneity (Glaeser and Maré, 2001; Yankow, 2006; Baum-Snow and Pavan, 2012; D'Costa and Overman, 2014; Steskal, 2015; De La Roca and Puga, 2017). This often the case in U.S.-based work because one of the few longitudinal surveys with appropriate data for this investigation, the National Longitudinal Survey of Youth, is based on a predominantly white, male sample. Carlsen et al. (2016) and Korpi and Clark (2019) observe a fixed positive effect on wages of being male (and native-born, as per Korpi and Clark, 2019), but do not test for gender- or immigrant-differentiated returns to urban migration, experience or tenure in job. Bacolod (2017) and Nisic (2017) find that male/female earnings gaps are reduced in larger metropolitan areas in the US and Germany, respectively. Though this evidence is consistent with the hypothesis of a greater female urban premium, it is not definitive because there may be inter-gender differences in the distributions of observable and unobservable skills across different levels of urbanization. Studies of internal migration and occupational achievement more often pay attention to gender (see review in Fielding, 2007) but do not employ rigorous statistical approaches to avoid selection bias.

[7] He compares estimates from OLS, fixed-effect, and a new selection control method and finds that they yield few substantial differences in the measured urban earnings premium for either gender.

[8] Note, however, that they do not identify the impacts of a rural-urban migration, nor do they control for worker education, skills or other unobserved heterogeneity in the pools of male and female workers.

[9] The results from Bennett et al. (2022) are not strictly comparable to either ours or Wessel and Magnusson Turner's (2020) because they use a significantly different operationalization of "urban" than we. They include hinterland towns in the range of 15,000–49,999 population as "urban." We exclude observations from such places from either our rural or urban samples, as we believe their urban-rural status is ambiguous.

[10] Sometimes the individual FE estimator is augmented with controls for observable measures of underlying ability such as educational credentials or test scores (Yankow, 2006), or proxies thereof such as parents' education (Ahlin et al., 2014; Wessel and Magnusson Turner, 2020). Other methods for dealing with selection have included family fixed effects and difference-in-differences (Steskal, 2015), education-level fixed effects (Carlsen, et al., 2016), experience fixed effects (Yankow, 2006), Heckman two-stage correction combined with course exact matching (Korpi and Clark, 2019), panel sample selection (Phimister, 2005) and instrumental variables (Glaeser and Maré, 2001).

this decision could be shaped by child-bearing, which also could affect earnings profiles (Phimister, 2005). Even if migration were exogenous, further difficulties for FE and DiD estimators arise if individuals experience different growth rates in: (1) some unmeasured productive attribute(s); (2) earnings payoffs from a fixed unobserved attribute (e.g., "ability to network"); or (3) earnings profiles before and after migration (as would be expected with the "learning" and "coordination" aspects of agglomeration). For a graphic illustration of how these factors can lead to biased estimates from FE and DiD estimators, see Appendix A.[11]

A final challenge in estimating the urban earnings premium relates to potentially heterogeneous treatment impacts (De la Roca and Puga, 2017). If earnings effects of urbanization are heterogeneous in worker characteristics (as they have proven to be in education, e.g.), the FE estimator would not consistently recover the mean premium (Baum-Snow and Pavan, 2012). Only a few studies have addressed this challenge, mostly via stratifying FE models by two or three education or skill levels (Baum-Snow and Pavan, 2012; Andersson et al., 2014; Matano and Naticchioni, 2016; Carlsen et al., 2016; Korpi and Clark, 2019).[12] Only Phimister (2005) and Wessel and Magnusson Turner (2020) stratify by gender. To our knowledge, no study in this literature has jointly stratified by gender and education, as we do.

2.4. Our contribution

In sum, previous investigations into the earnings impacts of rural-urban migration exhibit shortcomings in internal and external validity. The assumptions required for FE and DiD models to yield unbiased and consistent estimates are implausible and there have been inadequate efforts to test for generality beyond males. In this paper, we offer a superior identification strategy and apply it to specific tests of heterogeneity via joint sample stratification by gender and education.

We employ an augmented DiD estimator that controls for mover-stayer differences in not only pre-move *level* but also *trend* of earnings (DiD-TR). Though this model has often been used in urban economics (e.g., Greenstone et al., 2010; Ahlfeldt et al., 2017) and program impact analyses (e.g., Galster et al., 1999, 2002, 2004, 2006), it has rarely appeared in the urban wage premium literature.[13] This specification assumes that all unobserved time-varying and invariant characteristics of individuals that jointly influence both their propensity to move from rural to urban areas of Norway and their earnings trajectories will be captured by any pre-move differences in *either the level and/or the trend* in earnings between the treatment group (rural-urban migrants) and the control group (rural non-movers). A plausibly causal measure of the rural-urban move's impact on earnings is indicated if either there is a post-treatment (move) difference in the pre-treatment differences in control/treatment group *levels* of earnings (*the static urban wage premium*) or in *trends* of earnings (*the dynamic urban wage premium*). Parameters are identified by rural-urban migrants in our application. For a graphical comparison and evaluation of FE, DiD and DiD-TR approaches, see Appendix A.

We will employ DiD-TR model to answer the following research questions:

1. What is the earnings premium that can be causally attributed to a move during young adulthood from rural Norway to one of its four major metropolitan areas ("urban areas," hereafter), and is any premium manifested as an initially higher level of (static) and/or enhanced growth (dynamic) in earnings?
2. Do these static and dynamic earnings premiums differ across groups jointly distinguished by gender and educational attainment?

3. Empirical approach and data

3.1. Analytical strategy

In order to ascertain whether the DiD-TR model retrieves substantially different estimates of the earnings premium from rural-urban migration compared to conventional specifications, we first establish a baseline by estimating with OLS the premium with no controls for selection, then the premium with DiD and FE controls for selection. The differences in mean earnings and the value of experience between workers in rural areas and those with similar observed characteristics who have moved from rural to urban areas will be revealed in the OLS model:

$$\boldsymbol{E}_{it} = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \boldsymbol{TR}_{it} + \boldsymbol{\beta}_1 \boldsymbol{D}_{it} + \boldsymbol{\beta}_2 \boldsymbol{TR}_{it} \cdot \boldsymbol{D}_{it} + \boldsymbol{\psi}_k [\boldsymbol{Z}_{ki}] + \boldsymbol{\varphi}_j [\boldsymbol{Z}_{jit}] + \lambda [\boldsymbol{Y}_t] + \boldsymbol{\varepsilon}_{it} \quad (1)$$

where:

$\boldsymbol{E}_{it}$ = ln (inflation-adjusted annual earnings from full-time work of individual i during year t).

$\boldsymbol{TR}_{it}$ = 1 if first year of earnings by the individual post-school, = 2 if second year; 3 if third year, etc. (i.e., a time trend over the entire analysis period measuring labor market experience).

$\boldsymbol{D}_{it}$ = 1 if observed earnings occur during the year of individual i's rural-urban migration or after; zero otherwise

$[\boldsymbol{Z}_{ki}]$ = a vector of k observed, time-invariant characteristics of individual i

$[\boldsymbol{X}_{jit}]$ = a vector of j observed, time-varying characteristics of individual i during year t

$[\boldsymbol{Y}_t]$ = year fixed effects[14]

$\boldsymbol{\varepsilon}_{it}$ = a random error term with usual assumed i.i.d. properties

Parameters $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ from equation (1) will not provide unbiased estimates of the static or dynamic premiums in mean earnings and experience, respectively, caused by rural-urban migration since model (1) does not control for any observed or unobserved individual characteristics that may be jointly related to migration selection and earnings.

The DiD model can be expressed:

$$\boldsymbol{E}_{it} = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \boldsymbol{TR}_{it} + \boldsymbol{\beta}_1 \boldsymbol{D}_{it} + \boldsymbol{\beta}_2 \boldsymbol{TR}_{it} \cdot \boldsymbol{D}_{it} + \boldsymbol{\alpha}_2 \boldsymbol{T}_i + \boldsymbol{\psi}_k [\boldsymbol{Z}_{ki}] + \boldsymbol{\varphi}_j [\boldsymbol{X}_{jit}] + \lambda [\boldsymbol{Y}_t] + \boldsymbol{\varepsilon}_{it} \quad (2)$$

where:

$\boldsymbol{T}_i$ = 1if individual i is in the "urban treatment" group, i.e., undertakes a rural-urban move, else 0

Parameters $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ from equation (2) will represent the unbiased static and dynamic premiums under the assumption that differences in unobserved characteristics distinguishing movers and non-movers only affect the level of earnings and that this difference $\boldsymbol{\alpha}_2$ is constant over time (both pre- and post-move).

The FE model can be expressed:

$$\boldsymbol{E}_{it} = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \boldsymbol{TR}_{it} + \boldsymbol{\beta}_1 \boldsymbol{D}_{it} + \boldsymbol{\beta}_2 \boldsymbol{TR}_{it} \cdot \boldsymbol{D}_{it} + \boldsymbol{\delta}_i + \boldsymbol{\varphi}_j [\boldsymbol{X}_{jit}] + \lambda [\boldsymbol{Y}_t] + \boldsymbol{\varepsilon}_{it} \quad (3)$$

where:

$\boldsymbol{\delta}_i$ = the fixed effect for individual i; a proxy for time-invariant observed and unobserved characteristics

Parameters $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ from equation (3) will represent the unbiased

[11] Also see the discussion in De la Roca and Puga (2017).

[12] Glaeser and Maré (2001) experiment with education - urban residence interaction terms.

[13] To our knowledge, it has only been employed by De la Roca and Puga (2017: 130–131) as a robustness check. In the urban program impact evaluation literature this specification has been labelled the Adjusted Interrupted Time Series model (Galster et al., 2004, 2006).

[14] Note the $\mathbf{D}_{it}$ variables are not equivalent to single-year fixed effects $\lambda[\mathbf{Y}_t]$ because the former denote unique multi-year periods for each individual.

static and dynamic premiums under the assumption that differences in unobserved characteristics distinguishing movers and non-movers only affect the level of earnings and that this difference δ_i is constant over time (both pre- and post-move).[15]

Of course, these DiD and FE estimates are subject to all the limitations noted above. Thus, we specify our core model as DiD-TR:

$$\begin{aligned} E_{it} = {} & \alpha_0 + \alpha_1 TR_{it} + \alpha_2 T_i + \alpha_3 T_i \cdot TR_{it} + \beta_1 D_{it} + \beta_2 TR_{it} \cdot D_{it} \\ & + \psi_k [Z_{ki}] + \varphi_j [X_{jit}] + \lambda [Y_t] + \varepsilon_{it} \end{aligned} \tag{4}$$

Unlike the traditional DiD and FE models previously employed to measure urban earnings premiums, which must assume parallel trends in treated and untreated groups' earnings before the treatment, DiD-TR explicitly estimates and controls for such pre-treatment trends. This is the central distinction and advantage of our specification that, as we will demonstrate below, yields radically different measures of the urban earnings premium. Parameters β_1 and β_2 from equation (4) will represent the unbiased static and dynamic premiums under the assumption that differences in unobserved characteristics distinguishing movers and non-movers (after controlling for observed time-varying and -invariant characteristics) will be fully reflected in differences in their overall level of earnings and/or pre-migration linear trend in earnings (measured by α_2 and α_3, respectively), and that these differences remain constant over time (pre-move). The earnings trend (mean annual return on experience) while employed in rural areas is measured by α_1 for those who do not migrate and by $(\alpha_1+\alpha_3)$ for those who eventually migrate. The comparable return on experience in urban areas for those who migrate is $(\alpha_1+\alpha_3+\beta_2)$..

To answer our first research question, the DiD-TR measures of static and dynamic urban earnings premiums (level and growth) are provided by coefficients β_1 and β_2, respectively. To address the second research question of heterogeneous treatment effects from rural-urban migration, we stratify our sample jointly by education and gender and compare premium estimates across them. We stratify by education not only because we suspect heterogeneous treatment effects, but also given that Carlsen et al. (2016) find that Norwegian migration selection on unobserved ability is stronger for better-educated workers (especially when they are younger, as during our period of analysis), and thus we expect heterogeneity in α_2 and α_3.

3.2. Data and variables

3.2.1. Data sources and sample

In this study we use anonymized annual data about individuals that we assembled from various governmental social registers made available for us by Statistics Norway.[16] Our dataset has unusual breadth (encompassing all residents), length (years of observations), and comprehensiveness (variety of economic, demographic, educational and geographic information). Specifically, we selected the cohorts born in any municipality in Norway in 1979, 1980 and 1981, the people for whom we have the longest panel of complete annual information from1995[17] through the most recent register year available, 2018 (N = 153,327). Within these cohorts, we identified those who met these baseline criteria: (1) lived in Norway all years from birth through 2018; (2) were registered in the social welfare system all years from age 16 (when compulsory schooling completed) through 2018 (N = 139,622).[18]

Our focus is on the marginal earnings payoffs associated with the subset of these individuals who moved from rural to urban areas in Norway during their young work lives while working full-time.[19] We operationalize rural and urban using a centrality scheme developed by Høydahl (2017, 2020) and adopted by Statistics Norway to identify a six-fold hierarchy of all 423 municipalities in Norway.[20] The centrality index is based on how many jobs and service institutions one can reach within 90 min from each basic statistical unit (grunnkrets), aggregated to the municipal level, and adjusted for population size (Høydahl, 2017, 2020).

Oslo is the most central (class 1) municipality in Norway, with and index value of 1000 and a population of 673,469 in 2018. Some other municipalities surrounding Oslo are also characterized as class 1. One example is Bærum, which has a population of 12,454 in 2018 and a centrality index of 967 (Høydahl, 2017). Outside the Oslo area, the cities of Stavanger (902), Bergen (900) and Trondheim (894) are the most important class 2 municipalities (centrality index values in parentheses). The number of inhabitants in these cities are 141,186; 279,792; and 193, 501, respectively. The least central (class 6; index 295) municipality in Norway is Utsira, an island in Western Norway, which had a population of 208 in 2018. The most central municipality among the rural areas is Sør-Odal, a class 4 (index 769) municipality, with a population of 7884 in 2018, located in Eastern Norway around 75 km driving distance from central Oslo.

In this paper we define a rural municipality as centrality level 4, 5 or 6. We study moves from these municipalities to the most central (urban) municipalities, which is either class 1 or 2, depending on the model. Note that the urban centrality classes do not necessarily consist of coherent areas, such as a travel-to-work areas. For this reason, we also test robustness of our results by restricting the sample to moves to urban level 1 and 2 municipalities surrounding the capital, Oslo, which is a spatially coherent travel-to-work area. People who reside in or move to class 3 municipalities (62,295 observations) are excluded from the sample because of the ambiguity of whether to classify their centrality index as urban or rural.

Our sample starts with individuals from the 1979, 1980 and 1981 cohorts meeting the above baseline criteria who resided in a rural area at age 16 (i.e., 1995–1997), the earliest age after they complete compulsory schooling and can legally begin work (N = 30,459). From this group we analyze in our core model only those who after completing their education work full time, so we can focus on the pure earnings premium effect of urbanization purged from associated changes in hours worked (N = 28,473).[21] We observe them annually from ages 16 through ages 37–39, respectively (i.e., 2018). For each of these years we measure their

[15] In the most sophisticated application of the FE approach, De la Roca and Puga (2017) allow the value of experience TR to vary not only by urban-rural, but by metropolitan areas of different sizes and by where the person is currently employed. Nevertheless, their specification cannot control for unobserved heterogeneity in the rural earnings profiles of those who will become migrants to larger cities. As they note, the coefficient of experience is identified in their model by profiles of both movers to cities and those who have always resided there, whereas in the AITS model it is identified only by movers. Put differently, their specification cannot control for migration selection based on unobservables influencing the pre-move value of experience.

[16] These data are proprietary and can be accessed only by designated researchers using secure computers.

[17] Information about whether an individual worked fulltime only became available in 1995; since such work can legally begin at age 16 we could not select any younger cohorts for fear of censoring their earlier work experiences.

[18] These criteria excluded those who died between 1979 and 2018, emigrated abroad, lived abroad at least one year and returned, or were missing registration status.

[19] In the employment registers, planned full-time is defined as averaging 30 h or more of work weekly.

[20] The structure of municipalities changes over time, mainly via merging of municipalities. Definitions of centrality levels back in time are based on table 11727 from Statistics Norway.

[21] Unfortunately, Norwegian registers do not contain information on the hourly wage or precise number of hours worked per year in our study period (only three broad categories, defined as planned number of working hours).

educational and employment status, earnings (if any), and a variety of demographic and household characteristics. To operationalize DiD-TR, we need sufficient observations to establish earnings trends reliably both before and after migration to an urban area. This implies that we can only analyze movers once they have completed their formal schooling and have entered the labor force on a regular basis (as in Yankow, 2006), *and* have worked at least two years in rural areas post-school and at least two years in an urban area after they moved there. We believe that this sacrifice of some external validity is justified by the improved internal validity provided by the DiD-TR model. In sum, those rural residents (at age 16) who since completing their formal schooling have worked full-time at least two years in *both* rural areas and urban areas by ages 37–39 (depending on cohort, implying that we observe their migration between 1997 and 2016) comprise our "treatment group" (N = 5613).[22] Rural residents (at age 16) who since completing their formal schooling have never worked in urban areas by ages 37–39 comprise our "control group" (N = 22,860). Removal of 102 cases with zero values recorded for full-time earnings produced our final sample size of 28,391 individuals.

Since stratification by educational attainment is a key component of our analytical strategy, a brief overview of the Norwegian educational system is appropriate. Children between the ages of 6 and 16 must be enrolled in elementary education. Youth between the ages of 16 and 20 are entitled to (but not required to attend) free secondary schooling of up to three or four years in one of the college preparatory or vocational/technical education tracks. Post-secondary education requires a college preparatory secondary diploma but public colleges and universities are tuition-free. By and large, grants and loans are provided by the Norwegian State Educational Loan Fund to those accepted to any accredited institution. The rates of graduation from secondary schools and from colleges in Norway are comparable to those for similar cohorts in the U.S. (OECD, 2020). In our analysis, we stratify our sample into "lower" and "higher" attainment categories using the approximate median attainment for our cohorts: those who have only secondary school diploma (typically 13 years of school) or less vs. those with supplemental secondary or college education (typically 14 years of school or more).

3.2.2. Variables

Our dependent variable is inflation-adjusted (base 2015), pre-tax annual earnings from wages, plus taxable transfers associated with parental leave or sick leave that are associated with being employed. We transform this sum by the natural logarithm, as is conventional given positive skew.[23] As shown in Table 1, the mean annual earnings associated with fulltime work during young adulthood differ in the expected ways across geographic, gender and educational groups within our analytical sample. Rural-urban movers earn 18% more annually than those who remain in rural areas, on average. Those with above-median educational attainment earn more than their less-educated counterparts: 31% for males; 33% for females. There are considerable gendered earnings gaps: males with above-median educational attainment earn 32% more than comparably educated females; males with below-median educational attainments earn 33% more than comparably educated females.

Our key explanatory variables relate to those who move from rural to urban areas and the timing of move, and the comparative earnings levels and trajectories of these movers pre- and post-move. Following equation (3), we specify a dummy indicating the individual is a mover from a rural to an urban area, a dummy indicating a year after that move has occurred, a trend starting at "one" the first year of work post-schooling (i.e., labor force experience), a comparable trend variable but only for movers, and another trend variable starting at "one" the year the urban move occurred.

Covariates measure standard individual and family characteristics in earnings equations. We denote with dummy variables each year whether the individual is: cohabiting or married; divorced or widowed; single, never married is the reference category. We specify a dummy indicating the year a child was born and the succeeding year, as a proxy for earnings disruptions associated with taking parental leave (which is generously supported for both parents in Norway). We also denote the first calendar year when an individual takes a job after completing their education to account for the likelihood of a partial work year post-graduation. We control for macroeconomic conditions affecting the Norwegian labor market with calendar year fixed effects. Gender and education is addressed via sample stratification, though in preliminary models we specify dummy variables denoting officially defined Norwegian educational attainment categories[24]: secondary (13–14 years); university degree/low (14–17 years); university degree/high (18–19 years); PhD degree (21+ years of education); those with less than a secondary school diploma are the excluded reference category.

4. Results

4.1. Different model specifications yield substantially different results

One of the central claims of this paper is that the internal validity of conventional models (i.e., FE, DiD) for measuring urban earnings premiums are vulnerable if movers' unmeasured attributes correlated with moving produce different earnings trajectories, not merely different levels, both *before* and after the urban migration. We find this vulnerability clearly manifested in our results. Table 2 compares the OLS parameters estimated for the naïve (selection uncontrolled), FE, DiD and DiD-TR models, using the full, non-stratified sample of fulltime workers described above. Corresponding to these parameters, Fig. 1 provides a graphic portrayal of the estimated earnings profile for a representative mover during work year eight both before and after rural-urban migration, compared to an otherwise-comparable individual who continued working in a rural area. It is noteworthy that the naïve OLS model *understates* the static urban earnings premium compared to any of the alternative specifications that attempt to control for selection of movers. This implies that the rural-urban movers in this particular Norwegian sample were being compensated *less than* other members of their rural cohort before they moved, on average. Indeed, this might well have provided a motivation for their moves out of rural areas. Moreover, it might suggest (regardless of econometric specification) that these migrants were *negatively* selected on their unobserved motivation/ability/skill-related characteristics, which matches the conclusion of a previous study of post-1970 Norwegian rural north-urban south migrants (Steskal, 2015). However, it is also consistent with the oft-observed finding that the overqualified are disproportionately underpaid (Kracke et al., 2018); this could be more likely in rural areas, particularly for those with the highest educational credentials.

The DiD-TR model demonstrates that those who move to urban areas are indeed a select group of rural residents insofar as prior to moving they not only had a substantially *lower* (25 percent) *mean level* of rural earnings (vs. nine percent lower in DiD),[25] but they also had a 3.4 percent *greater annual growth* in rural earnings,[26] which is not revealed by the other models and, indeed must be assumed away for them to have

[22] We do not analyze any individuals who moved back to a rural area after living and working in an urban area (N = 11,064) or who moved several times between these areas (N = 2170).

[23] This operationalization follows Carlsen et al. (2016).

[24] See the NUS2000 standard: https://www.ssb.no/a/publikasjoner/pdf/nos_c617/nos_c617.pdf.

[25] Computed as exp (coefficient of T)-1.

[26] Computed as exp (coefficient of T * TR)-1.

Table 1
Descriptive Statistics, Mean (std. dev.).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Total	Rural	Rural-Urban	Males		Females	
		Stayers	Movers	Lower Educ	Higher Educ	Lower Educ	Higher Educ
Earnings (Norwegian Kroner; 1 NOK = 8.9 USD) (deflated by CPI, 2015 = base year)	443,926 (214,291)	428,144 (195,584)	503,416 (265,229)	438,991 (207,934)	576,514 (280,560)	328,952 (138,003)	437,131 (151,912)
Highest Education Completed is Lower Secondary School = 1; else = 0	0.157 (0.364)	0.178 (0.383)	0.079 (0.270)	0.194 (0.396)		0.093 (0.290)	
Highest Education Completed is Upper Secondary diploma (13–14 years) = 1; else = 0	0.512 (0.500)	0.567 (0.496)	0.304 (0.460)	0.610 (0.488)		0.340 (0.474)	
Highest Education Completed is University, lower degree (14–17 years) = 1; else = 0	0.280 (0.449)	0.226 (0.418)	0.488 (0.500)		0.156 (0.363)		0.500 (0.500)
Highest Education Completed is University, higher degree (18–19 years) = 1; else = 0	0.049 (0.217)	0.028 (0.166)	0.128 (0.334)		0.039 (0.194)		0.067 (0.250)
Highest Education Completed is PhD degree (21+ years) = 1; else = 0	0.000 (0.020)	0.000 (0.016)	0.001 (0.033)		0.000 (0.017)		0.001 (0.026)
Age when complete education (years)	23.350 (4.700)	23.391 (4.923)	23.127 (3.172)	21.599 (3.785)	23.221 (4.497)	21.620 (4.550)	25.914 (4.173)
Year when complete education	2003	2003	2003	2001	2003	2001	2005
Age when move to the city (movers only)			27.457 (3.102)	26.801 (3.628)	27.915 (3.087)	25.193 (3.20)	27.2145 (2.754)
TRit (years work experience after completing school)	16.126 (4.548)	16.128 (4.882)	16.118 (3.108)	18.381 (3.802)	13.206 (3.902)	18.212 (4.140)	13.840 (3.800)
Ti (rural-urban mover = 1; else = 0)	0.210 (0.407)	0.000 (0.000)	1.000 (0.000)	0.106 (0.308)	0.199 (0.399)	0.159 (0.366)	0.276 (0.447)
Dit (year residing in urban area = 1; else = 0)	0.162 (0.368)	0.000 (0.000)	0.772 (0.420)	0.073 (0.261)	0.150 (0.357)	0.123 (0.328)	0.223 (0.416)
Years work experience in rural areas (movers only)			6.759 (3.810)	9.0737 (4.146)	5.682 (2.961)	7.286 (3.649)	4.889 (2.569)
Years work experience in urban areas (movers only)			10.581 (3.384)	11.412 (3.938)	9.766 (2.946)	12.454 (3.906)	10.437 (3.019)
Cohabiting/married status during year = 1, else = 0	0.518 (0.500)	0.527 (0.499)	0.485 (0.500)	0.434 (0.496)	0.485 (0.500)	0.579 (0.494)	0.633 (0.482)
Divorced/Widow (er) during year = 1, else = 0	0.013 (0.116)	0.014 (0.120)	0.010 (0.099)	0.011 (0.105)	0.010 (0.098)	0.023 (0.151)	0.019 (0.137)
Child born this or preceding year = 1, else = 0	0.182 (0.386)	0.185 (0.389)	0.169 (0.375)	0.169 (0.375)	0.186 (0.389)	0.160 (0.366)	0.193 (0.394)
N of individuals	28391	22792	5599	11024	4765	5572	7026

Note: N = Full-time, native workers born in Norway 1979–1981 (see text for other sample restrictions); urban movers to either 1st *or 2nd tier cities.*

internal validity.[27] If, like De La Roca (2017), we assume that earnings rank in the local income distribution serves as a proxy for unobserved productivity, the lower mean rural earnings is consistent with the aforementioned evidence of negative selection on this attribute. Of course, as Baum-Snow and Pavan (2013) and Wessel and Magnusson Turner (2020) remind us, FE and DiD are imperfect proxies for unobserved individual characteristics that are associated with *changes* in personal productivity over time, such as internal locus of control or drive to succeed. At a deeper level, there may even be a causal relationship between unobserved fixed and time-varying individual attributes that may sometimes lead to different conclusions about productivity. As illustration, suppose that urban migrants systematically possess some attribute(s) that employers can observe (but researchers cannot, such as physical appearance or sociability) and use such as justification for allocating them to somewhat lower-paying jobs. These workers, in response to their frustration over being undervalued, may search more actively for alternative jobs and, indeed, successfully change into better-paying rural jobs more frequently than those who remain in rural areas and are content with their compensation. This interpretation holds considerable empirical validity, as we demonstrate below in section 5 where we explore the nature of this rural-urban selection process in greater depth.

Not surprisingly, the DiD-TR also paints a considerably different picture of the static and dynamic urban earnings premium: a substantial initial gain of 38 percent[28] but with a relative experience *penalty* of 3.4 percent annually *compared to the pre-move trend.*[29] The net result is a one percent real annual earnings growth after a rural-urban move, exactly the same value of experience as reaped by their counterparts who remained in rural areas.[30] By comparison, the DiD and FE models indicate only 23 percent and 25 percent static urban earnings premiums, respectively, with no marginal value from urban experience.[31] The failure of these two conventional specifications to account accurately for the pre-move lower level and upward trajectory of future mover's earnings leads to a distorted view of both the static and intertemporal dimensions of the premium.

Viewed holistically, the distinctive results from our DiD-TR model—greater static but virtually no dynamic premiums—hold implications for the theories of urban wage premiums discussed above. They suggest that the forces providing a one-time boost to earnings upon moving to an urban area—agglomeration economies and the superior

[27] This perhaps explains why this result differs from those of Phimister (2005), Baum-Snow and Pavan (2012), Combes et al. (2015), De La Roca and Puga (2017) and Wessel and Magnusson Turner (2020), who find that selection plays only a modest role. Although Steskal (2015) also found that migrants had a lower mean wage while working in rural areas than their counterparts who stayed there, she did not control for their returns from rural experience.

[28] Computed as exp (coefficient of D)-1.

[29] Computed as exp (coefficient of D * TR)-1.

[30] Computed as exp (sum of coefficients of TR, T * TR, and D * TR)-1.

[31] Baum-Snow and Pavan (2013) similarly find that FE models substantially understate the static premium during the first 15 years of experience, which dominates our analysis period.

Table 2
comparative earnings models estimated by naïve OLS, FE, DiD and DiD-TR specifications.

Variable	OLS	FE	DiD	DiD-TR
Trit (annual change in earnings)	0.010***	0.005	0.010***	0.010***
	(0.001)	(0.011)	(0.001)	(0.001)
Ti (=1 if rural-urban mover; else = 0)			−0.090***	−0.248***
			(0.006)	(0.011)
Ti*Trit (movers' earnings change increment)				0.034***
				(0.002)
Dit (=1 if residing in urban region during current year; else = 0)	0.156***	0.252***	0.228***	0.382***
	(0.006)	(0.009)	(0.008)	(0.012)
Dit*Trit (annual earnings change increment while residing in urban region)	−0.002***	0.001	−0.001	−0.034***
	(0.001)	(0.001)	(0.001)	(0.002)
Gender (=1 if male; = 0 if female)	0.243***		0.243***	0.243***
	(0.002)		(0.002)	(0.002)
Secondary Diploma (13–14 years)	0.288***		0.290***	0.291***
	(0.004)		(0.004)	(0.004)
University, lower degree (14–17 years)	0.376***		0.384***	0.388***
	(0.005)		(0.005)	(0.005)
University, higher degree (18–19 years)	0.529***		0.544***	0.554***
	(0.007)		(0.007)	(0.007)
PhD degree (21+ years)	0.648***		0.671***	0.682***
	(0.030)		(0.030)	(0.029)
1st year of fulltime work after highest education completed	−0.215***	−0.250***	−0.212***	−0.199***
	(0.007)	(0.007)	(0.007)	(0.007)
Married/cohabiting current year	0.051***	−0.017***	0.046***	0.048***
	(0.002)	(0.004)	(0.002)	(0.002)
Separated/Widowed current year	0.038***	−0.029*	0.034***	0.035***
	(0.008)	(0.012)	(0.008)	(0.008)
Child born this or preceding year	−0.024***	−0.030***	−0.027***	−0.027***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	9.929***	10.482***	9.934***	9.933***
	(0.099)	(0.139)	(0.099)	(0.099)
Observations	3,16,733	3,16,733	3,16,733	3,16,733
Individuals	28,391	28,391	28,391	28,391
R-squared	0.305	0.334	0.305	0.306
Year FE	YES	YES	YES	YES
Individual FE	NO	YES	NO	NO

Robust standard errors clustered by individuals shown parenthetically; ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.
Sample includes only full-time, native workers; moves are from rural to urban centrality levels 1 and 2.

Fig. 1. Earnings Trajectories for Rural Stayers and Rural-Urban Movers Estimated by Alternative Models
Note: baseline = earnings during 1st full year of fulltime work after completing education by similar individuals who remain in rural employment; intercept of urban movers in FE model is undetermined. Graphs are based on parameters shown in Table 2.

job-employee match associated with the migration—dominate the forces related to ever-increasing urban worker productivity—learning on-the-job or in the larger urban context, or better intra-urban job matching over time.

4.2. Heterogeneity of the premium by education and gender

4.2.1. Core results

Before turning to the urban earnings premium estimates, Table 3 offers several noteworthy results related to gender inequalities in the Norwegian labor market. First, being married/cohabiting (compared to being single/never married) is associated with higher earnings for males but lower earnings for females. Since we are holding fulltime work constant in our sample, these findings are consistent with the notion that partnered females are more likely to "satisfice" by accepting lower-paying jobs whereas partnered males are less likely to do so. The fact that these differences are especially acute for those with less education suggests that these behaviors may be associated with greater salience of traditional gender occupational roles. It is also reasonable to attribute these findings to endogenous selection into partner status: males

Table 3
comparative earnings models estimated by DiD-TR by gender and education.

Variable	Males			Females	
	Lower Education	Higher Education		Lower Education	Higher Education
Trit (annual change in earnings)	0.016***	0.018***	0.020***	0.013***	
	(0.001)	(0.001)	(0.001)	(0.001)	
Ti (=1 if rural-urban mover; else = 0)	−0.241***	−0.223***	−0.107***	−0.211***	
	(0.022)	(0.020)	(0.030)	(0.016)	
Ti*Trit (movers' earnings change increment)	0.028***	0.037***	0.011	0.036***	
	(0.003)	(0.004)	(0.006)	(0.003)	
Dit (=1 if residing in urban region during current year; else = 0)	0.317***	0.316***	0.239***	0.340***	
	(0.027)	(0.023)	(0.035)	(0.017)	
Dit*Trit (annual earnings change increment while residing in urban region)	−0.026***	−0.033***	−0.009	−0.038***	
	(0.003)	(0.004)	(0.006)	(0.004)	
1st year of fulltime work after highest education completed	−0.027*	−0.402***	−0.024	−0.375***	
	(0.011)	(0.017)	(0.015)	(0.013)	
Married/cohabiting current year	0.099***	0.066***	−0.056***	−0.022***	
	(0.004)	(0.005)	(0.005)	(0.004)	
Separated/Widowed current year	0.033*	0.061*	−0.009	0.021	
	(0.015)	(0.029)	(0.013)	(0.011)	
Child born this or preceding year	−0.012*	(0.011)	−0.077***	−0.077***	
	(0.005)	(0.006)	(0.007)	(0.004)	
Constant	10.107***	11.317***	9.416***	12.007***	
	(0.116)	(0.348)	(0.172)	(0.050)	
Observations	1,50,472	51,207	48,593	66,461	
Individuals	11,024	4769	5572	7026	
R-squared	0.267	0.271	0.360	0.286	
Year FE	YES	YES	YES	YES	
Individual FE	NO	NO	NO	NO	

Robust standard errors clustered by individuals shown parenthetically; ***p < 0.001, **p < 0.01, *p < 0.05.
Sample includes only full-time, native workers; moves are from rural to urban centrality levels 1 and 2.

(females) with lower earnings are more (less) likely to remain single; coupled males will have greater labor force experience than their partners. Second, it is notable that earnings in the one or two years following the birth of a child are substantially lower for females, but not for males, suggesting that females are more likely to take parental leave and suffer some financial penalty (even while still officially registered as working fulltime).

Based on parameters shown in Table 3, Fig. 2 provides a graphic portrayal of the estimated earnings profile for a representative mover from each stratum during work year eight both before and after rural-urban migration, compared to a representative member of the same stratum who remained in the rural area. In overview, the evidence strongly supports our contention of heterogeneous urban migration effects by both gender and education, though patterns are nuanced.

First consider the differential migration selection process across strata. Examination of the rural-urban movers prior to migration shows that all four strata were, on average, underpaid relative to their cohorts who remained in rural areas, but these differentials were higher for males (24 and 22 percent for lower- and higher-education) than females (11 and 21 percent for lower- and higher-education).[32] This weaker selection for urban female migrants in terms of rural underpayment (especially for less-educated females) is consistent with their more often being "trailing partners" in the primarily male-driven economic motive to move (Mulder and van Ham, 2005; Nisic, 2017). Similarly, all but lower-education females demonstrated substantially faster annual earnings growth prior to migration than their rural counterparts. Lower- and higher-education male movers exhibited three and four percent higher rates, respectively; the corresponding figures for lower- and higher-education females were zero and four percent. As expected, workers with higher educational attainments evinced faster annual earnings growth prior to migration than the less-educated members of their gender. Our parameters indicate that higher-educated male and female migrants would eventually earn as much as their rural counterparts who do not migrate after the seventh year of rural work experience; see Fig. 2. By contrast, less-educated male migrants would take much longer to catch up and less-educated female migrants may never earn as much as their rural counterparts if they did not move to the city. In sum, while urban movers generally shared common traits of more rapid growth in earnings while being underpaid on average in their rural contexts, there were clear differences by gender and education.

Second, as for static urban earnings premiums, we find that males made substantial, immediate gains from migration (32 percent for both education groups), which is a considerably larger estimate than in prior empirical work (cf. Glaeser and Maré, 2001; Wheeler, 2006; Gould, 2007; Baum-Snow and Pavan, 2012; D'Costa and Overman, 2014; Steskal, 2015; Matano and Naticchioni, 2016; Wang, 2016; Carlsen et al., 2016; De la Roca and Puga, 2017; Korpi and Clark, 2019). Females also experienced substantial static gains from urban migration: 24 and 34 percent for lower- and higher-education females, respectively. This is opposite the finding from Phimister's (2005) FE model's results that females with higher education reaped a smaller static premium than those with less. Note that our observed male-female "gender gap" in static earnings premiums is positive for those with less education (eight percentage points) but *negative* for those with more education (two percentage points). This suggests that the inconsistency of prior studies' conclusions about gender gaps in the payoffs from urban migration (cf. Combes et al., 2017; Bacolod, 2017; Wessel and Magnusson Turner, 2020; Bennett et al., 2022) might be traced to unexplored gendered heterogeneity in effects by education, as well as differences in model specification.

Third, as for dynamic earnings premiums, Table 3 shows that urban migration for all but lower-education females resulted in a significantly *lower marginal* gain from experience than would have been manifested had these individuals remained in the rural area. Net estimated returns from urban experience[33] remained positive for males (1.8 and 2.2 percent for lower- and higher-education). These findings agree with others (Möller and Haas, 2003; Wheeler, 2006; Gould, 2007;

[32] Estimated by exp (coefficient of Ti)-1.

[33] Computed at the sum of coefficients for TR, T * TR, and D * TR.

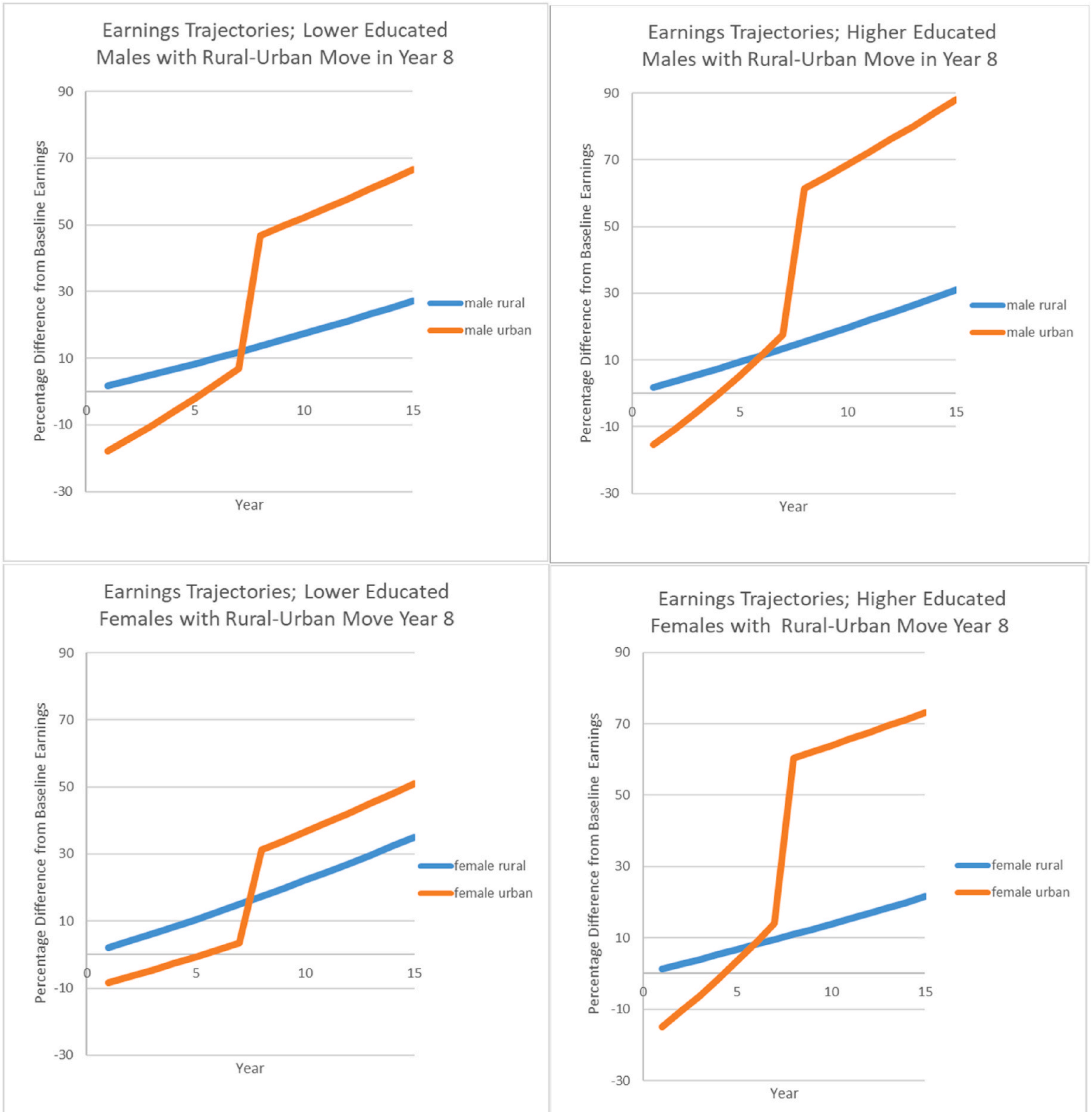


Fig. 2. Earnings Trajectories for Rural Stayers and Rural-Urban Movers Estimated by DiD-TR Model, by Gender and Education (for 15-years of employment, urban move year 8)
Note: baseline = earnings during 1st full year of fulltime work after completing education by similar individuals who remain in rural employment; all graphs based on coefficients shown in Table 3.

Baum-Snow and Pavan, 2012; Gordon, 2015; Carlsen et al., 2016) that urban earnings grow faster for better-educated males, but we attribute this primarily to an extension of their exceptional returns to experience exhibited *before* the rural-urban migration, not to urbanization effects *per se*. Similarly, net estimated returns from urban experience remained slightly positive for higher-education females (1.1 percent), but less than males' (replicating Phimister, 2005). By contrast, female migrants with less education did not experience a statistically significantly different earnings-experience profile either before or after migration, compared to their counterparts who remained in rural areas. These inconsequential returns to female urban experience are opposite to what would be predicted by the "productivity" theory of urbanization that focuses on ongoing worker skill enhancement though learning from the urban context, both on-the-job and through the broader milieu (Duranton and Puga, 2004; Yankow, 2006), and observed in much previous empirical work focusing on males (Glaeser and Maré, 2001; Wheeler, 2006; Gould, 2007; Baum-Snow and Pavan, 2012; D'Costa and Overman, 2014; Steskal, 2015; Matano and Naticchioni, 2016; Wang, 2016; Carlsen et al., 2016; De la Roca and Puga, 2017; Korpi and Clark, 2019). We posit that the reason for our distinctive finding is the failure of models previously employed (FE, DiD) to control the pre-move positive earnings trajectories that distinguished the urban movers. This apparent failure resulted in their overestimation of the dynamic effect (and underestimation of the static effect) of urbanization on earnings.[34] The overestimate of dynamic premiums was especially apparent in the FE models estimated

[34] In a robustness check, De La Roca and Puga (2017: 4.4) allow the value of experience for male movers and stayers to differ while they were employed in rural areas. They find, like us, male movers had greater returns from experience than stayers in both rural and urban settings, but it was not clear in their reporting of results whether their earnings growth was considerably smaller in the latter.

for males; see Appendix Table A2.

What might explain our key heterogeneous result that lower-education females reap the smallest urban earnings premium, higher-education females reap the highest, and males of either education group represent the intermediate case? There are several plausible reasons why lower-educated females gain the least from rural-urban migration. First, they are most likely to sublimate their own career aspirations to follow their partners' urban job opportunities (Mulder and van Ham, 2005; Nisic, 2017). This implies that such females may exhibit less migration selection on highly remunerated observed and unobserved attributes (consistent with their modest pre-move earnings growth profile) and/or likely "satisfice" in choosing urban employment opportunities. Second, lower-education females may be especially constrained by less-developed urban social networks through which employment and other business information is conveyed (Rosenthal and Strange, 2012; Bacolod, 2017). Third, they may be more reluctant to take on longer commutes because of their greater likelihood of assuming a larger share of (more traditional, gendered) household responsibilities, resulting in less efficient job matches. By contrast, higher-education females likely gain the largest premium for different reasons. First, they may benefit the most from enhanced matching with more specialized and technical jobs associated with thicker urban labor markets. This compares favorably to being "over-educated" compared to the requirements (and earnings) of the rural job (Frank, 1978; Kracke et al., 2018). Second, well-educated women may offer urban employers a preferable package of both interactive and intellectual skills, presaging that they will be the most productive workers (Duranton and Puga, 2004; Rosenthal and Strange, 2004; Bacolod, 2017).

That males in our sample reaped similar static premiums regardless of education is surprising at one level, but not if one considers the rural-urban industrial context and selectively of migration. By the time of migration considered by our study (1997–2016), many of the huge changes in rural employment (loss of farming, fishing, forestry, mining jobs; increases in public services and volatility in manufacturing) had sorted themselves out. Subsequently, there were virtually no differences in rural and urban shares of male employment in any major industrial sectors except manufacturing—where the rural share grew progressively higher over time (Bennett et al., 2022). For less-educated rural males in the manufacturing sector this meant ever-eroding prospective relative gains from urban migration. We think, therefore, that it is a highly selective subset of low-education rural males who chose to move to an urban area. We cannot discover all their unobserved traits, but we do observe that they were much more likely to switch industrial sectors while in rural employment than their counterparts who did not migrate, and were as likely to change sectors again at the point of moving as high-education males (more on this below; see Appendix Table A12). These and other unobserved characteristics of low education male migrants during our analysis period apparently were sufficient to reap them a premium comparable to that of better-educated males.

4.2.2. Different specifications yield different portraits of heterogeneity

We return to a theme raised earlier in 4.1 about how traditional DiD and FE econometric specifications yield different estimates of the urban earnings premium than DiD-TR. Here we reinforce that conclusion by comparing these models disaggregated by gender and education. Table 3 shows the stratified earnings model parameters estimated with DiD-TR, and Appendix Tables A1 and A2 present corresponding parameters estimated with DiD and FE specifications. A comparison of these three tables again makes it clear that alternative specifications yield significantly different portraits of gender and education differentials in both static and dynamic urban earnings premiums. First, although both DiD and FE models substantially understate the static premium across all four strata compared to the DiD-TR estimate, it is especially the case for males—ranging from 13 to 18 percentage points less—compared to 4 to 12 percentage points for females. This yields different implications for gender gaps: the static premium for males with lower education is 8 percentage points higher than comparable females as estimated by DiD-TR, but only 3 to 6 percentage points as measured by DiD and FE. On the contrary, the static premium for females with higher education is only 2 percentage points higher than comparable males as estimated by DiD-TR, but is 4 percentage points as measured by DiD or FE. Perhaps most clearly, DiD-TR finds that lower-education females reap the smallest static premium by far, whereas both DiD and FE find that they get a higher premium than either male group. Second, the three specifications provide distinct pictures of males' dynamic urban earnings premium. Regardless of education, the DiD-TR model estimates a negative dynamic premium, the DiD an insignificant one, and the FE a positive one for males. In sum, we have argued above the DID-TR specification has greater internal validity than the FE and DiD specifications conventionally used. Here we have demonstrated that conclusions about the heterogeneity of the urban earnings premium by gender and education are dramatically shaped by which econometric specification employed.

4.3. Supplemental analyses and robustness tests

4.3.1. Flexible functional form for experience

Our first robustness test involves relaxing the assumption of linearity in the value of experience in (4). One might reasonably argue that if the value of experience is convex (i.e., increasing at a decreasing rate) for movers and stayers alike, that: (1) the apparent slower growth of earnings for stayers compared to movers early in their work lives is (partially) an artifact from fitting a linear function for stayers; and (2) the break in trajectory we observe for migrants may not be (solely) related to the urban move but rather to this nonlinearity.[35] To test for this, we re-specified (4) as a "binned" model, replacing the linear trend for experience (TR_{it}) with a series of dummy variables denoting two-year increments of experience.[36] Results are presented in Appendix Table A3, and portrayed graphically in Appendix Figure A2.

Several salient points present themselves in this Table and Figure, all of which support the main findings of the linear DiD-TR model. First, with the possible exception of higher-education females, there is no obvious, substantial convexity in the value of experience for non-movers. Second, urban migrants (with the exception of lower-education females, as before) exhibit substantially greater gains from early workforce experience in rural areas than their counterparts who do not migrate, but this trajectory attenuates significantly after moving to the city (especially for those with higher education, as before). Third, all eventual urban migrants are, on average, paid less while working in rural areas than their counterparts; the degree of underpayment is the same as that measured in the linear DiD-TR model. Finally, the measured static urban earnings premium here is somewhat smaller for all groups, with higher-educated males exhibiting the highest value. This result is not directly comparable to that of the linear DiD-TR model, however, since in the binned model we also included dummy variables denoting both the year of the urban move and the year before (to test for an "Ashenfelter dip").[37] The former's coefficient suggested for migrants with higher education that the rural-urban transition year resulted in a one-time decrement in earnings (seven percent for males, three percent for females), likely due to the associated break between fulltime jobs. The latter's coefficient indicated that all strata exhibited a six to nine

[35] This potential challenge to validity was raised by a reviewer, whom we thank for this insight.

[36] This excluded the dummy denoting having years 2 and 3 of experience occurring in cities because we imposed a sample restriction that only those who (after completing schooling) worked two years or more in rural areas before moving to the city were analyzed.

[37] This potential phenomenon of a lower-than-typical earnings in the year prior to moving would bias our estimated static premium upward. For background on this phenomenon see De La Roca and Puga (2017: 4.4).

Table 4
Estimated parameters of Cox hazard model of rural-urban migration.

VARIABLES	(1) All	(2) Male Low	(3) Male High	(4) Fem Low	(5) Fem High
Married/Cohabiting	0.097***	0.153***	0.124***	0.084***	0.071***
Prior Year	(0.008)	(0.025)	(0.018)	(0.022)	(0.011)
Separated/Widowed	0.253***	0.626	0.363	0.317	0.171**
Prior Year	(0.086)	(0.368)	(0.258)	(0.317)	(0.100)
Had a Child 1-2	0.654***	0.648*	0.931	0.613	0.371***
Years Prior	(0.058)	(0.112)	(0.129)	(0.194)	(0.075)
ln CAGR Wages	**1.017*** **	**1.018***	**1.008**	**1.039***	**1.052*** **
Thru Prior Year	**(0.005)**	**(0.008)**	**(0.009)**	**(0.018)**	**(0.013)**
ln Wages, 1st Full Year	**0.861*** **	**0.922**	**0.777*** **	**1.147**	**0.977**
Full-time Work	**(0.021)**	**(0.054)**	**(0.029)**	**(0.133)**	**(0.049)**
Health Sector	0.919	1.664*	1.269*	0.686	0.810**
1st Year Work	(0.054)	(0.419)	(0.142)	(0.146)	(0.059)
Primary Sector	0.399***	0.465**	0.360***	0.306	0.457*
1st Year Work	(0.057)	(0.113)	(0.077)	(0.220)	(0.148)
Education Sector	0.825	0.872	1.049	0.573	0.658**
1st Year Work	(0.082)	(0.508)	(0.147)	(0.244)	(0.099)
Public Admin. Sector	0.994	1.740***	0.902	0.763	0.773
1st Year Work	(0.087)	(0.274)	(0.121)	(0.264)	(0.142)
Female w/Lower	2.158***				
Education	(0.164)				
Male w/Higher	5.290***				
Education	(0.331)				
Female w/Higher	5.788***				
Education	(0.373)				
Observations	176,798	97,517	23,773	25,652	29,856
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Notes: Robust standard errors in parentheses; CAGR = cumulative annual growth rate.
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

percent *higher* earnings the year before a move.

4.3.2. Omitting potentially endogenous control variables

In prior DiD-TR models we controlled for partnering status and childbearing. It well might be the case, however, that these two time-varying outcomes are themselves influenced by one's earnings and/or the decision to move from a rural to an urban area.[38] To ascertain if our core results were influenced by this potential "over-controlling," we re-estimated parameters of (4) omitting partnering status and childbearing variables; results are presented in Appendix Table A4. In sum, they alter none of our prior conclusions in substantial ways. When these covariates are omitted, the estimated: (1) selection based on being underpaid on average but with greater payoffs from experience in rural areas persists; (2) static earnings premiums are about one percentage point higher for all four strata; (3) the dynamic premiums are slightly less negative for males but slightly more so for females. The latter result is unsurprising, given the earnings penalty that females bear by marriage/cohabitation and/or bearing children (see Table 3).

4.3.3. Altering the reference group to rural-rural movers

In all previously discussed models, we employed as our baseline reference group all those who met our sample inclusion criteria and lived (and worked) solely within rural areas during the span of our analysis. This group included those who both always lived in the same location and those who changed (rural) residence. Here we test the extent to which the exclusion of the former subset from the analysis affects our DiD-TR model, thereby comparing earnings functions for rural-rural movers and rural-urban movers. Results are presented in Appendix Table A5. Comparison with Table 3 shows that this alteration of the rural reference group reduces the absolute magnitudes of all parameters related to selection and the urban earnings premium, implying that all movers tend to gain from their mobility. None of or prior qualitative conclusions are altered, however. Urban migrants are still underpaid while working in rural jobs, on average, but their growth in earnings is greater compared to their counterparts who move among rural jobs. The static urban earnings premium estimate drops by about ten percent for those with higher education, and by about 30 percent for those with lower education. The ranking of these premiums are: higher-education females, males, lower-education males, females, as in the original sample. The dynamic premiums are all slightly less negative here. In concert, these results suggest that while the starkest differences in earnings trajectories exist between those who never move rural residence and urban migrants, there remain substantially different selection factors and payoffs from mobility between those who move among rural locations and those who move from rural to urban locations.

4.3.4. Distinguishing moves to Oslo

Our next robustness check assesses the sensitivity of results to the definition of "urban." Prior work has suggested that industries sort across the urban size hierarchy, with those requiring higher skill and non-routine tasks locating disproportionately at the top. This implies different earnings premiums may be generated at different points in the hierarchy, with the greatest being associated with the highest tier (Combes et al., 2008; Korpi and Clark 2019), especially for those with more education (Carlsen et al., 2016). Wessel and Magnusson Turner (2020) found that migration from rural Norwegian areas to modest-scale, tier 2 cities benefitted males' earnings more than females', whereas the opposite was true for moves to tier 1 Oslo. Moreover, selection patterns may also be different at different points in the urban hierarchy (De la Roca, 2017).

Here we re-estimated our gender/education-stratified DiD-TR model employing only those moving from rural areas to the Oslo Travel to Work Area (N = 2917), dropping those from the sample those moving to the second-tier metropolitan areas of Bergen, Stavanger and Trondheim

[38] Equally plausible, of course, is that partner status and childbearing might influence the decision to migrate, in which case it is important to include them in the model as controls for selection.

(N = 3015). Oslo TTWA is the acknowledged top of the urban hierarchy in Norway (Høydahl, 2017; Wessel and Magnusson Turner, 2020), in which roughly a fifth of the nation resides.

Results are presented in Appendix Table A6; they strengthen virtually all of our prior conclusions. They replicate our prior findings regarding selection patterns that movers are underpaid, on average, compared to their rural counterparts yet demonstrate stronger growth in earnings before they move, though magnitudes of both these features are *larger* here for all strata; i.e., selection on these characteristics appears stronger in the case of movers to Oslo than to lower-tier urban areas. As expected (and consistent with Norwegian studies by Carlsen et al., 2016; Wessel and Magnusson Turner, 2020), the static earnings premium is considerably larger for each stratum when only movers to Oslo are considered.[39] Uniquely in our study, however, the same (or larger, in the case of those with less education) diminution in growth of earnings pre-/post migration is observed in the Oslo-only sample for all strata. Moreover, we find nuanced variation *jointly* by gender and education that belies the aforementioned generalizations by gender (Wessel and Magnusson Turner (2020) and education (Carlsen et al., 2016) about who gains more from moving to a larger instead of smaller city. Compared to moving to tier 2 urban areas, we find that the increment in static premium from moving to Oslo is largest for lower-education males, followed by lower-education females, higher-education males, and higher-education females.

4.3.5. Alternative adjustment for real earnings

Our next robustness check involves an alternative way of specifying real (inflation-adjusted) earnings. In our core model we followed convention by deflating nominal earnings by the Norwegian national Consumer Price Index and employing this as the (logged) dependent variable. Dumond, Hirsch and MacPherson (1999) have argued that a superior specification employs nominal earnings and uses the CPI as a control variable. When we do this we find virtually no alterations in the coefficients presented in Table 3; cf. Appendix Table A7.

4.3.6. Controlling for industrial sector

In our final sensitivity test, we control for the industrial sector in which the individual is working each year. Controlling for industrial sector is a way of adjusting observed rural-urban earnings disparities for inter-industry pay differentials that (at least partially) are independent of geography and thus theoretically unrelated to the urban premium. However, by doing so we may be over-controlling inasmuch as it rules out the potential influence of enhanced urban productivity through better worker-employer matches that might be gained through changing industrial sector concomitant with urban migration. Specifically, for this test we use the Norwegian NACE (Nomenclature of Economic Activities) standard, a Norwegian adaptation of the conventional European Union NACE classification system. We apply two digits for our classification, resulting in 54 mutually exclusive and exhaustive groups.[40]

The results of adding these controls to our core DiD-TR model are presented in Appendix Table A8. Overall, they show remarkably little differences between the core model's estimates and those produced when industrial sector is controlled, suggesting that alterations in sector associated with rural-urban migration are not the prime driver of the observed urban premium for this full-time worker sample.[41] Specifically, controlling for sector does nothing to alter the conclusion that those (especially males) who are underpaid, on average, compared to their compatriots in rural jobs, but also exhibit greater-than-average gains from rural experience are most likely to migrate. The measured static earnings premium is slightly lower for lower-education males and females, but this premium is slightly higher for higher-education males and females when sector is controlled (both differences are two percentage points or less). The estimated dynamic premium is virtually identical across the two models.

4.3.7. A replication test

A central claim of this paper is that failure to control for the unusually strong returns from experience that urban migrants exhibit prior to moving leads conventional FE and DiD models to understate the static urban earnings premium and overstate the dynamic premium. Here we demonstrate this directly by replicating (to the extent feasible[42]) Carlsen et al.'s. (2016) FE model of the Norwegian urban earnings premium and then examining how the results change when our DiD-TR model is applied. Results produced by these alternative specifications, stratified by gender and education, are presented in Appendix Tables A9 and A10.

Table A9 replicates the core conclusions of Carsen et al.'s (2016: Table 4) FE model: higher static and dynamic premiums for those moving to Oslo (compared to tier-2 cities) and for those with a higher education. Comparison of Tables A9 and A10 reveals that, indeed, the FE estimate of the static premium is smaller and the dynamic premium is larger than those estimated with our DiD-TR specification employing the same observations and covariates. The FE model's static premium parameters for the full sample of migrants to Oslo and tier-2 urban areas are 55% and 48% smaller, respectively, that those estimated with DiD-TR. The FE model's dynamic premium parameters are slightly positive for urban migrants, whereas in the DiD-TR model they are significantly negative, mostly offsetting their pre-move annual value of experience. As for the question of heterogeneous impacts from urbanization, the two models give distinctly different answers. The FE model indicates that the static premiums from moving either to Oslo or tier-2 cities are highest for higher-education females, followed by lower-education females, higher-education males, and lower-education males. By contrast, the ordinal rankings for the DiD-TR model are: (1) higher-education female, higher-education male, lower-education female, and lower-education male movers to Oslo; and (2) higher-education male, higher-education female, lower-education female, and lower-education male movers to tier-2 cities. Thus, for addressing all the research questions of this paper, it is clear from this replication exercise that the econometric specification matters greatly.

5. Supplemental analysis of rural-urban migration selection

The conclusion from our core DiD-TR model that younger, full-time workers who migrate from rural to urban Norway exhibit lower but faster-growing rural earnings, on average, remains robust to a wide variety of tests. This notable, new observation is thus worthy of deeper exploration. In this section, we briefly review what has been found by

[39] This differs from Baum-Snow and Pavan's (2012) finding that static effects were more important for generating wage premiums between medium and small locations, and dynamic effects were more important for large and small area comparisons.

[40] More information about the standard and the various versions is found at: https://www.ssb.no/en/klass/klassifikasjoner/6. The NACE codes are available from 1995, but because the classifications were modified over time we were forced to recode them using algorithms made available by Statistics Norway; details ay: https://www.ssb.no/virksomheter-foretak-og-regnskap/naeringsstandard-og-naeringskoder [read 11.09.2023].

[41] This is consistent with the conclusions of Baum-Snow and Pavan (2012) that better job matches available in the larger cities are an insignificant contributor to the earnings premium. It is not consistent, however, with the findings of Korpi and Clark (2019) that controlling for industry and occupation substantially reduces the measured premium.

[42] Like Carlsen et al. (2016), we omit workers from the public and primary (farming, fishing, forestry) industrial sectors, allow static and dynamic premiums to vary between Oslo and tier 2 urban areas (though our geographic definitions differ slightly), and control for industrial sector, year and age. We cannot control for occupation or job tenure as they do because of unavailable information for many working years of our cohorts.

previous research about who is most likely to migrate to the city and then estimate a Cox hazard model of rural-urban migration as means of gaining insight into the predominant drivers of this process for our analysis sample. Finally, we conduct an exploratory set of regressions in an effort to identify the characteristics of our migrants associated with their distinctive pre-migration earnings profiles.

5.1. Cox hazard model of rural-urban migration

Previous international work on the nature of selection into rural-urban migration has been quite consistent in which observable characteristics are salient. Multivariate probability analyses have demonstrated that a move to the city is more likely if the individual is: female, more educated, higher-scoring on achievement tests, younger, single, without school-age children, receiving unemployment benefits, or working in white-collar (instead of blue-collar) occupations (Rye, 2006; Gould, 2007; Korpi and Clark, 2019; Niedomysl and Fransson, 2014).[43] We draw upon these findings in specifying predictors in our own model of the decision to move from rural to urban areas of Norway.

Specifically, to investigate this geographic selection we employ the well-known Cox proportional hazard model, with the hazard function denoted by h(t).[44] The hazard function can be interpreted as the risk of moving from a rural to an urban area during year t, and can be expressed:

$$h(t) = h_0(t) \times exp\left(\beta_1 x_1 x_2 + \ldots + \beta_p x_p\right) \tag{5}$$

where.

- t represents the survival time (years) in rural residence
- $(x_1, x_2, \ldots, x_p)$ is a set of p time-varying and time-invariant covariates
- coefficients $(\beta_1, \beta_2, \ldots, \beta_p)$ measure the impact (i.e., the effect size) of covariates
- h_0 is the baseline hazard (i.e., hazard if all x_i equal zero)

Besides employing as many of the foregoing demographic, educational and economic characteristics as possible, we are especially keen on examining the impacts of variables that our DiD-TR analyses indicated distinguished movers from stayers: level of earnings and value of experience while residing in rural areas. We measure the former by the earnings exhibited in the first full-year of full-time work after education is completed (i.e., when $T_{it} = 2$). We measure the latter by the (natural logarithm of) compound annual rate of earnings growth occurring between the workers' first full year of full-time experience after they finish their education and the year prior to when the move may occur (t-1).[45]

Parameters estimated for the hazard model are presented in Table 4 model (1); results fully conform to expectations and prior research. The gender/education relationships with the hazard of moving to an urban area in the full sample are as follows: lower-educated males are the least likely to move, followed by lower-educated females, higher-educated males, and higher-educated females, all else equal. Having a child one or two years prior is associated with a lower hazard of moving in a given year, as is being married/cohabiting or being divorced/widowed in the prior year, compared to never being partnered. Also unsurprisingly, those who start their careers in the primary sector (mining/extraction, forestry, and fishing) are much less likely to move from rural areas than those in public or other private sectors.Of more salience to this research, those who start their rural careers with higher earnings (controlling for gender, education, and industrial sector) have a lower baseline hazard of migrating, presumably because they feel relatively well-compensated where they reside. By contrast, those who experience a faster growth of their earnings during their rural work lives are more likely to leave for an urban area. Below we probe further the potential reasons for these intriguing patterns. Suffice it to note here that the Cox model provides further confirmation of the results from our DiD-TR model regarding Norwegian rural-urban geographic selection on the basis of lower average earnings and faster earnings growth in the rural employment context.

It is also of interest to note the substantial heterogeneity across our four gender/education strata in the predictors of rural-urban migration; cf. Table 4 models (2)–(5). Having a child recently or a partner appear to be much larger deterrents to migration for females than males, regardless of education. Starting in the health sector significantly increases the hazard of moving for males, but just the opposite for females, regardless of education. Higher-educated females are much less likely to migrate if they start in the education sector, whereas lower-educated males are more likely to do so if they start in public administration. The positive association between growth of rural earnings and hazard of moving is quite general across the strata, but the negative association between starting level of earnings and this hazard is most dominant for higher-educated males.

5.2. Correlates of earnings trajectories of rural-urban movers: the key role of inter-sector job changes

Our Cox hazard model confirmed that those in our analysis sample who migrate have distinctive earnings profiles. We probe deeper into the multivariate correlates of these profiles by running OLS regressions on the sample of movers using our two earnings trajectory variables—initial level of full-year, full-time earnings and cumulative growth rate of earnings (measured here through the year before the observed move)—as dependent variables. Besides the aforementioned variables used in the hazard model, we introduce as a predictor in the earnings growth model the number of changes between industrial sectors that the individual undertook before moving to an urban area.

The model of initial earnings did not yield any valuable insights,[46] but that for growth of earnings revealed what we believe is a key relationship: those who changed industrial sectors more often exhibited a much faster growth of their earnings before they moved. Specifically, we estimated that each additional switch between sectors was associated with a 1.6 percentage-point-higher cumulative rate of growth in earnings (see Appendix Table A11). This parameter was highly statistically significant in all four gender/education strata and deviated in magnitude among them by only 0.2. Although these findings are not necessarily causal, they are consistent with the hypothesis that changing industrial sectors results in superior matches between employers' job requirements and workers' skills, which redound to the productivity benefit of both and thus yield wage gains.

If indeed their more frequent changes in industrial sector are a key reason for the exceptionally high growth of earnings exhibited by rural-urban migrants prior to moving, their observed slowing in earnings growth after moving (see Table 3 and Fig. 2) should also be associated with a decrease in their frequency of inter-sectoral changes of employment. This proves to be true; see Table A12 showing the annual average

[43] The evidence how unobserved ability/skills (as measured by worker fixed effects) relates to urban migration is less consistent. Combes et al. (2012) find that both extremely high and extremely low-skill workers are more prone to migrate to bigger cities. De La Roca and Puga (2017) find that this is an artifact of the failure to control for city-specific gains to experience, and thus agree with Baum-Snow and Pavan (2012) that there is no significant sorting on unobserved ability/skills. Carlsen et al. (2016) find that De La Roca and Puga are correct for those with less education, but those with the most unobserved ability/skills among the college-educated are more likely to select urban residence.

[44] Our preliminary chi-squared test revealed that the assumption of proportionality could not be rejected, thus we employ the Cox instead of the accelerated frailty time model.

[45] where is the initial earnings value, is the current earnings value, and is the number of years elapsed between them.

[46] Results are available from the authors upon request.

number of inter-sectoral job changes for rural stayers, urban migrants before they move, and urban movers after they move, all stratified by gender and education. On average, those who undertake rural-urban moves change industrial sectors 2.8 times for every ten years they work in a rural area, but only 1.1 times for every ten years they work in an urban area. This relationship holds across all gender/education strata: switching between sectors occurs much more often (roughly by a factor of two to three) before the move than after. Note also that future migrants exhibit much higher rates of rural inter-sectoral job changes than those who remain employed in rural areas (2.8 vs. 1.5 changes per ten years of employment, respectively); this relationship appertains to all gender/education strata.

5.3. Conclusions and implications regarding selection

In sum, we draw the following conclusions from the foregoing analyses of the nature of selection into rural-urban migration by our Norwegian sample. Our hazard model confirmed more directly our prior inferences drawn from our DiD-TR model: those who migrate from rural areas exhibit lower mean earnings but faster growth in earnings before they move, net of other observable characteristics. Movers also changed industrial sectors while working in rural areas more often than their counterparts who did not move; movers who changed sectors more often experienced greater cumulative growth of earnings. Having moved, however, their frequency of inter-sector changes declines. These patterns appertain qualitatively to all gender/education strata, though those with higher education change sectors more often before they move and less often after, compared to movers with lower education.

What might be going on here? We think several scenarios are consistent with our data, though admittedly what follows is speculative. After completing their education, rural residents who we know will migrate to the city eventually find themselves for one reason or another mismatched in their first jobs. Moreover, they are relatively low-paying. This relationship may be due to bad luck or a poor employment search process. Or, it may be due to negative selection by employers on the basis of unobserved individual characteristics. Regardless, some employees leave the rural area relatively quickly if they perceive bleak prospects for a better match. Others stay but switch industries in an effort to find a better match; those with higher education find this harder, so change more often. As they change sectors they see their rural earnings increase substantially. At some point, a more attractive urban opportunity may present itself, even if they may have found a suitable sector in the rural area. Typically, the move to an urban area does not involve a change of industrial sector (only 16 percent of our sample migrants did so at the point of moving; see Appendix Table A12), but a substantial wage premium nevertheless. This static premium could be due to agglomeration economies or superior urban job-worker matching within the same industrial sector. Regardless, the quality and compensation of the urban job match is sufficient from the employee's perspective that fewer inter-sectoral changes are required subsequently, with the concomitant slowing of earnings growth.

6. Study limitations

The primary limitation of our study is generality. Our core estimates of young adult earnings trajectories apply only to those born from 1979 to 1981 in rural Norway who, after completing their education and starting consistently full-time work, moved from a rural[47] to an urban area (and did not return) and completed at least two years of full-time employment in both contexts. Such sample restrictions are necessary to operationalize the DiD-TR model that we argue has superior internal validity than conventional FE and DiD approaches. Though our conclusions are robust, we acknowledge that they may not necessarily appertain when rural-urban moves occur at a different point in the life course or when they involve changes in hours worked. For illustration, Wessel and Magnusson Turner (2020) found that Norwegian female rural-urban migrants (especially to Oslo) gained the most when they moved *before* completion of education, and such gains were primarily due to more hours of working and changing industrial sector, not higher wages for the same jobs.

We did not consider here those who undertook urban-rural moves. We believe that it is beyond the scope of this paper to delve into this sample expansion. Given prior findings that learning gained from prior urban residence can be indelible and persists when one moves back from the city to a rural area (Andersson et al., 2014; De la Roca and Puga, 2017), we believe that the expansion of the sample to include return migrants would only dilute and obfuscate our main findings. Nevertheless, the economic returns from such moves are worthy of future study.

We have measured urban earnings premiums for migrations taking place over the entire period between 1997 and 2016 in Norway. It may be the case (e.g., Bennett et al., 2022) that there are temporal variations in these premiums that our study overlooked.

There are other study shortcomings related to data limitations. It is unfortunate that Norwegian social registers for the overall period under investigation do not provide accurate information on hours worked that would allow us to probe this issue more fully. For our complete analysis period they also do not provide information about occupations or specific employers, nor how long individuals have been employed by them. Finally, they do not record a complete lifetime of data for cohorts born earlier than the ones we analyzed, thus we could not consider workers who moved to a city in their later working years. De la Roca (2017) found for Spanish migrants that those who moved later in their careers may have had more certain employment prospects and were less likely to return to a rural context, implying that gender and educational gaps in premiums may be different if measured at different ages of migration.

7. Conclusion

We have explored urban earnings premiums gained by native individuals born in 1979, 1980 and 1981 in rural Norway who, after completing their education and starting fulltime work, moved from a rural to an urban area. Our difference-in-differences with pre-move differences in trends controlled (DiD-TR) model demonstrated that estimates produced by conventional fixed effects and difference-in-differences specifications of urban wage premiums substantially understate the static premium (and often overstate the dynamic premium) because they fail to account for the atypically low starting level but faster growth of earnings that urban migrants (regardless of gender and education) evince before they move. Our hazard model of rural-urban migration confirms the findings from the DiD-TR model that those who are paid less when they start their careers but subsequently evince a more rapid increase in rural earnings are more likely to migrate. This increase appears to be driven by more frequent changes of employment between rural industrial sectors, which diminish considerably after moving to an urban area. If these findings can be generalized, they challenge on methodological grounds longstanding claims about the size and sources of the urban earnings premium.

Parameters estimated from DiD-TR models jointly stratified by gender and education reveal more heterogeneity in urban earnings premiums than has typically been revealed by prior work that has stratified by either one characteristic or the other. Females with above-median educational attainment (college degree or more) gain the largest static premium and less-educated females the smallest, with males

[47] As a reminder, the largest town among our "rural" observations (centrality levels 4, 5, 6) has a population of only 7,884, so the geographic origins for migration to the central areas classified 1 or 2 is indeed a non-urbanized baseline. Our sample does not involve those who might undertake this move starting from a somewhat higher level 3 of centrality (since it is ambiguously "rural"), nor does it consider movers from levels 2 to 1.

gaining an intermediate amount regardless of education. Once again, estimates produced by conventional fixed effects and difference-in-differences specifications come to different and, we believe, incorrect conclusions about these gender differentials in urban earnings premiums.

If our results may be generalized, they would suggest that cities primarily generate wage premiums through agglomeration efficiencies and by better job-employee matches achieved at the point of the rural-urban migration, rather than by increases in worker productivity associated with urban work experience. These mainly static premiums do not accrue uniformly, however, but differ substantially by gender and education jointly. Regardless of the underlying processes at work, it is clear that urbanization in Norway has led to widening rural-urban gaps in earnings among full-time workers, especially for well-educated women.

Declarations

The authors declare no conflicts of interest.

Funding

This research was supported by the Norwegian Research Council grant number 302528.

CRediT authorship contribution statement

George C. Galster: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Liv Osland:** Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

This research was supported by the Norwegian Research Council grant number 302528.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors thank Stuart Rosenthal, Anna Maria Santiago, Lena Magnusson Turner and Terje Wessel for their helpful comments on an earlier draft.

Appendix A. Graphic Illustration of Estimating Earnings Premiums with FE, DiD and DID-TR Approaches

A simplified graphic illustration not only illuminates the potential shortcomings of FE and DiD in measuring urban earnings premiums in a more intuitive way, but also provides a basis for demonstrating the superiority of the DID-TR estimator that we employ in this study. Consider a stylized nation with one urban and one rural location and two working adults (or, equivalently, two equally sized groups of workers with different but internally homogenous characteristics) currently living in the rural location. Let us collect wage information about them annually for n years. At this point, one worker moves to the urban location and we again collect wage information about both workers annually for n additional years. For simplicity and with no loss of generality, let the non-migrating rural worker's annual earnings be fixed at A throughout. Let the migrating worker have four alternative earnings profiles, detailed below, which by construction all produce the same mean annual earnings differential with the non-migrating rural worker, as measured over the 2n period.[48] We next consider the validity of the conventional FE model for estimating the coefficient of a dummy variable indicating earnings in the urban location. We compare four hypothetical scenarios involving different assumptions about the unobserved characteristics of the migrating worker and the effect of the rural-urban migration on earnings. We conduct a parallel analysis for the well-known difference-in-difference (DiD) estimator since it is more closely related than FE to our DID-TR method. In all scenarios, we assume that all observable characteristics of both workers are identical and consider the consequences of controlling for experience.

[Figure A1 about here: Four Hypothetical Earnings Profile Scenarios of Rural-Urban Migration].

Scenario 1. Migrant has a distinct but **time-invariant bundle of unobserved** characteristics that affect both the probability of migrating and earnings; no gain from job experience or urban migration → flat migrant earnings profile CG in Figure A1. After comparing the annual average earnings of the migrant to the non-migrant, the FE would be calculated as (C-A). Since there would be no residual for the post-move period conditioning on this FE, the model would correctly imply that no urban wage premium existed. In the case of the DiD estimator, the inter-worker difference in earnings would be (C-A) both before and after migration. Since the difference between the differences would be zero, DiD would also correctly imply that no urban wage premium existed.[49] These conclusions are not sensitive to whether experience is controlled in the model since there are no gains from experience observed.

Scenario 2: Migrant has distinct but **time-invariant** bundle of **unobserved characteristics** that have a constant effect on earnings pre- and post-move; immediate and fixed gain from urban migration; no gains from experience → step migrant earnings profile BIMN in Figure A1. Here the FE model would accurately measure a static urban wage premium of (D-B) after controlling for the FE (B-A), and the DiD model would yield the same: (D-A)-(B-A). As in Scenario 1, both conclusions are insensitive to the inclusion of experience in the model.

Scenario 3: Migrant has a distinct but **time-varying bundle of unobserved characteristics** that affect the probability of migrating and generate constant marginal gains in earnings; no gain from urban migration → positively sloped linear migrant earnings profile AK in Figure A1. If worker experience were not controlled, the FE would still be (B-A) by construction. However, there would still be a positive residual for earnings in the post-move period after conditioning on this FE, yielding an incorrect premium estimate of (D-B). Similarly in the case of the DiD estimator, the mean inter-worker difference in earnings would be (B-A) before and (D-A) after migration. Since the difference between the differences is positive (D-B, which by construction = C-A), the DiD model also would incorrectly imply that an urban wage premium existed. If work experience were controlled the upward

[48] In Figure A1, the four scenarios each yield the same total earnings superiority for the migrant compared to the non-mover because the areas of the polygons formed by the alternative migrant excess earnings profiles are identical. I.e., in (1) rectangle area = J*G but by construction G = 0.5*K; (2) triangle area = 0.5*J*K; (3) two rectangle areas = (H*I) + (J-H)*N but by construction H==H = 0.5*J, I = 0.25*K, N = 0.75*K; (4) triangle area = 0.5*H*L but by construction H = 0.5*J and L = 2*K. Thus, in each case the annual earnings of the migrant that are in excess of those for the non-mover = (0.5*J*K)/2n.

[49] This reinforces that point made by Wooldridge (2002) that FE and DiD are formally equivalent if certain assumptions are met.

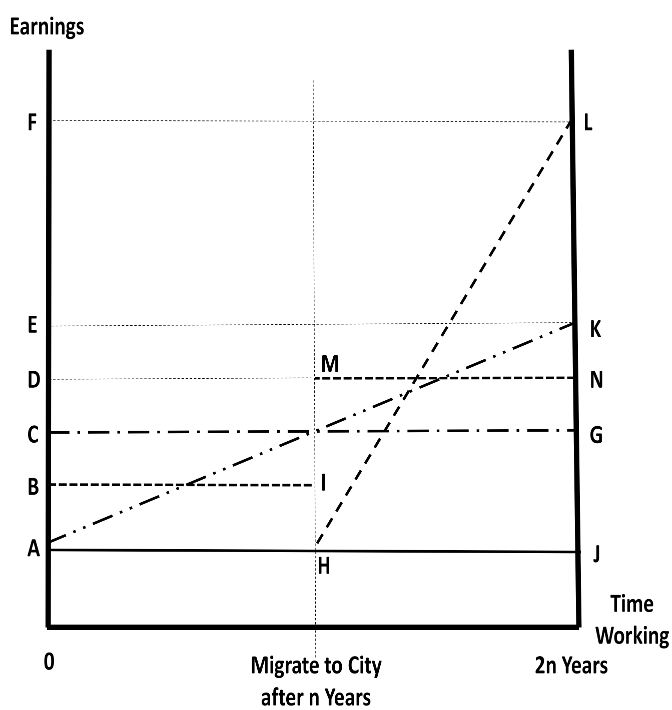
bias in the measured urban earnings premium would be smaller but would persist. The marginal gain from experience on average across both workers would be measured as half the slope of AK, by construction,[50] thereby leaving positive residuals in the post-migration period that would be (erroneously) picked up by the dummy variable indicating urban earnings.[51]

Scenario 4: Migrant has **no distinct unobserved characteristics** that affect either the probability of migrating or earnings post-move; urban wage premium takes form of constant marginal gain from urban job experience only → kinked migrant earnings profile AHL in Figure A1. In this scenario, the true total earnings gain from urban-rural migration is the area of triangle HJL (the true coefficient of a dummy variable indicating urban earnings would be this area divided by n years). Unfortunately, FE would underestimate this gain because the urban earnings premium would contribute to the estimation of the FE. Controlling for experience would not eliminate this bias because it also would erroneously attribute some of the gain from urban-only experience to generic experience.[52] The DiD estimator in this case (E-A) would provide a superior measure of the true urban earnings premium, though it could not distinguish that migration affected only the growth of earnings. Controlling for experience would only serve to bias the estimated premium downward, for the same reasons as in Scenario 3.

These scenarios demonstrate two lessons about the conventionally employed FE model with control for experience. First, when earnings trajectories for rural and urban workers are parallel, the FE estimator yields unbiased estimates of true urban earnings premiums, regardless of whether experience is controlled. In Scenario 1, it correctly indicated that higher earnings for the migrant were due purely to selection, and in Scenario 2 it accurately captured the full premium. Second, the internal validity of the FE estimator erodes when rural and urban workers' earnings grow at different rates over time, as can occur when unobserved worker characteristics influence intertemporal earnings trajectories or ability is valued more highly at the margin by urban employers. In Scenario 3, the FE estimator generated a "false positive" indicator about the existence of an urban earnings premium and in Scenario 4 it gave a "false negative." In neither case did controlling for experience eliminate these biases. Analogous strengths and shortcomings apply to the DiD model.

Heuristically, DID-TR represents an amalgamation of interrupted time-series (ITS) and DiD approaches inasmuch as it measures both the level and the trend in the outcome indicator in the treatment group (rural-urban migrants) before and after the treatment (moving to the city), then adjusts these measures based on comparable ones made for the control group (rural non-movers). An impact is indicated if either there is a post-treatment difference in the pre-treatment differences in control/treatment group levels of the outcome (as in DiD), or in trends of the outcome (as in ITS).

Even at this level of generality, it should be clear how DID-TR sidesteps the FE pitfalls harming internal validity illustrated in Figure A1. In Scenario 3, DID-TR would not discern any pre-/post-migration break in the trend of migrant earnings, so would correctly attribute no impact from migration to the city. Just the opposite discernment and conclusion would be correctly made by DID-TR in Scenario 4.

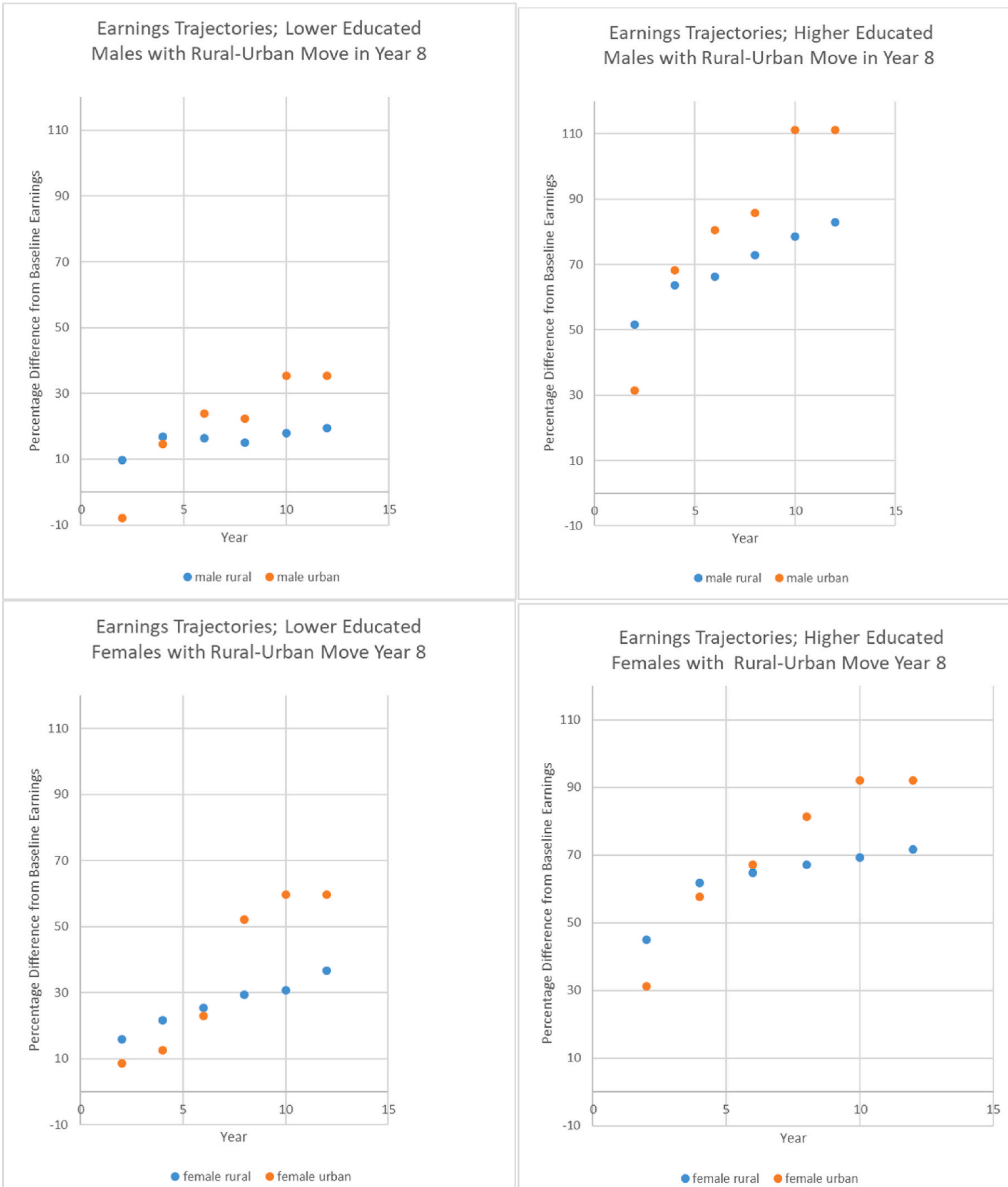


Appendix Fig. A1. Four Hypothetical Earnings Profile Scenarios of Rural-Urban Migration
Note: Graphs based on parameters in Table A3; baseline = earnings during 1st (partial) year of fulltime work after completing education by similar individuals who remain in rural employment; year before move assumed to correspond with year 6–7 and year of move assumed to correspond with year 8–9

[50] I.e., the marginal gain from a generic year of experience for the non-mover is zero and for the mover is the slope of AK, resulting in a regression coefficient on experience equal to one-half the slope of AK.

[51] Yankow (2006) avoids this problem by specifying individual FEs for experience.

[52] I.e., the marginal gain from a generic year of experience for the non-mover is zero and for the mover is the slope of AK (= the area HJL divided by n years, by construction). As in scenario 3, this results in a regression coefficient on generic experience of one-half the slope of AK.



Appendix Fig. A2. Earnings Trajectories Estimated by Binned DiD-TR Model, by Gender and Education.

Appendix Table A1
DiD Estimates of Urban Earnings Premiums, by Education and Gender

VARIABLES	(1)	(2)	(3)	(4)	(5)
	All	Male Low	Male High	Fem Low	Fem High
Trit (annual change)	0.010***	−0.012***	0.014***	0.007***	0.006***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ti (rural-urban mover)	−0.090***	−0.060***	−0.076***	−0.041**	−0.049***
	(0.006)	(0.011)	(0.011)	(0.015)	(0.009)
Dit (residing in urban area)	0.228***	0.171***	0.187***	0.199***	0.229***
	(0.008)	(0.020)	(0.016)	(0.023)	(0.011)
Dit*Trit (change in urban)	−0.001	0.001	0.001	0.000	−0.005***

(*continued on next page*)

Appendix Table A1 (*continued*)

VARIABLES	(1)	(2)	(3)	(4)	(5)
	All	Male Low	Male High	Fem Low	Fem High
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
1st (Partial) Year Working	−0.212***	−0.002	−0.406***	0.010	−0.373***
	(0.007)	(0.011)	(0.017)	(0.016)	(0.013)
Married/Cohabiting	0.046***	0.117***	0.064***	−0.048***	−0.029***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)
Separated/Widowed	0.034***	0.033*	0.056	−0.007	0.002
	(0.008)	(0.015)	(0.029)	(0.013)	(0.011)
Year Child Born & Prior	−0.027***	−0.013*	−0.008	−0.073***	−0.069***
	(0.003)	(0.005)	(0.006)	(0.007)	(0.004)
Constant	9.934***	10.097***	12.398***	9.388***	11.724***
	(0.099)	(0.116)	(0.156)	(0.171)	(0.013)
Observations	316,733	150,472	51,207	48,593	66,461
Individuals	28,391	11,024	4769	5572	7026
R-squared	0.305	0.248	0.255	0.344	0.265
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.
Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

Appendix Table A2
FE Estimates of Urban Earnings Premiums, by Education and Gender

VARIABLES		(1)	(2)	(3)	(4)
	All	Male Low	Male High	Fem Low	Fem High
TRit (annual change)	0.005	0.001	0.022	−0.000	0.077*
	(0.011)	(0.013)	(0.078)	(0.017)	(0.033)
Dit (residing in urban area)	0.252***	0.139***	0.179***	0.201***	0.223***
	(0.009)	(0.026)	(0.017)	(0.026)	(0.012)
Dit * Trit (annual change	0.001	0.006***	0.003*	0.001	−0.003**
when in urban area)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
1st (Partial) Year Working	−0.250***	−0.060***	−0.462***	−0.076***	−0.416***
	(0.007)	(0.011)	(0.015)	(0.016)	(0.012)
Married/Cohabiting	−0.017***	−0.010	−0.003	−0.070***	−0.034***
	(0.004)	(0.006)	(0.007)	(0.008)	(0.005)
Separated/Widowed	−0.029*	−0.054*	−0.055	−0.017	−0.000
	(0.012)	(0.023)	(0.046)	(0.020)	(0.012)
Year Child Born & Prior	−0.030***	−0.014**	−0.014**	−0.076***	−0.068***
	(0.002)	(0.004)	(0.005)	(0.006)	(0.003)
Constant	10.482***	10.418***	12.332***	9.627***	11.861***
	(0.139)	(0.165)	(0.444)	(0.184)	(0.172)
Observations	316,733	150,472	51,207	48,593	66,461
Individuals	28,391	11,024	4769	5572	7026
Year FE	YES	YES	YES	YES	YES
Individual FE	YES	YES	YES	YES	YES
R-squared	0.254	0.239	0.245	0.341	0.138

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.
Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

Appendix Table A3
Binned DiD-TR model of earnings, by gender and education

Variable		Males		Females	
	Full Sample	Lower Education	Higher Education	Lower Education	Higher Education
Employment Experience:^					
2–3 years	0.211***	0.093***	0.416***	0.147***	0.372***
	(0.008)	(0.012)	(0.020)	(0.018)	(0.014)
4–5 years	0.290***	0.155***	0.493***	0.196***	0.481***
	(0.008)	(0.012)	(0.020)	(0.018)	(0.014)
6–7 years	0.299***	0.152***	0.508***	0.227***	0.500***
	(0.008)	(0.013)	(0.020)	(0.019)	(0.015)
8–9 years	0.299***	0.141***	0.547***	0.257***	0.514***
	(0.008)	(0.013)	(0.020)	(0.019)	(0.015)
10–11 years	0.307***	0.165***	0.580***	0.268***	0.527***
	(0.008)	(0.013)	(0.021)	(0.020)	(0.015)
12–13 years	0.311***	0.178***	0.604***	0.313***	0.540***
	(0.008)	(0.014)	(0.021)	(0.020)	(0.015)
14+ years	0.331***	0.223***	0.631***	0.360***	0.548***
	(0.008)	(0.014)	(0.020)	(0.019)	(0.015)
Ti (=1 if rural-urban mover; else = 0)	−0.319***	−0.331***	−0.201***	−0.105*	−0.200***

(*continued on next page*)

Appendix Table A3 (*continued*)

Variable	Full Sample	Males		Females	
		Lower Education	Higher Education	Lower Education	Higher Education
	(0.018)	(0.040)	(0.033)	(0.045)	(0.025)
Employment Experience by Movers:					
2–3 years	0.186***	0.157***	0.059	0.040	0.100***
	(0.020)	(0.046)	(0.038)	(0.052)	(0.029)
4–5 years	0.233***	0.234***	0.118**	0.019	0.107***
	(0.021)	(0.046)	(0.038)	(0.053)	(0.030)
6–7 years	0.302***	0.315***	0.213***	(0.004)	0.155***
	(0.022)	(0.046)	(0.041)	(0.059)	(0.036)
8–9 years	0.344***	0.352***	0.230***	0.073	0.230***
	(0.026)	(0.049)	(0.049)	(0.066)	(0.042)
10–11 years	0.350***	0.338***	0.219***	0.134	0.277***
	(0.032)	(0.054)	(0.065)	(0.080)	(0.058)
12–13 years	0.350***	0.342***	0.322***	0.146	0.306**
	(0.042)	(0.063)	(0.078)	(0.102)	(0.101)
14+ years	0.337***	0.375***	0.324**	(0.057)	(0.038)
	(0.059)	(0.071)	(0.098)	(0.238)	(0.100)
Dit (=1 if residing in urban region during current year; else = 0)	0.307***	0.158	0.329***	0.133	0.295***
	(0.021)	(0.097)	(0.037)	(0.082)	(0.024)
If Employment Experience is in Urban Area:					
4–5 years	−0.106***	(0.003)	−0.127**	0.079	−0.110***
	(0.023)	(0.102)	(0.041)	(0.087)	(0.027)
6–7 years	−0.158***	(0.025)	−0.207***	0.138	−0.153***
	(0.024)	(0.101)	(0.043)	(0.092)	(0.034)
8–9 years	−0.191***	(0.074)	−0.222***	0.053	−0.217***
	(0.027)	(0.101)	(0.052)	(0.095)	(0.040)
10–11 years	−0.196***	(0.046)	−0.222***	(0.013)	−0.252***
	(0.033)	(0.104)	(0.066)	(0.105)	(0.057)
12–13 years	−0.194***	(0.045)	−0.306***	(0.019)	−0.288**
	(0.044)	(0.109)	(0.080)	(0.122)	(0.101)
14+ years	−0.173**	(0.076)	−0.253*	0.206	0.074
	(0.060)	(0.114)	(0.100)	(0.248)	(0.100)
Year moved to urban area	−0.037***	(0.045)	−0.064***	0.008	−0.027*
	(0.010)	(0.025)	(0.017)	(0.027)	(0.013)
Year prior to moving to urban area	0.076***	0.078***	0.071***	0.089***	0.058***
	(0.006)	(0.014)	(0.012)	(0.018)	(0.009)
Constant	9.755***	10.104***	10.953***	9.414***	11.643***
	(0.099)	(0.116)	(0.343)	(0.171)	(0.048)
Observations	3,16,733	1,50,472	51,207	48,593	66,461
Individuals	28,391	11,024	4769	5572	7026
R-squared	0.309	0.267	0.273	0.359	0.291
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors clustered by individuals shown parenthetically; ***p < 0.001, **p < 0.01, *p < 0.05.
Sample includes only full-time, native workers; moves are from rural to urban centrality levels 1 and 2.
Note: all models include controls for relationship status & childbearing; full sample model includes controls for gender & education; see Tables 2-3.
ˆ Excluded reference category is first (likely partial) year of fulltime work.

Appendix Table A4
DiD-TR Model Estimates of Urban Earnings Premiums, by Gender and Education, Omitting Partnering Status and Childbearing Covariates

VARIABLES	(1)	(2)	(3)	(4)	(5)
	All	Male Low	Male High	Fem Low	Fem High
TRit	0.011***	0.017***	0.019***	0.019***	0.013***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ti	−0.254***	−0.242***	−0.233***	−0.097**	−0.192***
	(0.011)	(0.022)	(0.020)	(0.030)	(0.016)
Ti*TRit	0.034***	0.026***	0.035***	0.014*	0.038***
	(0.002)	(0.003)	(0.004)	(0.006)	(0.003)
Dit	0.387***	0.320***	0.326***	0.248***	0.341***
	(0.012)	(0.027)	(0.023)	(0.035)	(0.017)
Dit*TRit	−0.034***	−0.024***	−0.031***	−0.013*	−0.042***
	(0.002)	(0.003)	(0.004)	(0.006)	(0.004)
1st year work	−0.199***	−0.029**	−0.403***	−0.017	−0.364***
	(0.007)	(0.011)	(0.017)	(0.015)	(0.013)
Constant	9.938***	10.108***	11.316***	9.409***	11.984***
	(0.099)	(0.116)	(0.346)	(0.171)	(0.050)
Observations	316,733	150,472	51,207	48,593	66,461
R-squared	0.306	0.264	0.268	0.355	0.279
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.

Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

Appendix Table A6

DiD-TR Estimates of Urban Earnings Premiums, by Education and Gender; Rural-Urban Moves to Tier 1 Oslo Only

VARIABLES		(1)	(2)	(3)	(4)
	All	Male Low	Male High	Fem Low	Fem High
Trit (change in earnings)	0.010***	0.015***	0.016***	0.020***	0.012***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ti (rural-urban mover)	−0.277***	−0.284***	−0.276***	−0.146***	−0.210***
	(0.016)	(0.037)	(0.030)	(0.039)	(0.021)
Ti*Trit (movers change)	0.037***	0.031***	0.040***	0.018**	0.037***
	(0.003)	(0.005)	(0.006)	(0.006)	(0.005)
Dit (residing urban region)	0.415***	0.368***	0.329***	0.284***	0.348***
	(0.018)	(0.046)	(0.035)	(0.046)	(0.024)
Dit*Trit (change in urban)	−0.036***	−0.030***	−0.032***	−0.014*	−0.037***
	(0.003)	(0.005)	(0.006)	(0.006)	(0.005)
1st (partial) year after educ.	−0.190***	−0.030**	−0.394***	−0.024	−0.376***
	(0.007)	(0.011)	(0.019)	(0.016)	(0.013)
Married/cohabiting	0.052***	0.100***	0.072***	−0.052***	−0.017***
	(0.003)	(0.004)	(0.006)	(0.006)	(0.004)
Seprated/Widowed	0.042***	0.036*	0.096***	−0.001	0.022
	(0.008)	(0.015)	(0.027)	(0.013)	(0.012)
Child born this/last year	−0.028***	−0.012*	−0.021**	−0.078***	−0.077***
	(0.003)	(0.005)	(0.007)	(0.007)	(0.004)
Constant	9.935***	10.125***	11.072***	9.424***	12.072***
	(0.102)	(0.118)	(0.326)	(0.180)	(0.050)
Observations	280,852	141,043	40,230	44,675	54,904
	25,376	10,347	3842	5228	5959
R-squared	0.295	0.260	0.249	0.357	0.281
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.

Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

Appendix Table A7

DiD-TR Estimates of Nominal Urban Earnings Premiums, by Education and Gender; Using CPI as Covariate

VARIABLES		(1)	(2)	(3)	(4)
	All	Male Low	Male High	Fem Low	Fem High
Trit (change in earnings)	0.010***	0.016***	0.018***	0.020***	0.013***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ti (rural-urban mover)	−0.248***	−0.241***	−0.223***	−0.107***	−0.211***
	(0.011)	(0.022)	(0.020)	(0.030)	(0.016)
Ti*Trit (movers change)	0.034***	0.028***	0.037***	0.011	0.036***
	(0.002)	(0.003)	(0.004)	(0.006)	(0.003)
Dit (residing urban region)	0.382***	0.317***	0.316***	0.239***	0.340***
	(0.012)	(0.027)	(0.023)	(0.035)	(0.017)
Dit*Trit (change in urban)	−0.034***	−0.026***	−0.033***	−0.009	−0.038***
	(0.002)	(0.003)	(0.004)	(0.006)	(0.004)
CPI (2015 = 100)	0.062***	0.060***	0.014***	0.072***	0.032***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.000)
1st (partial) year after educ.	−0.199***	−0.027*	−0.402***	−0.024	−0.375***
	(0.007)	(0.011)	(0.017)	(0.015)	(0.013)
Married/cohabiting	0.048***	0.099***	0.066***	−0.056***	−0.022***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)
Seprated/Widowed	0.035***	0.033*	0.061*	−0.009	0.021
	(0.008)	(0.015)	(0.029)	(0.013)	(0.011)
Child born this/last year	−0.027***	−0.012*	−0.011	−0.077***	−0.077***
	(0.003)	(0.005)	(0.006)	(0.007)	(0.004)
Constant	5.752***	6.058***	10.323***	4.559***	9.712***
	(0.261)	(0.307)	(0.527)	(0.453)	(0.063)
Observations	316,733	150,472	51,207	48,593	66,461
Individuals	28,391	11,024	4769	5572	7026
R-squared	0.306	0.267	0.271	0.360	0.286
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.

Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

Appendix Table A8
DiD-TR Model Estimates of Urban Earnings Premiums, by Gender and Education, Controlling for Industrial Sector of Employment

VARIABLES	All	(1) Male Low	(2) Male High	(3) Fem Low	(4) Fem High
Trit (change in earnings)	0.011***	0.016***	0.015***	0.018***	0.012***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ti (rural-urban mover)	−0.270***	−0.251***	−0.250***	−0.120***	−0.204***
	(0.010)	(0.021)	(0.020)	(0.029)	(0.016)
Ti*Trit (movers change)	0.036***	0.029***	0.039***	0.013*	0.034***
	(0.002)	(0.003)	(0.004)	(0.005)	(0.003)
Dit (residing urban region)	0.370***	0.320***	0.314***	0.257***	0.322***
	(0.011)	(0.027)	(0.022)	(0.034)	(0.017)
Dit*Trit (change in urban)	−0.035***	−0.026***	−0.035***	−0.013*	−0.037***
	(0.002)	(0.003)	(0.004)	(0.005)	(0.003)
1st (partial) year after educ.	−0.205***	−0.034**	−0.407***	−0.026	−0.372***
	(0.007)	(0.011)	(0.017)	(0.015)	(0.013)
Married/cohabiting	0.042***	0.086***	0.063***	−0.051***	−0.025***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)
Seprated/Widowed	0.027***	0.013	0.050	−0.006	0.020
	(0.007)	(0.014)	(0.027)	(0.013)	(0.011)
Child born this/last year	−0.029***	−0.009*	−0.014**	−0.079***	−0.078***
	(0.003)	(0.005)	(0.006)	(0.006)	(0.004)
Constant	10.002***	10.114***	11.377***	9.423***	11.919***
	(0.100)	(0.116)	(0.338)	(0.172)	(0.050)
Observations	316,733	150,472	51,207	48,593	66,461
R-squared	0.375	0.352	0.362	0.396	0.320
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.
Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

Appendix Table A9
FE Model Estimates of Urban Earnings Premiums, by Gender and Education, Replicating Carlsen et al. (2016) Specification

VARIABLES	All	(1) Male Low	(2) Male High	(3) Fem Low	(4) Fem High
Working in Tier 1 (Oslo)	0.240***	0.104*	0.126***	0.189***	0.214***
	(0.017)	(0.052)	(0.029)	(0.042)	(0.023)
Working in Tier 2 City	0.239***	0.108***	0.190***	0.196***	0.222***
	(0.014)	(0.026)	(0.022)	(0.038)	(0.025)
T1*Trit (Oslo experience)	0.006***	0.008*	0.015***	0.004	0.004
	(0.001)	(0.003)	(0.003)	(0.003)	(0.002)
T2*Trit (Tier 2 experience)	0.001	0.005**	0.000	−0.001	−0.003
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)
Trit (years of experience)	0.012	0.011	−0.034	−0.003	−0.014
	(0.011)	(0.013)	(0.068)	(0.025)	(0.069)
1st (partial) year worked	−0.209***	−0.047***	−0.471***	−0.073***	−0.405***
	(0.008)	(0.010)	(0.016)	(0.021)	(0.024)
Married/cohabiting	−0.011**	−0.009	0.002	−0.062***	−0.032***
	(0.004)	(0.005)	(0.008)	(0.010)	(0.009)
Seprated/Widowed	−0.038**	−0.053***	−0.029	−0.025	0.010
	(0.012)	(0.016)	(0.039)	(0.025)	(0.031)
Had child this/prior year	−0.017***	−0.006	−0.013**	−0.073***	−0.073***
	(0.003)	(0.004)	(0.005)	(0.007)	(0.006)
Constant	10.453***	10.415***	11.733***	9.598***	11.630***
	(0.134)	(0.157)	(0.358)	(0.190)	(0.353)
Observations	219,458	134,099	35,634	30,140	19,585
R-squared	0.383	0.383	0.496	0.400	0.436
Number of individuals	21,102	10,550	3793	3898	2861
Year FE	YES	YES	YES	YES	YES
Individual FE	YES	YES	YES	YES	YES
Industry	YES	YES	YES	YES	YES

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.

Appendix Table A10
DiD-TR Model Estimates of Urban Earnings Premiums, by Gender and Education, Replicating Carlsen et al. (2016) Specification

VARIABLES		(1)	(2)	(3)	(4)
	All	Male Low	Male High	Fem Low	Fem High
Working in Tier 1 (Oslo)	0.469***	0.250***	0.346***	0.324***	0.423***
	(0.016)	(0.037)	(0.031)	(0.044)	(0.032)
Working in Tier 2 City	0.399***	0.250***	0.379***	0.272***	0.343***
	(0.015)	(0.027)	(0.028)	(0.038)	(0.034)
T1*Trit (Oslo experience)	−0.035***	−0.016***	−0.031***	−0.016***	−0.036***
	(0.002)	(0.003)	(0.005)	(0.005)	(0.006)
T2*Trit (Tier 2 experience)	−0.038***	−0.020***	−0.043***	−0.021***	−0.041***
	(0.002)	(0.003)	(0.005)	(0.005)	(0.007)
Trit (years of experience)	0.012***	0.017***	0.016***	0.024***	0.024***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Ti (urban migrant)	−0.309***	−0.223***	−0.310***	−0.144***	−0.213***
	(0.012)	(0.020)	(0.024)	(0.030)	(0.028)
Ti*Trit (migrant exp.)	0.038***	0.024***	0.045***	0.020***	0.036***
	(0.002)	(0.003)	(0.005)	(0.004)	(0.006)
1st (partial) year worked	−0.182***	−0.026**	−0.419***	−0.038	−0.361***
	(0.008)	(0.009)	(0.019)	(0.020)	(0.026)
Married/cohabiting	0.055***	0.073***	0.073***	−0.044***	0.012
	(0.003)	(0.003)	(0.006)	(0.006)	(0.008)
Seprated/Widowed	0.015	0.007	0.085**	−0.010	0.015
	(0.009)	(0.013)	(0.030)	(0.017)	(0.031)
Had child this/prior year	−0.015***	−0.005	−0.009	−0.077***	−0.078***
	(0.003)	(0.004)	(0.006)	(0.008)	(0.008)
Constant	9.953***	10.104***	10.978***	9.430***	11.704***
	(0.100)	(0.116)	(0.406)	(0.173)	(0.110)
Observations	219,458	134,099	35,634	30,140	19,585
R-squared	0.395	0.356	0.387	0.420	0.379
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.

Appendix Table A11
OLS Estimates of Urban Migrants' Cumulative Growth of Earnings while Working in Rural Areas, by Gender and Education

VARIABLES	(1)	(2)	(3)	(4)	(5)
	All	Male Low	Male High	Fem Low	Fem High
Married/Cohabiting	−0.141	−0.539	−0.349	−1.690	0.302
	(0.277)	(1.181)	(0.346)	(1.014)	(0.297)
Separated/Widow (er)	−1.684	−7.692	−0.322	0.713	−3.220***
	(2.221)	(5.782)	(0.477)	(3.807)	(0.660)
Had Child 1–2 Years	−1.083*	−3.918**	0.041	0.175	−0.372
Prior	(0.424)	(1.432)	(0.585)	(1.195)	(0.588)
Female w/Lower	−2.880***				
Education	(0.566)				
Male w/Higher	−8.771***				
Education	(0.478)				
Female w/Higher	−8.746***				
Education	(0.454)				
Public Admin. Sector	−1.882***	−1.729	−1.547**	−1.124	−1.638***
1st Year Work	(0.388)	(1.596)	(0.532)	(1.744)	(0.488)
Education Sector	0.882*	−8.344*	1.365*	3.579	1.223**
1st Year Work	(0.385)	(3.428)	(0.642)	(3.108)	(0.435)
Primary Sector	−0.718	−3.688	0.983	5.481	−2.496
1st Year Work	(1.420)	(4.686)	(1.690)	(4.685)	(1.635)
Health Sector	−0.455	−0.365	−0.987*	−3.200**	−0.072
1st Year Work	(0.253)	(2.113)	(0.502)	(0.982)	(0.277)
# Sectoral Changes	1.592***	1.403***	1.600***	1.539***	1.459***
Before Moving	(0.119)	(0.334)	(0.131)	(0.357)	(0.126)
Constant	5.872***	3.299***	−0.347	2.992	−0.312
	(0.566)	(0.000)	(0.544)	(.)	(1.082)
Observations	4070	795	1383	374	1518
R-squared	0.343	0.297	0.226	0.327	0.227
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors in parentheses; dependent variable is logged; sample is rural-urban movers only.
****p < 0.001, **p < 0.01, *p < 0.05.*

Appendix Table A12

Mean Annual Changes Between Industrial Sectors, by Movers & Stayers, Gender, Education

Mean annual changes in sector:	All	Male Low	Male High	Fem Low	Fem High
Stayers, while working in rural	0.146 (0.140)	0.155 (0.128)	0.157 (0.152)	0.142 (0.131)	0.124 (0.152)
Movers, while working in rural	0.279 (0.255)	0.231 (0.216)	0.307 (0.264)	0.265 (0.230)	0.285 (0.271)
Movers, while working in urban	0.110 (0.125)	0.133 (0.139)	0.111 (0.116)	0.118 (0.116)	0.094 (0.120)
Proportion changing sector when move	0.163	0.172	0.184	0.134	0.151

Appendix Table A5

Earnings Functions Estimated by DiD-TR, by Gender and Education, Excluding those Staying in Same Rural Location Throughout.

VARIABLES	(1)	(2)	(3)	(4)	(5)
	All	Male Low	Male High	Fem Low	Fem High
Trit (change in earnings)	0.014***	0.023***	0.025***	0.022***	0.016***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Ti (rural-urban mover)	−0.213***	−0.166***	−0.229***	−0.051	−0.191***
	(0.012)	(0.025)	(0.024)	(0.033)	(0.018)
Ti*Trit (movers change)	0.028***	0.020***	0.034***	0.004	0.030***
	(0.002)	(0.003)	(0.004)	(0.006)	(0.004)
Dit (residing urban region)	0.343***	0.252***	0.308***	0.202***	0.313***
	(0.012)	(0.030)	(0.024)	(0.038)	(0.018)
Dit*Trit (change in urban)	−0.029***	−0.019***	−0.030***	−0.004	−0.032***
	(0.002)	(0.004)	(0.004)	(0.006)	(0.004)
1st (partial) year after educ.	−0.261***	−0.059*	−0.412***	0.044	−0.400***
	(0.012)	(0.024)	(0.025)	(0.026)	(0.019)
Married/cohabiting	0.017***	0.077***	0.048***	−0.067***	−0.036***
	(0.004)	(0.007)	(0.007)	(0.008)	(0.005)
Seprated/Widowed	0.006	0.047*	0.043	−0.055*	−0.017
	(0.013)	(0.022)	(0.039)	(0.028)	(0.021)
Child born this/last year	−0.028***	−0.012	0.006	−0.079***	−0.073***
	(0.004)	(0.009)	(0.007)	(0.011)	(0.005)
Constant	10.088***	10.177***	9.668***	9.495***	12.052***
	(0.152)	(0.188)	(0.516)	(0.224)	(0.166)
Observations	121,256	40,347	27,602	18,173	35,134
Individuals	10,856	2987	2455	1960	3454
R-squared	0.364	0.327	0.321	0.398	0.302
Year FE	YES	YES	YES	YES	YES
Individual FE	NO	NO	NO	NO	NO

Robust standard errors in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05.

Variables related to gender, education have been included in the non-stratified regressions; their related coefficients are not shown here.

References

Ahlfeldt, G., Moeller, K., Waights, S., Wendland, N., 2017. Game of zones: the political economy of conservation areas. Econ. J. 127, F421–F445.

Ahlin, L., Andersson, M., Thulin, P., 2014. Market thickness and the early labour market career of university graduates: an urban advantage? Spatial Econ. Anal. 9, 396–419.

Altonji, J., Blank, R., 1999. Race and gender in the labor market. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3. Elsevier/North-Holland, Amsterdam, pp. 3144–3260.

Ananat, E., Shine, F., Ross, S.L., 2018. Race-specific urban wage premia and the black–white wage gap. J. Urban Econ. 108 (1), 141–153.

Andersson, F., Burgess, S., Lane, J., 2007. Cities, matching and the productivity gains of agglomeration. J. Urban Econ. 61 (1), 112–128.

Andersson, M., Klaesson, J., Larsson, J., 2014. The sources of the urban wage premium by worker skills: spatial sorting or agglomeration economies? Pap. Reg. Sci. 93, 727–747.

Bacolod, M., 2017. Skills, the gender wage gap, and cities. J. Reg. Sci. 57 (2), 290–318.

Bacolod, M., Blum, B., Strange, W., 2009. Skills in the city. J. Urban Econ. 65, 136–153.

Baum-Snow, N., Freedman, M., Pavan, R., 2018. Why has urban inequality increased? Am. Econ. J. Appl. Econ. 10 (4), 1–42.

Baum-Snow, N., Pavan, R., 2012. Understanding the city size wage gap. Rev. Econ. Stud. 79, 88–127.

Baum-Snow, N., Pavan, R., 2013. Inequality and city size. Rev. Econ. Stat. 95 (5), 1535–1548.

Behrens, K., Robert-Nicoud, F., 2014. Survival of the fittest in cities: urbanisation and inequality. Econ. J. 124 (581), 1371–1400.

Bennett, P., Butikofer, A., Salvanes, K.G., Steskal, D., 2022. Changes in Urban Wages, Jobs, and Workers from 1958-2017. https://doi.org/10.2139/ssrn.4123826. CESifo Working Paper No. 9766, Av at SSRN: https://ssrn.com/abstract=4123826.

Bleakley, H., Lin, J., 2012. Thick-market effects and churning in the labor market: evidence from US cities. J. Urban Econ. 72 (2–3), 87–103.

Borgonovi, F., Ferrara, A., Maghnouj, S., 2018. The Gender Gap in Educational Outcomes in Norway. OECD Education Working Paper No. 183. EDU/WKP, p. 19 (2018).

Borjas, G., Bronars, S., Trejo, S., 1992. Self-selection and internal migration in the United States. J. Urban Econ. 32, 159–185.

Ciccone, A., Hall, R., 1996. Productivity and the density of economic activity. Am. Econ. Rev. 86, 54–70.

Carlsen, F., Rattsø, J., Stokke, H., 2016. Education, experience, and urban wage premia. Reg. Sci. Urban Econ. 60, 39–49.

Champion, T., Coombes, M., Gordon, I., 2014. How far do England's second-order cities emulate London as human-capital 'escalators'? Popul. Space Place 20 (5), 421–433.

Combes, P.P., Démurger, S., Li, S., 2015. Migration externalities in China. Eur. Econ. Rev. 76, 152–167.

Combes, P.P., Démurger, S., Li, S., 2017. Productivity gains from agglomeration and migration in the people's Republic of China between 2002 and 2013. Asian Dev. Rev. Stud. Asian Pac. Econ. Issues 34 (2), 184–200.

Combes, P.P., Duranton, G., Gobillon, L., 2008. Spatial wage disparities: sorting matters. J. Urban Econ. 63 (2), 723–742.

Combes, P.P., Duranton, G., Gobillon, L., Roux, S., 2012. Sorting and local wage and skill distributions in France. Reg. Sci. Urban Econ. 42, 913–930.

Combes, P.P., Gobillon, L., 2015. The empirics of agglomeration economies. In: Duranton, G., Henderson, J.V., Strange, W. (Eds.), Handbook of Regional and Urban Economics, vol. 5. Elsevier/North-Holland, Amsterdam, pp. 247–348.

Davis, D., Dingel, J., 2019. A spatial knowledge economy. Am. Econ. Rev. 109 (1), 153–170.

D'Costa, S., Overman, H., 2014. The urban wage growth premium: sorting or learning? Reg. Sci. Urban Econ. 48, 168–179.

De la Roca, J., 2017. Selection in initial and return migration: evidence from moves across Spanish cities. J. Urban Econ. 100, 33–53.

De la Roca, J., Puga, D., 2017. Learning by working in big cities. Rev. Econ. Stud. 84 (1), 106–142.

Duranton, G., Puga, D., 2004. Micro-foundations of urban economies. In: Henderson, J. V., Thisse, J.-F. (Eds.), Handbook of Urban and Regional Economics, vol. 4. Elsevier/North-Holland, Amsterdam, pp. 2063–2117.

DuMond, J., Hirsch, B., MacPherson, D., 1999. Wage differentials across labor markets and workers: does cost of living matter? Econ. Inq. 37, 577–598.

Eeckhout, J., Pinheiro, R., Schmidheiny, K., 2014. Spatial sorting. J. Polit. Econ. 122 (3), 554–620.
Fielding, A., 1992. Migration and social mobility: south East England as an escalator region. Reg. Stud. 26 (1), 1–15.
Fielding, A., 2007. Migration and social mobility in urban systems: national and international trends. In: Geyer, H.S. (Ed.), International Handbook of Urban Policy. Edward Elgar, Cheltenham, UK, pp. 107–137.
Frank, R., 1978. Why women earn less: the theory and estimation of differential over-qualification. Am. Econ. Rev. 68, 360–373.
Fuchs, V., 1967. Differentials in Hourly Earnings by Region and City Size, 1959. NY: NBER Occasional Paper 101.
Galster, G., Pettit, K., Santiago, A., Tatian, P., 2002. The impact of supportive housing on neighborhood crime rates. J. Urban Aff. 24 (3), 289–315.
Galster, G., Tatian, P., Accordino, J., 2006. Targeting investments for neighborhood revitalization. J. Am. Plann. Assoc. 72 (4), 457–474.
Galster, G., Tatian, P., Smith, R., 1999. The impact of neighbors who use section 8 certificates on property values. Housing Policy Debate 10 (4), 879–917.
Galster, G., Temkin, K., Walker, C., Sawyer, N., 2004. Measuring the impacts of community development initiatives: a new application of the adjusted interrupted time-series method. Eval. Rev. 28 (6), 1–38.
Glaeser, E., 1998. Are cities dying. J. Econ. Perspect. 12, 139–160.
Glaeser, E., 1999. Learning in cities. J. Urban Econ. 46, 254–277.
Glaeser, E., Maré, D., 2001. Cities and skills. J. Labor Econ. 19, 316–342.
Glaeser, E., Resseger, M., 2010. The complementarity between cities and skills. J. Reg. Sci. 50 (1), 221–244.
Greenstone, M., Hornbeck, R., Moretti, E., 2010. Identifying agglomeration spillovers: evidence from winners and losers of large plant openings. J. Polit. Econ. 118 (3), 536–598.
Gordon, I., 2015. Ambition, human capital acquisition and the metropolitan escalator. Reg. Stud. 49 (6), 1042–1055.
Gordon, I., Champion, T., Combes, M., 2015. Urban escalators and interregional elevators: the difference that location, mobility, and sectoral specialization make to occupational progression. Environ. Plann. 47 (3), 588–606.
Gould, E., 2007. Cities, workers, and wages: a structural analysis of the urban wage premium. Rev. Econ. Stud. 74, 477–506.
Granovetter, M., 1995. Getting a Job: A Study of Contacts and Careers, second ed. Univ. of Chicago Press, Chicago.
Gyourko, J., Mayer, C., Sinai, T., 2013. Superstar cities. Am. Econ. J. Econ. Pol. 5 (4), 167–199.
Helsey, R., Strange, W., 1990. Matching and agglomeration economies in a system of cities. Reg. Sci. Urban Econ. 28, 189–212.
Høydahl, E., 2017. Ny sentralitetsindeks for kommunene, Statistics Norway. Documents 2017 (40).
Høydahl, E., 2020. Sentralitetsindeksen. Oppdatering med 2020-kommuner. Statistics Norway, Documents 2020 (4).
Hwang, H., Reed, W., Hubbard, C., 1992. Compensating wage differentials and unobserved productivity. J. Polit. Econ. 100, 835–858.
Kim, S., 1987. Diversity in urban labor markets and agglomeration economies. Pap. Reg. Sci. Assoc. Reg. Sci. Assoc. Meet. 62, 57–70.
Kim, S., 1990. Labor heterogeneity, wage bargaining, and agglomeration economies. J. Urban Econ. 28, 160–177.
Korpi, M., Clark, W.A.V., 2019. Migration and occupational careers: the static and dynamic urban wage premium by education and city size. Pap. Reg. Sci. 98 (1), 555–574.
Kracke, N., Reichelt, M., Vicari, B., 2018. Wage losses due to overqualification: the role of formal degrees and occupational skills. Soc. Indicat. Res. 139, 1085–1108.
Lee, S., 2010. Ability sorting and consumer city. J. Urban Econ. 68 (2010), 20–33.
Liu, S., 2017. Agglomeration, urban wage premiums, and college majors. J. Reg. Sci. 57 (4), 611–630.
Matano, A., Naticchioni, P., 2012. Wage distribution and the spatial sorting of workers. J. Econ. Geogr. 12 (2), 379–408.
Matano, A., Naticchioni, P., 2016. What drives the urban wage premium? J. Reg. Sci. 56 (2), 191–209.
Mion, G., Naticchioni, P., 2009. The spatial sorting and matching of skills and firms. Can. J. Econ. 42 (1), 28–55.
Möller, J., Haas, A., 2003. The agglomeration wage differential reconsidered: an investigation using German micro data 1984–1997. In: Bröcker, J., Dohse, D., Soltwedel, R. (Eds.), Innovation Clusters and Interregional Competition. Springer, Berlin and Heidelberg, pp. 182–217.
Moretti, E., 2004. Workers' education, spillovers and productivity: evidence from plant-level production functions. Am. Econ. Rev. 94, 656–690.
Moretti, E., 2013. Real wage inequality. Am. Econ. J. Appl. Econ. 5 (1), 65–103.
Mulder, C., van Ham, M., 2005. Migration histories and occupational achievement. Popul. Space Place 11 (3), 173–186.
Niedomysl, T., Fransson, U., 2014. On distance and the spatial dimension in the definition of internal migration. Ann. Assoc. Am. Geogr. 104, 357–372.
Nisic, N., 2017. Smaller differences in bigger cities? Assessing the regional dimension of the gender wage gap. Eur. Socio Rev. 33 (2), 292–3044.
OECD, 2020. Education at a Glance: OECD Indicators. https://doi.org/10.1787/69096873-en. Retrieved from.
Phimister, E., 2005. Urban effects on participation and wages: are there gender differences? J. Urban Econ. 58 (3), 513–536.
Piketty, T., Saez, E., 2014. Inequality in the long run. Science 344 (6186), 838–843.
Puga, D., 2010. The magnitude and causes of agglomeration economies. J. Reg. Sci. 50, 203–219.
Rauch, J., 1993. Productivity gains from geographic concentration of human capital: evidence from the cities. J. Urban Econ. 34, 380–400.
Roback, J., 1982. Wages, rents, and the quality of life. J. Polit. Econ. 90, 1257–1278.
Rosenthal, S., Strange, W., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, J.V., Thisse, J. (Eds.), Handbook of Regional and Urban Economics, ume 4. Elsevier/North-Holland, Amsterdam, pp. 2119–2171.
Rosenthal, S., Strange, W., 2012. Female entrepreneurship, agglomeration, and a new spatial mismatch. Rev. Econ. Stat. 94 (3), 764–788.
Rye, J., 2006. Leaving the countryside. Acta Sociol. 49 (1), 47–65.
Rosen, S., 1979. Wage-based indexes of urban quality of life. In: Mieszkowski, P., Strazheim, M. (Eds.), Current Issues in Urban Economics. Johns Hopkins University Press, Baltimore, pp. 74–104.
Sato, Y., 2001. Labor heterogeneity in an urban labor market. J. Urban Econ. 50, 313–337.
Slettebakk, M.H., 2021. Labour migration and increasing inequality in Norway. Acta Sociol. 64 (3), 314–330.
Statistics Norway, 2019. https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/disse-kommunene-har-storst-og-minst-inntektsulikhet. (Accessed 19 March 2021).
Steskal, D., 2015. Essays on Urban Wage Premium: Returns to Internal Migration and Migrants' Selection in Norway. PhD dissertation, Department of Economics, NHH - Norwegian School of Economics. Available at: https://openaccess.nhh.no/nhh-xmlui/bitstream/handle/11250/2412919/Darina%20Seskal_dissertation.pdf?sequence=1&isAllowed=y.
Verbeek, M., Nijman, T., 1992. Testing for selectivity bias in panel data models. Int. Econ. Rev. 33, 681–703.
Wang, Z., 2016. Wage growth, ability sorting, and location choice at labor-force entry: new evidence from U.S. Census data. J. Urban Econ. 96, 112–120.
Wessel, T., Magnusson Turner, L., 2020. The migration pathway to economic mobility: does gender matter? Popul. Space Place. https://doi.org/10.1002/psp.2419.
Wheeler, C., 2001. Search, sorting, and urban agglomeration. J. Labor Econ. 19 (4), 879–899.
Wheeler, C., 2006. Cities and the growth of wages among young workers: evidence from the NLSY. J. Urban Econ. 60, 162–184.
Wixe, S., 2015. The impact of spatial externalities: skills, education and plant productivity. Reg. Stud. 49 (12), 2053–2069.
Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, MA.
Yankow, J., 2006. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. J. Urban Econ. 60, 139–161.