

Rural-Urban Migration and Socioeconomic Mobility in Victorian Britain

JASON LONG

This article analyzes rural-urban migration in Great Britain in the latter half of the nineteenth century. Using a new dataset of 28,000 individuals matched between the 1851 and 1881 population censuses, I examine the selection process and treatment effect of migration, controlling for the endogeneity of the migration decision. I find that urban migrants were positively selected—the best of the rural labor pool—and that the economic benefits of migration were substantial. Migrants responded to market signals, and labor markets were largely efficient; however, not all gains from migration were exploited, potentially indicating some degree of inefficiency.

There is a large and very old literature examining internal migration in nineteenth-century Great Britain. Several factors combined to make the populace mobile. The uneven spread of industrialization and economic modernization created wage differentials that induced migration. Well-developed roadways and increasing rail coverage kept migration costs low. In addition, unlike in the early twentieth century, which saw the rise of national social welfare programs and widespread homeownership, virtually no adverse institutional incentives inhibited mobility. The resultant high rate of internal mobility, particularly from the rural to the fast-growing urban areas, was one of the most prevalent demographic features of nineteenth-century Britain. From 1841 to 1901 the rural areas of England and Wales lost more than 4 million people from internal migration, 3 million of whom moved to towns, at a rate of more than half a million per decade.¹ The 1851 population census re-

The Journal of Economic History, Vol. 65, No. 1 (March 2005). © The Economic History Association. All rights reserved. ISSN 0022-0507.

Jason Long is Assistant Professor, Department of Economics, Colby College, 5243 Mayflower Hill Drive, Waterville, ME 04901. E-mail: jmlong@colby.edu.

This article is based on the first chapter of my Ph.D. dissertation, *Labor Mobility in Victorian Britain*, Northwestern University, 2002. I am grateful to my dissertation advisors, Joel Mokyr, Joseph Ferrie, and Joseph Altonji for many helpful comments. I also benefited from conversations with Henry Siu, James Sullivan, and Christopher Taber, and from input from workshop participants at Northwestern University and participants of the 2002 All-UC Group in Economic History Conference, the 2002 Cliometrics Conference, and the 2002 Economic History Association meeting. Anonymous referees for this and another journal made careful, helpful comments on previous versions of this article. Justin Hayes and Humphrey Southall provided me with data. This research was supported by a Northwestern University Graduate Research Grant and a Pew Younger Scholars Program Fellowship.

¹ Crouzet, *Victorian Economy*, p. 93.

vealed a watershed demographic event: for the first time in the history of any large nation, more people lived in towns than in the countryside.²

This mobility also had important efficiency implications for Britain's growing economy, as labor reallocated itself in the wake of the Industrial Revolution from areas of surplus to areas where it was in demand to fill factories. This redistribution fueled growth in at least two ways. First, labor moved from places of lower marginal product (typically rural) to higher (typically urban), resulting in efficiency gains for the economy.³ Second, cities grew. Modern economic growth theory, with its emphasis on increasing returns and external scale economies, has recognized that the localized information and knowledge spillover of urbanization make cities the "engines of economic growth in an economy."⁴

Considering the magnitude and the demographic and economic implications of internal migration in nineteenth-century Britain, it is natural that scholars have studied it intensely. One recent survey lists 15 books and articles dating from the pioneering work of Earnst Ravenstein in 1885 to the present and covering every decade from the 1920s to the 1990s.⁵ There are two principal strands of the literature. The first examines the extent, patterns, and determinants of internal migrant flows. Important examples include the early work of Ravenstein, who developed the well-known "laws of migration," and the studies of later scholars such as Arthur Redford and Dudley Baines.⁶ Redford, Baines, and others have shown that most migrants were young, moved only short distances, and often moved from rural to urban areas. Migrants responded to wage differentials but were limited by the aforementioned tendencies—the old rarely moved, and the young favored short-distance moves. The second strand of the migration literature analyzes the efficiency of British labor markets. Many have claimed that nineteenth-century British labor markets failed, impeding economic growth. Large and persistent rural-urban wage gaps were indicative of insufficient mobility, and excessive urban labor costs slowed industrialization.⁷ Others have argued that, on the contrary, migrants did respond to wage gaps and that excess agricultural labor left for the cities early and fueled industrialization and growth.⁸

² Towns here are places of more than 2,500 inhabitants. See the "Data" section for the definition of "towns" and "cities."

³ These gains are often depicted graphically with the "Harberger triangle" of deadweight loss.

⁴ Black and Henderson, "Urban Growth," pp. 252–53. Also, Lucas, "Mechanics."

⁵ Boyer and Hatton, "Migration."

⁶ Ravenstein, "Laws"; Redford, *Labour Migration*; and Baines, *Migration*.

⁷ See, for example, Pollard, "Labour"; and Williamson, *Coping*.

⁸ See especially Crafts, *British Economic Growth*.

Though these studies, which generally analyze migration at some aggregate level such as the county, are tremendously valuable, each suffers from a limitation common to empirical investigations of pre-twentieth-century labor markets: a lack of nationally representative micro-level panel data with which to observe changes over time in the lives of individuals.⁹ Without such data, we can know little about the individual-level forces that drove the migrants, about the selection process by which some moved and others did not, and about the economic return to migration. We can observe the connection between average wages and aggregate migration rates, but we cannot observe the extent to which individuals did (or did not) move in response to expected advancement in the labor market.

This study addresses these questions with a new dataset of approximately 28,000 individuals matched between the 1851 and 1881 Censuses of the Population of England and Wales. I develop a structural econometric model to analyze the selection and treatment effects of migration, controlling for the endogeneity of the migration decision. I find that the urban migrants of nineteenth-century Britain had not been the most destitute rural residents. They in fact had better rural prospects than did those who remained behind, and they were positively selected for migration: they performed better in the urban labor markets than the rural persisters would have had they instead chosen to move to an urban place. Potential migrants responded to labor market signals: they were more likely to move if they anticipated economic gains. In addition, the treatment effect of moving to a city was positive and large across all socioeconomic strata. These findings indicate that labor markets were generally effective in allocating labor from rural to urban areas. However, it appears that not all gains from moving were exploited: those who remained in rural areas also could have benefited from urban migration, though to a lesser extent than those who actually moved. The results, then, are largely, but not entirely, positive with respect to the efficiency of nineteenth-century British labor markets.

THE DATA

This study uses data from a new sample of 28,474 males linked from the 1851 Census of the Population of England and Wales to the 1881 census. The population censuses are the most important source of indi-

⁹ Williamson, *Coping*, includes micro-level analysis, but on individuals observed only at one point in time. Pooley and Turnbull, *Migration*, examines impressive panel data on individuals, but the data are chronologically diffuse and may not accurately represent the population, and the authors do not include any econometric analyses of migration.

vidual-level data for Great Britain from the nineteenth century. Each census, however, is only a cross-sectional snapshot of the population. Successive censuses can be used to examine changes in the nation over time but not changes in the lives of individuals over time, as no continuity exists for individuals between the censuses. But continuity can be created. The Genealogical Society of Utah in conjunction with the Federation of Family History Societies has recently computerized the entire 1881 census of the population of England and Wales. With these data, one can search for any individual or group of individuals in the 1881 census. For the dataset used in this study, I searched for individuals from a computerized, nationally representative 2-percent sample of the 1851 census.¹⁰

For each individual, the censuses include name, address, relationship to head of household, marital status, age, sex, occupation, county and parish of birth, and whether blind or deaf and dumb. I used name, birth year, and county and parish of birth—the four pieces of information that should not have changed between enumerations—to link individuals between the two censuses. In order to be considered a true match for an individual from 1851, an individual from 1881 had to have the same name, a year of birth different by no more than five years, and the same county and parish of birth.¹¹ If an individual from the 1851 sample had more than one match in 1881, I dropped that individual from the sample.

Applying this matching process to an initial pool of 168,130 English and Welsh males from the 1851 2-percent sample yielded a set of 28,474 males observed both in 1851 and 1881, a success rate of 17 percent. Using age-specific mortality and emigration rates, I estimate that approximately 85,000 of the males in the initial sample would be expected to die by 1881, and that 13,500 would be expected to migrate out of England and Wales.¹² The expected match success rate, then, is 41 percent, substantially greater than the 17 percent I successfully matched. Part of the discrepancy comes from duplicate matching: I

¹⁰ The 1851 sample was compiled principally by Anderson, Collins, and Stott and is available from the Data Archive at the University of Essex as study number 1316. Regarding the construction of the sample, it is noteworthy that the clustering procedure ensures that family units remain intact; for every individual in the sample, information is included for each member of that individual's household, including immediate family members and anyone else residing in the same dwelling place. For a full description, see Anderson, *National Sample*. The 1881 census is also available from the Data Archive, as study 3643.

¹¹ I allowed name and parish of birth to have slight phonetic variation. I allowed the variation in birth year to account for age misreporting, which was common in nineteenth-century Britain. See the section "Estimation of the Model." I also visually inspected the data to remove common name-matching errors.

¹² Mortality rates are from Mitchell, *Abstract*, pp. 38–39. Emigration rates are from Baines, *Migration*, pp. 152–53.

eliminated 9,622 individuals from the 1851 sample because they matched to more than one person in the 1881 census. The gap between the expected 41 percent success rate and the 23 percent of individuals who actually were matched (uniquely or not) must be attributed to enumeration error—the individuals misreported their age by more than five years, or the birthplace or name information given in the two censuses could not be reconciled, or they simply were not enumerated in 1881.

The effect of emigration on the sample deserves particular attention. Unlike attrition due to mortality, emigration from England and Wales represents self-selection out of the sample. Emigration rates were high in the second half of the nineteenth century, particularly for young adult males.¹³ Overseas migration may have been an alternative to urban migration. It may also have been part of a stage migration process whereby rural residents migrated first to cities within Britain then subsequently to destinations overseas. The selection process of emigration is therefore an important part of the story, but one that I cannot address with these data. It is possible, however, to match males from the 1851 British census sample into the recently computerized 1880 U.S. population census in order to analyze the process of trans-Atlantic migration from England and Wales. This effort is part of an ongoing research project.¹⁴

Considering the substantial rate of enumeration error present in the matching process, it is important to consider carefully the representativeness of the matched sample. Table 1 shows a comparison between the sample of matched individuals and the entire group of males from the 1851 2-percent sample.¹⁵ The table includes information for the matched sample in 1881 for comparison, but it is the first two columns, which show 1851 values for both groups, which should be used to gauge sample representativeness. There are differences between the groups throughout the categories, but they all stem from one factor: the matched men are younger on average. The average age for the matched men is 18, whereas for the entire group it is 24. This should be expected, as younger men would have been more likely to survive the 30 years necessary to be matched.¹⁶ From this one difference follow the others: matched men were more likely to be sons rather than household

¹³ Baines finds that if the at-risk pool of potential male migrants is assumed to be those aged 15–24, the average emigration rate for the period was 15.4 percent per decade. See *Migration*, p. 153.

¹⁴ To date, I have been able to match 6,000 males from the 1851 sample into the 1880 U.S. census. See Long and Ferrie, “Tale.”

¹⁵ It also includes the smaller subsample of individuals used in the econometric analysis. I discuss this group in what follows.

¹⁶ Life expectancy in England at birth and at age 20 were, respectively, 39.5 and 40.4 years. See Woods, *Demography*, p. 224.

TABLE 1
COMPARISON OF ALL MALES, MATCHED SAMPLE, AND ESTIMATION
SUBSAMPLE
(percentages)

	All Males 1851	Matched Sample 1851	Matched Sample 1881	Estimation Subsample 1851
Age (mean)	24.36	18.09	47.91	16.02
Occupation				
Student	15.84	22.30	0	27.61
Ag. Laborer	11.22	11.96	9.49	18.76
Farmer	3.01	2.20	5.98	0.45
Laborer	2.91	2.30	4.38	2.81
Miner	2.58	2.11	2.97	3.05
Weaver	2.11	2.66	1.41	3.26
Tailor	1.68	1.59	1.66	1.51
Carpenter	1.61	1.68	2.59	1.80
Shoemaker	1.12	1.27	1.96	1.75
Relation				
Son	44.56	59.86	4.17	100
Head	33.80	25.24	85.03	0
Lodger	5.69	3.46	3.28	0
Servant	1.89	3.69	0.61	0
Visitor	1.86	1.32	0.49	0
Marital status				
Married	47.72	42.31	79.27	4.12
Unmarried	47.66	55.61	10.95	95.21
Widower	4.18	1.45	9.45	0.22
Region				
East	6.58	7.88	7.10	9.86
Lancashire-Cheshire	11.31	12.43	13.51	8.77
London	8.55	5.14	7.02	0
London Environs	12.70	13.71	13.50	13.99
Midlands	19.89	20.94	19.41	25.12
North	5.06	5.02	5.39	5.25
South	18.49	20.57	18.63	26.50
Wales	4.93	4.61	4.89	3.13
York	12.50	9.70	10.43	7.39
Lived in county of birth in 1851	71.56	88.41		95.12
<i>N</i>	187,117	28,474	28,474	3,774

Notes: The occupations listed are the most commonly reported among the sample of all males. Students were children either attending school or receiving formal instruction at home.

Sources: 1851 census 2 percent sample and new sample of matched individuals.

heads; they were less likely to be married; more likely to be students and weavers and less likely to be anything else; and more likely still to be living within their county of birth, having had less time to move away.

Because I eliminate duplicate matches, individuals with the most common names are less likely to be included in the sample. We might expect such individuals more frequently to come from the lower socio-

economic strata, and this might impart bias to the matched sample. To test for this possibility, I use all of the males in the 1851 census sample to estimate a probit regression in which the dependent variable is whether an individual's name occurs eight times or more in the 1851 sample and the explanatory variables are age, county of residence, and dummy variables indicating occupational status as laborer, farmer, or servant.¹⁷ Age and the occupational dummies are statistically significant, but none is practically significant; that is, the marginal effects are all small. The largest coefficient is for the servant dummy, which increases the probability of having a common name by only 1.6 percentage points, relative to a baseline probability of 25 percent. So the "common name" problem should not impart a troublesome degree of bias to the matched sample.¹⁸

For the purposes of this particular study, I extracted a smaller subsample from the 28,000 matched individuals in order to examine the rural-urban migration decision. First, in order to see what caused some to move to the city and others to remain behind, and what happened to movers versus stayers, it is necessary to consider only individuals who began the period in a rural area.¹⁹

Second, I include only those who were listed as sons in the 1851 census. This group represents the largest in the matched sample, at 60 percent. (25 percent were household heads; 15 percent were grandsons, visitors, and so on) More importantly, for the sons, not only can we ob-

¹⁷ I define "common" here so that 25 percent of the population has a common name; the eight-occurrence cutoff follows from this definition. Results are unchanged if the definition is changed to 10 percent of the population. Excluding the county variable also does not change the results.

¹⁸ This finding is consistent with other studies using matched data, which also fail to find evidence of a "common name" problem; see Ferrie, *Yankees*, pp. 20–31; and Steckel, "Census Matching." Also, the British censuses provide very specific birthplace information (parish) as opposed to the U.S. censuses (state), which allows for more common-name individuals to be uniquely matched and included in the sample.

¹⁹ There is no hard definition for what constitutes a "town" or a "city." The United Nations has recommended that all places with more than 20,000 inhabitants living close together be considered as "urban." The U.S. Census uses 2,500 inhabitants for its definition of urban. These two numbers are often considered to be meaningful cut-off points. In nineteenth-century Britain, cities of more than 20,000 inhabitants were experiencing the most rapid growth; I consider these "urban" for the purposes of this study. The populations of towns and cities in both 1851 and 1881 were calculated using the 1881 census to ensure that the effects of increasing population between the two years did not cause people to appear to urbanize simply by living in a growing town. A simple dichotomous division of all places into either rural or urban is not entirely realistic but greatly facilitates the development of the structural econometric model. The results are largely robust to the use of an alternate definition of urban as places as those with more than 10,000 inhabitants, rather than 20,000. The only noteworthy change is that the coefficient measuring the effect of anticipated socioeconomic gain ($y_1^* - y_0^*$) on migration probability loses its economic and statistical significance. Other important results are unchanged.

serve their own occupation, but also that of their father, allowing us to gauge the impact of rural-urban migration on both intra- and intergenerational socioeconomic mobility. Furthermore, the location choice of sons in 1851 is likely to be exogenously chosen by the head of household; including household heads, who would already have made a rational 1851 location choice, would introduce endogeneity into the problem.²⁰

Finally, in order to judge the effects of moving, or not moving, to the city, the sample must also be limited to those individuals for whom there is substantive economic information from both censuses. Unlike their U.S. counterparts, Victorian censuses provide no quantitative economic variables such as personal or real estate wealth. However, they do include detailed occupational information. The census distinguishes employer from employee and master from apprentice, and it records the number of persons employed. There were thousands of occupations listed in the enumerators' books. In the absence of wage information, some system of ranking jobs according to their desirability is necessary. Scholars generally use the ranking scheme proposed by W. A. Armstrong for this purpose. He argues that only the Registrar General's social classification schemes of 1921 and 1951 satisfy the dual requirements of being appropriate for the nineteenth-century census data and being available in published lists for categorizing the vast array of occupations listed in the census. Armstrong's scheme consists of five ranked classes of occupation: 1–Professional, 2–Intermediate, 3–Skilled, 4–Semiskilled, and 5–Unskilled.²¹

Two sources, both published by the General Register Office of Great Britain, are needed to classify each occupation fully. The *Classification of Occupations*, 1921, lists approximately 16,000 different occupations and gives a three-digit code number for each. The *Decennial Supplement* for 1921 then provides the appropriate class rank for every code number.²² Armstrong provides several modifications to this system to

²⁰ Only sons aged 9 (the minimum legal working age) to 29 in 1851 were included. The median age of individuals in the sample is 15 years; 73 percent are between the ages of 9 and 19. The inclusion of pre- and young adolescent sons is desirable to augment sample size and to allow a more thorough analysis of intergenerational mobility; however, it does limit the extent to which results on intragenerational mobility (from first job in 1851 to later job in 1881) truly measure the extent of career mobility. We may doubt the extent to which the job of a 12 year old, for example, truly represents that individual's starting point in the labor market. For this reason, the results using *intergenerational* mobility should be seen as primary. However, the results are largely robust to the exclusion of the younger sons (aged 9–13) from the sample. See Appendix 2.

²¹ See Armstrong, "Use," for a full description. Some common occupations by class are Class 1–solicitor, accountant; Class 2–farmer, carpenter (employer); Class 3–carpenter (not employer), butcher (not employer); Class 4–agricultural laborer, wool comber; Class 5–general laborer, porter.

²² I gratefully acknowledge the assistance of Humphrey Southall, whose electronic dictionary was very useful for coding many of the occupations.

bring it into somewhat better harmony with the nature of nineteenth-century occupational and class structure. Most important is his use of the employee information given under the occupational field in the census. Regardless of job title, all employers of 25 or more are included in Class 1, and all people with Class 3 or 4 occupations employing at least one person other than a family member are included in Class 2. This system of classification is normative: higher-class jobs were better than lower class.²³ I have modified Armstrong's basic scheme in one way, in order to take full advantage of the information offered by the census. Because I observe entire households in both 1851 and 1881, I can calculate the ratio of servants to household members for each household. The job class ranking of some individuals was upgraded (never downgraded) according to the following scheme, proposed by Stephen Royle: all heads whose households contained at least one servant per household member were placed in Class 1, all others with one servant per three household members in Class 2, and any others that employed at least one servant in Class 3.²⁴ Formulated in this way, the job class variable represents an index of socioeconomic status, and I use the terms occupational class and socioeconomic status interchangeably throughout the article. Of course, collapsing thousands of occupations down into only five categories does result in substantial loss of information. However, this scheme is both tractable and, more importantly, standardized; Armstrong's is the standard method used by scholars to classify nineteenth-century British occupations, and it offers the attractive feature of allowing direct comparisons of socioeconomic information across studies.

Out of the 28,474 matched individuals, 3,774 are sons living with their father in a rural area in 1851, with both son and father reporting occupational information. They are shown in the third column of Table 1 for comparative purposes, though of course this subsample is unrepresentative of the male population of England and Wales by construction. I use this group to analyze rural-urban migration.

Two important caveats concerning the use of these matched individuals to measure the economic return to rural-urban migration must be noted. First, an important feature of the matching process is that only survivors over the 30-year period are included in the sample. In nineteenth-century Britain, urban areas exhibited higher mortality rates than did rural areas. If this urban mortality bias was greater for the poor, es-

²³ Armstrong demonstrates that job class, defined according to this system, is positively correlated with the employment of servants and negatively correlated with the incidence of shared accommodation. See "Use," p. 212.

²⁴ For an explanation and defense of the modification, see Royle, "Social Stratification." This modification affects only 5 percent of the matched sample.

estimates of the economic return to migration will be biased upward. We do not fully understand the relationship between socioeconomic status, urban residence, and mortality in nineteenth-century Britain. There is, however, reason to doubt that the urban mortality bias was strongly related to socioeconomic status. The basic nature of disease transmission was poorly understood, and the overall quality of health care was low; both factors limited the ability of the wealthy to insulate themselves from disease. Mortality and socioeconomic status were related, but not unambiguously. Life expectancy at age 20 was highest for Class 1 individuals, at 42.6 years; for Class 2 individuals it was 37.7 years; for Class 3, 40.7 years; and for Class 4, 38.8 years.²⁵ One study that looks at the U.S. from 1850 to 1860 finds no significant relationship between real estate wealth and mortality, whereas another finds a strong negative relationship between family personal wealth and mortality, but only in rural America.²⁶ Finally, the size of the rural-urban mortality gap in Britain began to close in the last quarter of the nineteenth century.²⁷

The second caveat concerns return migrants. The matched data measure only permanent moves; rural-urban migrants who return to their rural origin before 1881 are observed to be rural persisters rather than urban migrants. It is entirely plausible that poor urban labor market outcomes motivated many of these return moves, which again would act to raise the estimate of the return to urban migration. It is impossible to fully address this shortcoming with the data at hand; indeed, the issue of return migration is present in virtually all empirical migration studies. Greater frequency of observation would permit a refinement of the analysis, but missing some moves due to return migration is inevitable. At issue is the question being asked. This study will estimate the economic effect of permanent relocation from rural to urban areas for those individuals who in fact decided to move. It will not ask what the effect would have been for a randomly assigned individual, nor for an individual who moved to the city but subsequently decided to return to a rural area.

A MODEL OF RURAL-URBAN MIGRATION

George Borjas and others have demonstrated the importance of considering the endogeneity of migration decisions.²⁸ Moving from a rural to an urban area in nineteenth-century Britain was not a randomly as-

²⁵ Woods, *Demography*, p. 234.

²⁶ The studies are, respectively, Steckel, "Health"; and Ferrie, "Poor."

²⁷ Wrigley and Schofield, *Population History*, p. 476.

²⁸ Borjas, "Self-Selection." See also Sjaastad, "Costs"; and more recently, Ferrie, *Yankees*.

signed treatment; an individual's expectation of his labor market outcome surely influenced his decision of whether or not to move. To account for this self-selection, I use a model of regime switching with endogenous switching to analyze the migration decision. The model is

$$y_{1i} = \beta_1' X_{1i} + \varepsilon_{1i} \quad \text{if } M_i = 1 \quad (1)$$

$$y_{0i} = \beta_0' X_{0i} + \varepsilon_{0i} \quad \text{if } M_i = 0 \quad (2)$$

$$M_i = \begin{cases} 1 & \text{if } \gamma_1' Z_i + \gamma_2(y_{1i} - y_{0i}) + u_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where y represents socioeconomic status and M the urban migration decision.²⁹ X and Z are the factors that influence an individual's labor market outcome and migration decision, respectively, β and γ are vectors of coefficients, and ε and u are unobservable factors.³⁰ The left hand side of the inequality in equation 3 represents the net benefit of migration. If it is positive, individual i will move ($M_i = 1$). A wide range of empirical questions has been examined with this model, typically in the form of switching regressions, where y is a continuous variable, often wage, and equations 1 and 2 are estimated by OLS.³¹ The switching regression model does not apply to the present setting, because here the labor market outcome variable—socioeconomic status—is discrete, taking one of five ranked values. The ordinal nature of the socioeconomic status variable suggests the use of ordered probit analysis. In place of the standard model, in Appendix 1 I develop a model of switching ordered probits which I use to estimate equations 1–3. The estimates of β_0 , β_1 , γ_1 , and γ_2 reveal the determinants of job class attainment and urban migration. To analyze the selection of urban migrants and rural persisters and the treatment effect of migration, it is necessary to define several more parameters to be estimated. The selection of urban migrants (s_1) and the selection of rural persisters (s_0), conditional on observables, are

$$\begin{aligned} s_1 &= E(y_1^* | M = 1) - E(y_1^* | M = 0) = \tilde{X}_1 \hat{\beta}_1 - \tilde{X}_0 \hat{\beta}_1 \\ s_0 &= E(y_0^* | M = 0) - E(y_0^* | M = 1) = \tilde{X}_0 \hat{\beta}_0 - \tilde{X}_1 \hat{\beta}_0 \end{aligned} \quad (4)$$

²⁹ For an overview of the identification and parametric and semiparametric estimation issues associated with this class of model, see Heckman, "Varieties."

³⁰ For a list of the variables included in the vectors X and Z , see Table 2. I discuss these variables in the following section.

³¹ In addition to the migration studies cited above, see Maddala, *Limited-Dependent and Qualitative Variables*, and Amemiya, *Advanced Econometrics*, for numerous examples.

where for $r = 0, 1$, $\tilde{X}_r = E(X|M = r)$ and y_r^* is a continuous variable representation of job quality derived in Appendix 1. \tilde{X}_1 and \tilde{X}_0 represent the endowments of the skills and attributes of the average urban migrant and average rural persister, respectively. They are simply vectors of the conditional means of the explanatory variables for each category. The selection parameters compare the outcomes predicted by the econometric model for the average migrant and the average persister. If $s_1 > 0$ then urban migrants were positively selected: they achieved higher job quality in the urban labor market than the rural persisters would have had they chosen to move to the city. The analogous characterization holds for s_0 and the selection of rural persisters.

I consider two different measures of the treatment effect of rural-urban migration. The standard measure is the effect of treatment on the treated—the economic effect of migration on the average migrant. It is

$$\tau_1 = E(y_1^* - y_0^* | M = 1) = \tilde{X}_1 \hat{\beta}_1 - \tilde{X}_1 \hat{\beta}_0 \quad (5)$$

Like the selection parameters, τ_1 is calculated conditional on the observables of the model and ignores any unobserved heterogeneity between individuals. If the treatment effect is positive, then migration to the city yielded a positive return in terms of socioeconomic status for the average migrant. This is an indicator of labor market efficiency—individuals were moving in response to market signals. The second treatment effect parameter is the effect of treatment on the untreated (those who remained in rural areas)

$$\tau_0 = E(y_1^* - y_0^* | M = 0) = \tilde{X}_0 \hat{\beta}_1 - \tilde{X}_0 \hat{\beta}_0 \quad (6)$$

This is a hypothetical measure of the economic effect that urban migration would have had on the average rural resident who in fact chose *not* to move to an urban area. If τ_0 is positive, then labor markets were not operating at full efficiency—migration was not sufficient to exploit all potential gains, in that some who could have benefited from moving failed to do so.³²

³² This overstates the case somewhat, in that it ignores moving costs, higher urban cost of living, and possible compensating differentials required to induce movement into urban areas, with their higher rates of mortality and morbidity; τ_0 would have to be large as well as positive to indicate inefficiency. I discuss this issue more fully in the “Empirical Results” section.

ESTIMATION OF THE MODEL

The determinants of maximum attainable job quality in 1881—the elements of X in equations 1 and 2—include class rank in 1851, father’s class rank, age, age squared, age discrepancy, whether the individual lived in a town in 1851, whether the individual was the eldest son still residing in the household, whether the father was a farmer or employer, an interaction between the previous two terms, the industrial classification of the individual’s occupation in 1851, the degree of age-heaping for the county of residence in 1851, and region in 1851. Table 2 shows summary statistics for these variables, and for those in the migration equation. Several of the variables require explanation. I include the individual’s class rank in 1851 to measure the effect of moving to the city on the *change* in an individual’s job class rank; it controls for individual-specific effects on the level of job quality. The job class of the father is also included. Estimating the effect of the individual’s class in 1851 on his class in 1881 allows us to measure the degree of occupational mobility; including father’s class in 1851 reveals the degree of intergenerational mobility. The model, then, estimates socioeconomic mobility, both intra- and intergenerational, and how mobility differed for migrants and nonmigrants.

One class included in 1851 is not a proper occupational classification at all: that of “student.” It is the second most common class in 1851, and the single most commonly reported occupation, in this sample as well as in the census as a whole. It does not fit into the Registrar General’s occupational classification system, but it may be regarded as a class of its own. The 1851 census instructed parents to report their children as students (“scholars” in the language of the Victorian censuses) if they were older than five years and were “daily attending school, or receiving regular tuition under a master or governess at home.” Though we cannot know substantive details of their instruction, children older than nine who were listed as students were receiving an education. Many of their peers were not; they were working.³³ By including the students as a separate class, we can gauge in a very rough fashion the effect of early schooling on the future attainment of job quality as compared to employment at a young age.

Literacy and specific educational information are not observed, but it is possible to construct a proxy. I define “age discrepancy” as $|\text{Reported Age}_{1881} - (\text{Reported Age}_{1851} + 30)|$. As discussed previously, the

³³ See Tuttle, “Role,” for a discussion of the high incidence of gainful employment for young children in Victorian Britain.

TABLE 2
SUMMARY STATISTICS, ESTIMATION SAMPLE

	Element of Vector	Mean, 1851			Mean, 1881		
		All	Movers	Stayers	All	Movers	Stayers
Individual Characteristics:							
Job class:							
1 - Professional	<i>X</i>	0.004	0.01	0.003	0.02	0.03	0.02
2 - Intermediate	<i>X</i>	0.02	0.02	0.02	0.15	0.14	0.16
3 - Skilled	<i>X</i>	0.30	0.32	0.30	0.45	0.56	0.42
4 - Partly skilled	<i>X</i>	0.28	0.20	0.30	0.25	0.12	0.29
5 - Unskilled	<i>X</i>	0.12	0.10	0.12	0.12	0.15	0.11
Student	<i>X</i>	0.28	0.36	0.26	—	—	—
1851 class to 1881 class:							
Moved up					0.32	0.32	0.32
Moved down					0.18	0.19	0.18
Father's class:							
1 - Professional	<i>X</i>	0.02	0.03	0.01			
2 - Intermediate	<i>X</i>	0.16	0.17	0.16			
3 - Skilled	<i>X</i>	0.38	0.45	0.36			
4 - Partly skilled	<i>X</i>	0.33	0.27	0.35			
5 - Unskilled	<i>X</i>	0.10	0.08	0.11			
Father's class 1851 to 1881							
Moved up					0.29	0.30	0.28
Moved down					0.25	0.30	0.24
Age discrepancy (years):							
0	<i>X,Z</i>				0.49	0.49	0.49
1	<i>X,Z</i>				0.29	0.28	0.29
2	<i>X,Z</i>				0.10	0.08	0.11
3	<i>X,Z</i>				0.05	0.07	0.05
4	<i>X,Z</i>				0.03	0.04	0.03
5	<i>X,Z</i>				0.03	0.04	0.03
Age (years)	<i>X,Z</i>	16.02	15.26	16.26			
Eldest	<i>X</i>	0.57	0.57	0.58			
Inheritance	<i>X</i>	0.12	0.11	0.12			
Married	<i>X,Z</i>	0.02	0.02	0.03			
Not in town of birth	<i>Z</i>	0.26	0.32	0.24			
Industry:							
Agriculture	<i>X,Z</i>	0.24	0.15	0.27	0.24	0.04	0.30
Building	<i>X,Z</i>	0.03	0.03	0.03	0.08	0.10	0.07
Distributive	<i>X,Z</i>	0.02	0.02	0.02	0.09	0.13	0.08
Mining	<i>X,Z</i>	0.04	0.02	0.04	0.05	0.03	0.05
Textiles	<i>X,Z</i>	0.08	0.12	0.06	0.04	0.06	0.04
Iron and Steel	<i>X,Z</i>	0.02	0.03	0.02	0.04	0.06	0.03
Other Manufacturing	<i>X,Z</i>	0.15	0.15	0.15	0.20	0.23	0.19
Other	<i>X,Z</i>	0.42	0.48	0.40	0.26	0.36	0.23
Location Characteristics:							
Living in city		0	0	0	0.24	1	0
Living in town	<i>X,Z</i>	0.51	0.63	0.47	0.40	0	0.53
Region:							
Lancashire	<i>X,Z</i>	0.06	0.11	0.04	0.08	0.20	0.04
London	<i>X,Z</i>	0	0.00	0.00	0.04	0.17	0.00
London environs	<i>X,Z</i>	0.14	0.15	0.14	0.14	0.14	0.14

TABLE 2 — continued

	Element of Vector	Mean, 1851			Mean, 1881		
		All	Movers	Stayers	All	Movers	Stayers
Location Characteristic — continued							
Region — continued							
Wales	<i>X,Z</i>	0.03	0.02	0.04	0.03	0.02	0.04
Yorkshire	<i>X,Z</i>	0.07	0.08	0.07	0.08	0.09	0.07
Others	<i>X,Z</i>	0.70	0.64	0.72	0.63	0.39	0.71
Age-heaping (index)	<i>X,Z</i>	61.00	61.58	60.82			
Distance to city (km)	<i>Z</i>	23.35	19.60	24.51			
Wage gap (sh/wk)	<i>Z</i>	4.79	4.98	4.73			
Previous migrants	<i>Z</i>	13,716	15,339	13,211			
Unemployment (percent)	<i>Z</i>	7.97	8.84	7.70			
Nearby cities	<i>Z</i>	32.56	36.77	31.25			
Pct. in agriculture (percent)	<i>Z</i>	29.29	0.26	0.30			
Pct. in manuf. (percent)	<i>Z</i>	23.00	0.26	0.23			
<i>N</i>		3,774	896	2,878	3,774	896	2,878

Notes: Variable definitions and sources are given in the text. “All” refers to all individuals in the estimation subsample. “Movers” refers to those who moved to an urban place between 1851 and 1881; “Stayers” to those who remained in a rural place, though not necessarily the same rural place. Vector *X* contains the determinants of socioeconomic mobility; *Z* contains the determinants of migration.

age matching criterion is that an individual’s reported age in 1881 cannot be more than five years different from what it should be: the reported age in 1851 plus 30.³⁴ Approximately half of all individuals in the sample reported consistent values for age in the two censuses. Another quarter was off only by a year, and the remaining quarter was off by two to five years. This was a time before systematic record keeping, and many people had only an approximate idea of their age.³⁵ The observed age discrepancy for each individual gives an indication of that individual’s familiarity with arithmetic (“numeracy”) and the precision of his concept of time. In general, it may be regarded as a proxy for literacy, education, intelligence, and the like, and therefore it should have a positive impact on job quality attainment.³⁶ For the population as a

³⁴ In most cases for this sample, the parents would have reported the age in 1851. It could be possible, therefore, for an individual to report his correct age and still have a nonzero value of age discrepancy, if his parents had incorrectly reported his age in 1851. This seems unlikely, though, as most people would have learned their age from their parents.

³⁵ On age-enumeration in the nineteenth-century censuses, see Higgs, *Clearer Sense*.

³⁶ There is, in fact, a correlation between age discrepancy and being a student in 1851: students were 8 percent more likely to know their exact age than were nonstudents. Under-reporting was more common at every level than over-reporting, by about 23 percent overall. There may be reason to think that those who under-report their age are less likely truly to be ignorant of their exact date of birth or of arithmetic than those who over-report; it may be simple vanity at work. To check this, the model was re-estimated with age discrepancy defined as the actual discrepancy, rather than the absolute value. I found no statistical difference between positive and negative discrepancies.

whole, widespread ignorance of precise age manifests itself as age heaping—the tendency of any population to over-report rounded ages. The age heaping variable is a county-level index of this tendency and is intended to capture otherwise unmeasured community effects that might influence the labor market performance of individuals.³⁷

The remaining variables are more straightforward. People living in a town in 1851, as opposed to those in the truly rural countryside, may have possessed a form of locational human capital that would ease their transition into the big-city labor market. I estimate this effect by including a dummy variable that equals one if the individual lived in a town of between 2,500 and 19,999 inhabitants. I include dummy variables indicating whether a person in the sample was the eldest son currently residing in the household and whether his father was a farmer or an employer to capture an inheritance effect.³⁸ Fathers who owned land were likely to leave it to their eldest son, as primogeniture was the norm with regard to property inheritance in England. Likewise, sons of fathers who owned a business were likely heirs to an occupational inheritance, though in this case the inheritor could have been any son. Finally, I include industrial and regional dummy variables, the former classified according to the well-known scheme developed by Charles Booth.³⁹

In addition to expectation of future job prospects in both regimes, the factors that influence the migration decision—the components of Z in equation 3—include both individual- and location-specific elements. The individual-specific characteristics include age and age squared, marital status, age discrepancy, whether the person lived in a town in 1851, whether the person was living in the parish of his birth in 1851, and industrial classification. The human-capital interpretation of migration (that people migrate in order to maximize lifetime net benefit from moving) would suggest that the likelihood of migrating will decrease with age, as individuals have a shorter time span over which to reap the gains from moving and as they make more location-specific investments in their place of residence. More educated, knowledgeable people would have a lower tendency to report their age inconsistently; these same people should be better able to gather information on potential urban moves. Moving to a big city might have been a less drastic change, and so have carried a lower psychic cost, for those already living in towns in 1851. There are only two

³⁷ For a more thorough discussion of the economic usefulness of age heaping information, with emphasis on migration analysis, see Mokyr, *Ireland*.

³⁸ Note that the eldest son currently residing in the household may not in fact be the eldest son. If the eldest son has moved out of the household and is living on his own, he will not be identified in the data set.

³⁹ Booth, *Life*.

new variables in this group. The first is marital status, the expected effect of which is not clear *a priori*. Although studies of overseas migration generally find the typical migrant to be a young, single male, this may not be the case for shorter-distance, often local moves from countryside to city.⁴⁰ Indeed, during the second half of the nineteenth century, the decline in rural employment affected the job prospects of women more than those of men, so women became more likely than men to leave the countryside in order to find work in the city, most often as domestic servants.⁴¹ The second new variable is an indicator of whether an individual was living in his parish of birth in 1851; such individuals might be expected to have closer ties to their community and be less likely to move.

The elements of Z that depend on the individual's location in 1851 include the distance to the nearest city, the number of large cities within 100 km, a proxy for the total number of previous migrants in each nearby city, the average rural-urban wage gap, the urban unemployment rate, an interaction between the unemployment rate and distance to the nearest city, the percentage of the male workforce engaged in agriculture and in manufacturing in the home county in 1851, the degree of age heaping of the home county in 1851, and the regional dummies.⁴² Of these variables only age heaping and the regional dummies are in X .⁴³ Distance should be negatively correlated with tendency to migrate, as it would be both more costly to move to a distant city and more difficult to gather pertinent information. Knowing people who had already moved to a particular city was another important source of information for the potential migrant. I constructed a proxy for this friends-and-family effect. First, I tabulated the county of birth of all urban residents in England and Wales in 1881 by city using the 1881 census data. Second, I defined the set of most likely destination cities for each individual in the sample to be the nine closest cities to his parish of residence in 1851 plus London. Finally, for each individual i with likely destination cities $j = 1, \dots, 10$, the proxy for the stock of previous migrants is

$$\text{previous migrants}_i = \sum_{j=1}^{10} \left(\frac{d_{ij}^{-1}}{\sum_j d_{ij}^{-1}} \right) m_{ij} \quad (7)$$

⁴⁰ For results on overseas migration, see Hatton and Williamson, "Mass Migrations."

⁴¹ Baines, "Population."

⁴² Share in agriculture and manufacturing are from 1871 census, in Lee, *British Regional Employment Statistics*.

⁴³ I calculated distance to the nearest city and number of large cities nearby using modern grid point references for place names from the 1991 British Census. I am grateful to Justin Hayes with MIMAS at the University of Manchester for providing me with these data.

where m_{ij} is the number of people living in city j in 1881 born in the county of residence of individual i , and d_{ij} is the distance between city j and the place of residence of individual i . This is a proximity-weighted average of the m values of all likely destination cities. I calculated the average urban-rural wage gap and the urban unemployment rate along similar lines. The wage gap is the average wage of laborers in the building trade in nearby cities minus the average wage of agricultural laborers in the origin county.⁴⁴ These wages would not have applied specifically to the majority of workers making migration decisions; rather, they serve as a rough proxy for the general magnitude of the wage gap that existed between various rural areas and nearby cities. I calculated the urban unemployment rate with data on joblessness among members of the Amalgamated Society of Engineers, available for 56 cities in England and Wales from 1858 through 1909.⁴⁵ It is the proximity-weighted average—as in equation 7—of the rates from 1862, 1868, and 1879 for the two nearest cities and London. This is the urban unemployment rate, so its expected effect on tendency to migrate should be negative; people would be less likely to move to the city if there were high unemployment there. However, we do not observe the relevant local unemployment rate and so cannot control for it. Therefore, because the urban unemployment rate is taken from the cities nearest to each individual's place of origin, it proxies, to an extent, for the local rural unemployment rate facing each individual. Insofar as this is true, the effect of unemployment would be positive.⁴⁶ To test for this effect, an interaction term between the unemployment rate and distance to nearest city is included. This term is anticipated to have a negative coefficient: for any given unemployment rate, an increase in the value of the interaction term indicates that the unemployment pertains to a city which is farther away, thus more truly the unemployment of the destination city and less that of the place of origin.

The three parameter vectors β_1 , β_2 , and γ are identified under the assumptions of the model; however, if the variables in X and Z are identi-

⁴⁴ Agricultural laborers' wages are from Hunt, "Industrialization," table 6. They cover the years 1867–1870. Building laborers' wages are from Hunt, *Regional Wage Variation*, table 1.5, for the year 1886.

⁴⁵ These data are from Southall, "Regional Unemployment," table 3. For a discussion of its appropriateness as a proxy for the unemployment experiences of other groups, see Southall and Gilbert, "Good Time to Wed."

⁴⁶ Unfortunately, the occupational information from the 1851 census is not well suited to determining individual unemployment within the sample. Specific instructions regarding the unemployed were not given to enumerators until 1861; many people out of work in 1851 reported the fact, but many simply reported the job they most recently held.

cal, this holds only if the structure and normality assumptions of the model are exactly correct. Fortunately, in this case there are some reasonable exclusion restrictions. Z includes quite a few variables not present in X : marital status, whether the person was living in the parish of his birth in 1851, and all of the location-specific variables save age heaping and the regional dummies. X does not include quite so many unique variables: the eldest and inheritance dummies, father's job class, and own job class in 1851. These, of course, do influence the individual's migration decision, but only through their effect on $(y_{1i}^* - y_{0i}^*)$. With these exclusion restrictions, the parameters of the model are identified, even if the assumptions of the model do not hold exactly.

EMPIRICAL RESULTS

Occupational Class Attainment

Estimating the two job class equations, 1 and 2, yields some general insight on the nature of socioeconomic mobility in nineteenth-century Britain, and on how it differed between the rural and urban areas. Results from estimating equations 1 and 2 and as well as from estimating what may be considered the baseline model—an ordered probit estimation of job class in 1881 on X plus a dummy variable for moving to a city by 1881—are presented in Table 3. The columns labeled “Movers” show the estimates of β_1 , the coefficients for those who chose to move to (and remain in) an urban area by 1881, and the columns labeled “Stayers” show the estimates of β_0 , the coefficients for those who remained in a rural place. The ancillary parameters ρ and k do not vary between the two groups. Interpretation of ordered probit results is not straightforward. The coefficient estimates represent the effect of a change in the explanatory variable on the unobserved latent variable y^* , job quality in this case. Job quality is a theoretical construct of the model, with no units or simple interpretation. The coefficient estimates reveal the sign of the effect of the explanatory variable and the precision with which the effect is estimated. For an understanding of magnitude, the marginal effect of each dependent variable must be calculated. These are shown in Table 4, with the “Movers” and “Stayers” columns as in Table 3. Also shown are the baseline probabilities of attaining each job class conditional on moving to a city or remaining in a rural place. Each explanatory variable has a different marginal effect for each possible outcome of the ordered probit, so there is a different effect for each of the five job classes that an individual could attain in 1881. The

TABLE 3
 MAXIMUM LIKELIHOOD RESULTS: DETERMINANTS OF JOB CLASS ATTAINMENT

Variable	Class, 1881	Class, 1881		Variable	Class, 1881	Class, 1881	
		Movers	Stayers			Movers	Stayers
Equation	(1)	(2)	(3)	Equation	(1)	(2)	(3)
Constant	0.449 (0.319)	-0.131 (0.758)	0.699 (0.359)	Industry	(***)		(***)
Own class 1851	(***)	(***)	(***)	Agric.	-0.181 (0.106)	-0.199 (0.234)	-0.128 (0.122)
1,2 - Prof Inter.	1.371 (0.147)	1.958 (0.289)	1.156 (0.172)	Building	0.280 (0.128)	0.305 (0.273)	0.301 (0.144)
3 - Skilled	0.354 (0.089)	0.403 (0.193)	0.317 (0.101)	Distrib.	0.319 (0.148)	0.126 (0.308)	0.361 (0.168)
4 - Partly Skilled	0.136 (0.109)	0.311 (0.227)	0.054 (0.126)	Mining	0.065 (0.122)	0.167 (0.309)	0.083 (0.136)
Student	0.515 (0.074)	0.785 (0.162)	0.384 (0.086)	Textiles	-0.114 (0.102)	0.033 (0.211)	-0.200 (0.119)
Father's class	(***)	(***)	(***)	Iron & Steel	-0.022 (0.144)	-0.049 (0.289)	-0.016 (0.166)
1 - Prof.	1.427 (0.159)	0.930 (0.283)	1.630 (0.203)	Other Manuf.	0.204 (0.089)	0.271 (0.189)	0.192 (0.101)
2 - Inter.	0.463 (0.101)	0.444 (0.198)	0.366 (0.121)	Region Lancash.	(*) 0.041 (0.089)	(***) -0.015 (0.155)	-0.037 (0.120)
3 - Skilled	0.342 (0.064)	0.213 (0.142)	0.348 (0.074)	London envir.	-0.075 (0.053)	-0.353 (0.109)	0.001 (0.061)
4 - Partly Skilled	-0.134 (0.065)	-0.238 (0.148)	-0.117 (0.072)	Wales	0.166 (0.116)	0.113 (0.313)	0.219 (0.126)
Age dis (yrs)	(***)		(***)	Yorkshire	0.126 (0.075)	0.282 (0.154)	0.045 (0.086)
0	0.201 (0.046)	0.133 (0.095)	0.224 (0.053)	Moved to city	-0.005 (0.043)		
1	-0.008 (0.050)	-0.028 (0.103)	0.004 (0.057)	Column	(1)	(2) + (3)	
Age	0.052 (0.027)	0.167 (0.059)	0.009 (0.031)	r_1			-0.053 (0.100)
Age ² / 100	-0.152 (0.072)	-0.505 (0.162)	-0.023 (0.081)	r_2			-0.324 (0.088)
Eldest	0.030 (0.039)	0.111 (0.081)	0.002 (0.045)	k_2	0.959 (0.028)		0.948 (0.031)
Inheritance	0.335 (0.110)	-0.102 (0.213)	0.552 (0.130)	k_3	2.429 (0.038)		2.408 (0.050)
Eldest × Inheritnc.	-0.039 (0.108)	-0.252 (0.237)	-0.012 (0.121)	k_4	3.665 (0.060)		3.642 (0.076)
Age heaping	-0.003 (0.003)	-0.005 (0.007)	-0.003 (0.004)	Pseudo R^2	0.092		0.090
Living in town	0.051 (0.038)	-0.084 (0.083)	0.050 (0.045)	Loglikelihood.	-4,572.715		-6,457.798
				LR c^2	925.530		1,284.230
				Prob > c^2	0.000		0.000

* = Jointly significant at the 10-percent level.
 *** = Jointly significant at the 1-percent level

** = Jointly significant at the 5-percent level.

TABLE 3 — continued

Notes: $N = 3,774$. The two columns of equation 1 represent all 3,774, those for equation 2 represent the 896 who moved to a city, and those for equation 3 the 2,878 who remained in a rural place. Standard errors are in parentheses. All explanatory variables use 1851 values. Pseudo $R^2 = 1 - L_1/L_0$, where L_1 is the log likelihood of the model and L_0 is the log likelihood of the “constant only” model. The LR χ^2 statistic represents a test of all parameters of the model being equal to zero. Omitted dummies are *Class*, *Father's class*: 5, *Age discrepancy*: 2–5, *Industry*: Others, *Region*: Others. k_1 is normalized to zero.

sign of the overall effect of each explanatory variable is unambiguous, but the sign will vary across 1881 job classes. A variable that positively influences job quality will positively influence the probability of being in the highest job class but will negatively influence the probability of being in the lowest job class.

As expected, the 1851 class variables, both own and father's, are strong influences on subsequent job quality attainment. In each case save one, the effects are as anticipated: being in a higher class in 1851 and being the son of a father in a higher class strongly predict attaining a higher quality job in 1881. For example, Table 4 indicates that migrants who began in Class 1 or 2 were 31 percentage points more likely to end up in Class 2 than were other migrants, and they were 36 percentage points more likely to end up in Class 1. The two baseline attainment probabilities of 15 and 1 percent, respectively, indicate the extreme difficulty of moving up into a Class 1 or 2 job from a lower initial class.

Perhaps most interesting with regard to the class variables is to compare their effect on those who moved to urban areas versus on those who remained in rural places. A chi-square test reveals no significant difference between the combined effect of both own and father's class on those who moved versus those who remained. However, as Table 3 shows, for urban migrants the effect of their own class in 1851 was stronger than the effect of their father's class. The reverse was true for those who remained behind, the job class of their father being a stronger influence on their 1881 class attainment than was their own class in 1851. In addition, the effect of own class was stronger for those who moved than those who did not, and the effect of father's class was weaker.⁴⁷ It appears that leaving the countryside and moving to the city offered migrants a better chance to escape the intergenerational career trajectory inherited from their father. It is also interesting to note that the benefit of being a student in 1851 was nearly twice as great for those who moved to a city as it was for those who remained rural. Students

⁴⁷ Only the first is statistically significant. The two χ^2 statistics differ from zero with 97 and 80 percent confidence, respectively.

TABLE 4
 MAXIMUM LIKELIHOOD RESULTS: DETERMINANTS OF JOB CLASS ATTAINMENT — (marginal effects of explanatory variables)

Job class	$\partial [P(y=1)] / \partial x$		$\partial [P(y=2)] / \partial x$		$\partial [P(y=3)] / \partial x$		$\partial [P(y=4)] / \partial x$		$\partial [P(y=5)] / \partial x$	
	Movers	Stayers	Movers	Stayers	Movers	Stayers	Movers	Stayers	Movers	Stayers
1,2 - Prof., Intermed.	0.3639	0.0921	0.3075	0.2739	-0.3423	-0.0452	-0.2437	-0.2151	-0.0855	-0.1057
3 - Skilled	0.0161	0.0081	0.0888	0.0596	0.0335	0.0515	-0.0847	-0.0636	-0.0538	-0.0555
4 - Partly Skilled	0.0120	0.0012	0.0682	0.0097	0.0273	0.0098	-0.0656	-0.0106	-0.0419	-0.0100
Student	0.0402	0.0103	0.1797	0.0734	0.0334	0.0591	-0.1594	-0.0774	-0.0939	-0.0654
Father's class										
1 - Professional	0.0811	0.2007	0.2297	0.3490	-0.0668	-0.1784	-0.1720	-0.2581	-0.0720	-0.1131
2 - Intermediate	0.0206	0.0108	0.1022	0.0726	0.0237	0.0510	-0.0930	-0.0751	-0.0535	-0.0593
3 - Skilled	0.0074	0.0084	0.0453	0.0640	0.0228	0.0591	-0.0450	-0.0690	-0.0306	-0.0626
4 - Partly Skilled	-0.0072	-0.0024	-0.0480	-0.0202	-0.0316	-0.0224	0.0494	0.0225	0.0374	0.0225
Age discrep. (years)										
0	0.0044	0.0049	0.0277	0.0396	0.0155	0.0413	-0.0279	-0.0436	-0.0197	-0.0421
1	-0.0009	0.0001	-0.0059	0.0007	-0.0034	0.0007	0.0060	-0.0008	0.0042	-0.0007
Age	0.0055	0.0002	0.0349	0.0016	0.0197	0.0017	-0.0352	-0.0018	-0.0248	-0.0017
Age squared / 100	-0.0166	-0.0005	-0.1055	-0.0040	-0.0595	-0.0042	0.1065	0.0044	0.0750	0.0043
Elders	0.0036	0.0000	0.0229	0.0004	0.0134	0.0004	-0.0233	-0.0004	-0.0166	-0.0004
Inheritance	-0.0031	0.0202	-0.0206	0.1167	-0.0136	0.0568	0.0213	-0.1139	0.0160	-0.0798
Elders*Inheritance	-0.0065	-0.0002	-0.0477	-0.0021	-0.0402	-0.0022	0.0509	0.0023	0.0435	0.0022
Age heaping	-0.0002	-0.0001	-0.0010	-0.0005	-0.0007	-0.0006	0.0011	0.0006	0.0008	0.0006
Living in town	-0.0028	0.0011	-0.0176	0.0088	-0.0099	0.0093	0.0177	-0.0098	0.0125	-0.0094
Industry										
Agriculture	-0.0059	-0.0025	-0.0398	-0.0218	-0.0275	-0.0251	0.0413	0.0244	0.0318	0.0251
Building	0.0138	0.0092	0.0704	0.0607	0.0173	0.0400	-0.0646	-0.0624	-0.0369	-0.0475
Distributive	0.0047	0.0119	0.0276	0.0745	0.0117	0.0431	-0.0268	-0.0750	-0.0172	-0.0544
Mining	0.0065	0.0020	0.0368	0.0153	0.0142	0.0143	-0.0354	-0.0166	-0.0222	-0.0149
Textiles	0.0011	-0.0035	0.0069	-0.0322	0.0037	-0.0425	-0.0069	0.0365	-0.0048	0.0417
Iron & Steel	-0.0015	-0.0003	-0.0100	-0.0028	-0.0062	-0.0030	0.0102	0.0031	0.0075	0.0031
Other manuf.	0.0110	0.0049	0.0605	0.0362	0.0209	0.0310	-0.0573	-0.0388	-0.0351	-0.0333
Pr[y X=E(X)]	0.0127	0.0078	0.1456	0.1105	0.5183	0.4905	0.2436	0.2807	0.0799	0.1104

TABLE 4 — continued

Notes: Dependent variable y is job class in 1881. Region variables not shown. Effects are calculated at mean of x for continuous variables; for discrete variables, effect is $P(y=c|x=1) - P(y=c|x=0)$, $c=1, \dots, 5$. Baseline probabilities calculated at mean of all variables.

who moved to the city were 18 percentage points more likely to end up with a Class 2 job than nonstudent urban migrants; for those who remained in a rural place, the advantage was only 7 percentage points.⁴⁸ As for other variables of interest, people who reported their age with consistency between the two censuses performed better in the labor market. The effect was particularly strong for nonmigrants, who were 8 percentage points more likely to end up with a Class 2 or 3 job and 4 percentage points less likely to find themselves in a Class 5 job than nonmigrants who inconsistently reported their age by more than one year. Having a potential inheritance also was a strongly positive force, but only for nonmigrants, as expected. The effect was particularly positive for the attainment of Class 2 jobs, where it increased the probability by 12 percentage points.

Self-Selection and the Determinants of Migration

Who were the urban migrants? Table 2 gives an overview of their characteristics. Of the 3,774 individuals in the estimation sample, 896 moved from a rural to an urban area between 1851 and 1881, for an overall urban migration rate of 24 percent. Out of the entire sample of 28,474 matched individuals (younger than the population as a whole because of the survival effect in the matching process), 18,740 (66 percent) were living in a nonurban area in 1851. Of these, 4,387 moved to a city by 1881, for a rate of 23 percent over the 30-year period. Notably, the migrants were not those at the bottom of the economic and social ladder, desperate for any sort of a change. There are relatively more people with Class 3 jobs in 1851 and sons of fathers with Class 3 jobs among the pool of urban migrants than among the nonmigrants. Just the reverse is true for Classes 4 and 5, which are represented more heavily among the rural persisters. Sons described as students in 1851 were also more prevalent within the group of urban migrants than the group who remained in rural places. It decidedly was not the case that the most destitute and poverty stricken, those occupying the society's lowest class, were pouring from the countryside into the cities. The pool of urban

⁴⁸ It is noteworthy that early school attendance appears to have conveyed substantial benefits on later adult labor market outcomes. The full extent of this effect and the nature of the schooling versus labor market participation selection mechanism is an interesting issue in its own right, which I explore in greater depth in ongoing work. See Long, "Economic Return."

migrants resembled the overall population, with the middle classes being somewhat overrepresented and the lowest classes being underrepresented.

Table 5 shows results from the probit estimation of the urban migration decision, which yields a more accurate and detailed picture of the characteristics of migrants. The large, positive coefficient on $(\hat{y}_{1i}^* - \hat{y}_{0i}^*)$ indicates that in fact people were moving to the cities in order to improve their socioeconomic status and that the pull of this factor was strong relative to other factors. Interpreting the magnitude of the effect is not straightforward, as y^* , job quality, is a construct of the model, not an observed variable. Every unit increase in anticipated job quality difference between the rural and urban regimes increased the odds of an individual moving to a city by 7 percentage points. The estimated cut points from Table 5 shed light on this figure. The range of y^* between different job classes varies from 0.95 for Class 4 to 1.46 for Class 3. So the prospect of being one class higher in the city than in the countryside (roughly an increase in y^* of 1 to 1.5) increases the odds of migration by about 7–10 percentage points—a large increase considering that the baseline predicted probability of moving is 22 percent. Understanding the other determinants of migration is simpler. As expected, people who were not living in their place of birth and those living in towns were more likely to move to a city by 1881, by 5 and 8 percentage points, respectively. Taken together, this suggests at least the possibility of a sort of intergenerational two-stage migration, with fathers moving from the truly rural areas to small towns, and their sons subsequently moving to the bigger cities. This conjecture is speculative but consistent with the data. Age did not yield the usual negative effect, almost certainly because the sample is already restricted to the young. Nor did consistent age reporting yield the expected positive informational effect. Those working in agriculture and mining in 1851 were considerably less likely to move to a city by 1881.

The location-level variables were also important factors in the migration decision. The effect of job quality did not entirely capture the economic incentive to move; the average wage gap was also a significant influence. For every shilling per week of expected wage difference between the city and the countryside, an individual was about 1 percentage point more likely to move. As anticipated, the effect of the unemployment rate was not clear-cut. The effect of urban unemployment was significantly positive, indicating that it may have been proxying for local unemployment. The test for this is to interact unemployment with distance to nearest city. The coefficient is negative, indicating that the farther away was the nearest city, the more the effect of high urban un-

TABLE 5
PROBIT RESULTS: DETERMINANTS OF URBAN MIGRATION

Variable (individual specific)	Move to City		Variable (location specific)	Move to City	
	Coefficient	Marginal Effect		Coefficient	Marginal Effect
Constant	-1.296 (0.522)		<i>Distance to city</i>	0.001 (0.004)	0.0003
$y_{1i}^* - y_{0i}^*$	0.239 (0.124)	0.071	<i>Wage gap</i>	0.035 (0.017)	0.011
<i>Age discrep. (years)</i>			<i>Previous migrants</i>	-4.43E-06 (-3.12E-06)	-1.32E-06
0	-0.055 (0.065)	-0.016	<i>Unemployment</i>	0.043 (0.013)	0.013
1	-0.044 (0.073)	-0.013	<i>Unemp*Distance</i>	-0.001 (0.000)	-0.0004
<i>Age</i>	0.020 (0.037)	0.006	<i>Nearby cities</i>	-0.006 (0.002)	-0.002
$Age^2 / 100$	-0.093 (0.106)	-0.028	<i>Age heaping</i>	0.005 (0.005)	0.002
<i>Married</i>	0.092 (0.176)	0.028	<i>Pct. in agriculture</i>	-0.685 (0.370)	-0.204
<i>Not in town of birth</i>	0.157 (0.054)	0.048	<i>Pct. in manuf.</i>	0.508 (0.553)	0.151
<i>Living in town</i>	0.281 (0.055)	0.083	<i>Region</i>	(*)	
<i>Industry</i>	(***)		Lancashire	0.108 (0.189)	0.033
Agriculture	-0.270 (0.087)	-0.076	London Envrns.	0.194 (0.095)	0.061
Building	0.025 (0.155)	0.007	Wales	-0.230 (0.184)	-0.063
Distributive	0.010 (0.209)	0.003	Yorkshire	0.156 (0.139)	0.049
Mining	-0.493 (0.170)	-0.12	Pseudo R ²	0.065	
Textiles	0.056 (0.107)	0.017	Log likelihood	-1,933.311	
Iron & Steel	0.083 (0.170)	0.025	LR χ^2	270.290	
Other Manuf.	-0.040 (0.082)	-0.012	Prob > χ^2	0.000	
			Pr[M=1 Z=E(Z)]	0.222	

* = Jointly significant at the 10-percent level. ** = Jointly significant at the 5-percent level.
*** = Jointly significant at the 1-percent level.

Notes: $N = 3,774$, of which 896 moved to a city by 1881. Standard errors are in parentheses, calculated by bootstrapping via data resampling with 1000 repetitions, which is necessary since $y_{1i}^* - y_{0i}^*$ is estimated rather than observed. All dependent variables are 1851 values. Omitted dummies are *Age discrepancy*: 2-5, *Industry*: Others, *Region*: Others. Marginal effects evaluated at the mean of z , except for discrete variables, where the marginal effect is $P(M=1|z=1) - P(M=1|z=0)$.

employment was indeed negative. Other factors, such as distance, nearby cities, and friends-and-family, exerted small or statistically insignificant effects. This result is somewhat surprising, considering that

previous studies have tended to find these variables important.⁴⁹ Most likely, the individual-level analysis used here is an imprecise way to estimate these location-level effects.

We have seen that the urban migrants did not begin the period at the bottom of the labor market. We turn now to the question of selection: whether the urban migrants and rural persisters were positively or negatively selected. Table 6 presents estimates for the four migration-effect parameters defined in equations 4–6. The first two summarize the selection process. Urban migrants were positively selected whereas rural persisters were negatively selected. Both results are statistically significant at the 1-percent significance level. Not only did the migrants perform better in the urban labor market than the persisters would have, they also would have outperformed the persisters in the rural labor market had they chosen not to migrate. In this sense, then, urban migrants were the “cream of the crop.” They were those whose labor market prospects were brightest.⁵⁰

Treatment Effect

One of the central questions of the present study concerns the effect of moving to a city on the ability of migrants to attain high quality jobs. The estimate of the treatment effect for both urban migrants and those who remained in rural areas, along with standard errors and confidence intervals, are shown in Table 6. The economic effect of moving to a city is positive for *both* groups, though it is larger for those who actually chose to migrate. Neither estimate is statistically significant according to the standard two-tailed test; however, the null hypothesis that $\tau_l \leq 0$ can be rejected at the 10-percent significance level. Moving to the city allowed the average mover to obtain a better job than he would have been able to get had he remained in a rural place, and it *would* have allowed the average rural persister to improve as well, had he chosen to move.

Magnitude is best interpreted by considering job class transitions rather than effects on the latent variable y^* . In addition, it is informative to examine exactly for whom the treatment effects were largest. This information is presented in Table 7. Here I define the treatment effect in terms of the effect of an urban move on intra- and intergenerational oc-

⁴⁹ See, for example, Boyer and Hatton, “Migration.”

⁵⁰ In this aspect they appear to be different from rural-urban migrants in the nineteenth-century United States. Ferrie, “Down on the Farm,” p. 11, reports that urban migrants in the United States between 1850 and 1860 were negatively, rather than positively, selected. They fared worse in the cities than rural persisters would have had they chosen to move.

TABLE 6
MEASURES OF THE MIGRATION EFFECT
(dependent variable is latent job quality in 1881, y^*)

		Estimate	Std Error	90% Confidence Interval	85% Confidence Interval
Selection of urban migrants	($s1$)	0.1337	0.0385	0.0791 <i>0.0864</i>	0.2082 <i>0.1996</i>
Selection of rural persisters	($s0$)	-0.0982	0.0307	-0.1500 <i>-0.1435</i>	-0.0500 <i>-0.0570</i>
Treatment effect, urban migrants	($\tau1$)	0.2087	0.1506	-0.0206 <i>0.0078</i>	0.4689 <i>0.4352</i>
Treatment effect, rural persisters	($\tau0$)	0.1733	0.1625	-0.0868 <i>-0.0626</i>	0.4344 <i>0.3956</i>

Notes: Standard errors and confidence intervals are calculated by bootstrapping via data resampling with 1,000 repetitions.

cupational mobility.⁵¹ The table gives the probabilities that an average migrant from each socioeconomic class would attain either a higher or a lower class if he were to move to an urban place and if he were to remain in a rural area.⁵² In the first five rows, moves are relative to the individual's father's class, and in the second five to the individual's own class in 1851. I define the net gain of moving to a city as

$$[P(\text{up} | \text{urban}) - P(\text{up} | \text{rural})] - [P(\text{down} | \text{urban}) - P(\text{down} | \text{rural})] \quad (8)$$

For example, the average mover whose father held a Class 4 job in 1851 had a 53 percent chance of moving up to a Class 1, 2, or 3 job if he moved to a city and a 42 percent chance if he chose to remain in a rural place. That same person would have a 15 percent chance of falling to a Class 5 job in the city and a 23 percent chance of falling in the countryside. So moving to the city confers an 11-percentage-point boost in the probability of making an upward move and offers an 8-percentage-point lower likelihood of falling to a lower class, for a net gain of 18 percentage points.

The first important thing to note from Table 7 is that individuals from across all socioeconomic strata realized the gains from urban migration. All individuals were more likely to improve their own socioeconomic status if they migrated to a city, and sons of fathers belonging to any class other than 1 were more likely to end up in a higher strata by 1881 than their fathers occupied in 1851 if they moved. The second feature to

⁵¹ Again, because of the relatively young average age of the men included in the sample, the results on intergenerational mobility should be seen as primary. In any case, the two sets of results are quite similar.

⁵² This is, then, simply another way to look at the effect of treatment on the treated.

TABLE 7
TREATMENT EFFECT FOR MIGRANTS, BY CLASS

	P(up urban)	P(up rural)	P(down urban)	P(down rural)	Net Gain
Father's class = 1	0	0	0.8376	0.7432	-0.0944
Father's class = 2	0.0377	0.0341	0.7067	0.7224	0.0193
Father's class = 3	0.2225	0.1602	0.2432	0.3206	0.1398
Father's class = 4	0.5326	0.4219	0.1516	0.2264	0.1855
Father's class = 5	0.8940	0.8140	0	0	0.0800
Class in 1851 = 1 or 2	0	0	0.1976	0.3874	0.1898
Class in 1851 = 3	0.2131	0.1641	0.2533	0.3148	0.1105
Class in 1851 = 4	0.5188	0.4282	0.1598	0.2215	0.1523
Class in 1851 = 5	0.8374	0.8260	0	0	0.0115
Class in 1851 = student	0.2510	0.1793	0.2153	0.2941	0.1505

Note: Individuals classed as students in 1851 were considered to have moved up if they obtained a Class 1 or 2 job by 1881, and they were considered to have moved down if they held Class 4 or 5 jobs in 1881.

note from the table is that the most prevalent groups in the sample also had the largest treatment effects. As Table 2 shows, about 70 percent of the sample had fathers with Class 3 or 4 jobs. These men realized net gains to urban migration, in terms of intergenerational occupational mobility, of 13.98 and 18.55 percentage points, respectively. Eighty-six percent of the sample began the period in Class 3 or 4 or as a student. The net gain of urban migration was large for all these individuals—15.05, 15.23, and 11.05 percentage points, respectively.⁵³

It is instructive to consider some common examples. The average rural son of a Class 4, partly skilled father (an agricultural laborer, for example) who subsequently decided to move to a city was 26 percent more likely to improve the quality of his occupation over that of his father than if he had remained in a rural area (53 versus 42 percent), and he was 35 percent less likely to fall into a Class 5, unskilled occupation (15 versus 23 percent). The average Class 4 worker in 1851 was 21 percent more likely to improve his socioeconomic status if he moved to an urban place than if he remained in the countryside, and he was 27 percent less likely to find himself with a Class 5 occupation. These are substantial gains. For people like this, moving to the city was an important avenue of socioeconomic improvement. Table 7 does not include comparable results for rural persisters. The results are similar, though

⁵³ Individuals with Class 4 jobs in 1851 and sons of Class 4 fathers realized strong net gains from moving to urban areas even though many of them made the downward move from agricultural laborer (Class 4) to general laborer (Class 5). But such individuals actually realized a 65 percent gain in annual earnings, on average (Tuttle, "Role," p. 173). If agricultural and general laborers were put into the same class, or if general laborers were ranked above agricultural, then the treatment effects would be even stronger.

slightly smaller in magnitude.⁵⁴ The important thing to note is that many people who chose to remain in rural areas, especially sons of Class 3 and 4 fathers, could have realized substantial labor market gains had they chosen to move.

CONCLUSIONS

The two strands of the literature on internal migration in nineteenth-century Britain focus on the extent and determinants of migration on the one hand, and the efficiency implications of mobility on the other. The results in this study speak to both strands. With respect to the nature of rural-urban migration, it is clear that urban migrants were the cream of the rural labor market crop, in that their prospects in both the urban and the rural labor markets were superior to those of the rural persisters. Migrants considered anticipated labor market outcomes and were more likely to move the larger were their anticipated gains. Typically, the decision to move was a fruitful one. On average, people from all socioeconomic strata who moved to the city were substantially more successful in improving their socioeconomic status than they would have been had they remained in rural areas, and they were more likely to experience upward intergenerational occupational mobility.

The implications of these results for the efficiency of British labor markets is largely, but not entirely, positive. First, potential urban migrants responded to labor market signals, as evidenced by the large, positive coefficient on $(\hat{y}_{1i}^* - \hat{y}_{0i}^*)$ in the migration-decision equation. In addition, migrants were well rewarded; the large, positive estimate of τ_1 indicates a strong economic return to urban migration for the average migrant. Both of these features are indicative of an efficiently functioning labor market drawing migrants from areas of low to areas of high marginal product. The only evidence that labor markets might have been functioning suboptimally is the positive estimate of τ_0 ; some rural persisters could have benefited from moving to urban areas but chose to remain rural anyway. This is not, of course, conclusive evidence of labor market failure. The present analysis does not account for higher living costs, psychic costs of moving, and any compensating differentials required to induce migration to cities with poor living standards. It is entirely possible that gains to migration went unexploited because they were insufficient to outweigh these unmeasured costs. As it stands,

⁵⁴ The intergenerational net gains for sons of Class 1–Class 5 fathers are -0.075 , -0.053 , 0.135 , 0.159 , and 0.074 , respectively. These are the net gains the average rural persister would have realized had he chosen to move to an urban area.

however, the presence of these unexploited labor market gains is at least suggestive of potential inefficiency. Even if this is the case, and the overall level of migration was less than optimal, the high *quality* of urban migrants must have served to offset the inefficiently low *quantity*. The results here indicate that urban migration was a positive selection process, whereas rural persistence was a negative selection process. The fact that urban migrants were the “cream of the crop” meant that urban, industrial labor markets were drawing the best of the rural labor pool, even if the magnitude of migration was less than optimal.

Appendix 1: The Econometric Model

First, assume a continuum of job quality, $Y^* \in (-\infty, +\infty)$. Individuals have a utility function $U(y^*, \theta)$, where y^* is a realization of Y^* and θ is a vector of other inputs. All individuals prefer higher quality jobs: $\partial U / \partial y^* > 0$. The maximum quality of job that an individual can attain in 1881, y_i^* , is a linear function of that person’s observable traits and skills, X_i , and unobservable characteristics, ε_i . The function may be different in the urban labor market than in the rural, though I assume the relevant elements of X_i are the same in each

$$\text{Rural:} \quad y_{0i}^* = \beta_0' X_i + \varepsilon_{0i} \quad (\text{A1})$$

$$\text{Urban:} \quad y_{1i}^* = \beta_1' X_i + \varepsilon_{1i} \quad (\text{A2})$$

The decision facing each individual in this model is whether to migrate to a city or remain in a rural place. Define the net benefit of moving, M_i^* , as a linear function of: Z_i , a vector of observable individual and location specific characteristics; the difference between the individual’s maximum attainable job quality in the urban regime versus that in the rural, $(y_{1i}^* - y_{0i}^*)$; and u_i , unobservable characteristics

$$M_i^* = \gamma_1' Z_i + \gamma_2 (y_{1i}^* - y_{0i}^*) + u_i \quad (\text{A3})$$

Only the outcome of the choice, M_i , is observed: $M_i = 1$ if $M_i^* \geq 0$, $M_i = 0$ otherwise.⁵⁵ In a standard endogenous regime switching model, either y_{0i} or y_{1i} is observed for each individual, depending on the value of decision variable, whereas in the current model neither y_{0i}^* nor y_{1i}^* is observed for any individual. What is observed is the job class, y_i , of every individual, either in the rural or urban regime

⁵⁵ Note that the decision is *not* whether to move or stay, as is typical in most migration studies. The decision is whether to move to an urban area or not. So an individual who moves from one rural area to another would have $M_i = 0$.

$$y_{0i} \in [1,2,3,4,5] \quad \text{if } M_i = 0$$

$$y_{1i} \in [1,2,3,4,5] \quad \text{if } M_i = 1$$

Assume the following relationship between job class, y_i , and job quality, y_i^*

$$y_{ri} = \begin{cases} 5 & \text{if } -\infty < y_{ri}^* \leq k_1 \\ 4 & \text{if } k_1 < y_{ri}^* \leq k_2 \\ 3 & \text{if } k_2 < y_{ri}^* \leq k_3 \\ 2 & \text{if } k_3 < y_{ri}^* \leq k_4 \\ 1 & \text{if } k_4 < y_{ri}^* < +\infty \end{cases} \quad r = 0,1 \quad (\text{A4})$$

where $k_1 - k_4$ are constants. Class 1 jobs are the most desirable, Class 5 the least. Under this formulation, not all jobs within a class are equivalent; indeed, the best job in any class is only marginally inferior to the worst job in the next highest class. In addition, the quality/class structure is identical in the rural and urban labor markets: the four threshold levels $k_1 - k_4$ do not vary between the two regimes. Otherwise, an individual could change job class without changing jobs simply by moving from one regime to the other.

I estimate the three-equation system defined by equations A3 and A4 by Full Information Maximum Likelihood (FIML), which proceeds in two stages. First, by substituting equations A1 and A2 into equation A3, equation A3 can be written in reduced form, in terms of observables, as

$$M_i = \begin{cases} 1 & \text{if } \gamma'W_i + v_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{A5})$$

where W contains all the elements of X and Z . I assume the disturbance terms— ε_0 , ε_2 , and v —to be i.i.d. draws from a trivariate normal distribution with mean vector zero.⁵⁶

With this distributional assumption, equation A4 becomes a standard ordered probit model with five outcomes and four threshold levels, where the outcome is determined by the latent variable y^* . I jointly estimate the three equations in equation A4 and A5 by maximizing the likelihood function

⁵⁶ This distributional assumption is standard but not innocuous, as the literature on nonparametric estimation of treatment effects demonstrates. See Heckman, “Varieties”; Manski, “Nonparametric Bounds”; and Newey et al., “Semiparametric Estimation.”

$$L(\beta_0, \beta_1, \gamma, k_1, k_2, k_3, k_4, \rho_0, \rho_1) =$$

$$\left[\begin{array}{l} \prod_{y_i=5} F(k_1 - \beta_0' X_{0i}, \gamma' W, -\rho_0) \\ \prod_{\substack{y_i=j \\ j=2,3,4}} [F(k_j - \beta_0' X_{0i}, \gamma' W, -\rho_0) - F(k_{j-1} - \beta_0' X_{0i}, \gamma' W, -\rho_0)] \\ \prod_{y_i=1} F(\beta_0' X_{0i} - k_4, \gamma' W, \rho_0) \end{array} \right]^{1-M_i}$$

$$\left[\begin{array}{l} \prod_{y_i=5} G(k_1 - \beta_1' X_{1i}, -\gamma' W, \rho_1) \\ \prod_{\substack{y_i=j \\ j=2,3,4}} [G(k_j - \beta_1' X_{1i}, -\gamma' W, \rho_1) - G(k_{j-1} - \beta_1' X_{1i}, -\gamma' W, \rho_1)] \\ \prod_{y_i=1} G(\beta_1' X_{1i} - k_4, -\gamma' W, -\rho_1) \end{array} \right]^{M_i}$$

where F and G are, respectively, the bivariate normal distribution functions of (ε_{0i}, v_i) and (ε_{1i}, v_i) , and ρ_0 and ρ_1 are the correlation coefficients of the two distributions. β_0 and β_1 include a constant, so the four threshold levels, k , are not all identified, and k_1 is normalized to zero. This estimation procedure produces consistent and asymptotically efficient estimates of the structural parameters of interest, β_0 and β_1 .

The second stage is the recovery of the structural parameters of equation A3. With estimates of β_0 and β_1 in hand, predicted values of y_{0i}^* and y_{1i}^* for each individual are $\hat{y}_{ri}^* = \hat{\beta}_r' X_i$, $r = 0, 1$. Substituting these predicted values into equation A3 in place of the unobservable y_{0i}^* and y_{1i}^* , I estimate the equation by probit maximum likelihood to obtain estimates of the structural parameters γ_1 and γ_2 .

Appendix 2: Excluding the Young

I include all male sons of working age in the estimation sample in order to achieve the maximum possible sample size and to allow for a thorough analysis of intergenerational mobility. With all sons included, results on intergenerational mobility should be seen as the primary results, and intragenerational mobility is secondary. To get a better understanding of the relationship between rural-urban migration and intragenerational mobility, the youngest males, whose jobs most likely do not represent a true starting point in the labor market, should be excluded.

Appendix Table 1 reports the key empirical results from an analysis that excludes the youngest males from the sample. The analysis follows that from the text; the only change is the omission of the 1,485 males aged 9–13 in 1851 and the 78 males older than 14 listed as “scholars” in 1851. The most important thing to note is that the results are largely unchanged. The parameter measuring the selection of the urban migrants is slightly greater, and the treatment effect for migrants is slightly smaller. For rural persisters, the selection parameter is slightly less negative, and the (implied) treatment effect is substantially lower. The effect of anticipated socioeconomic gain

APPENDIX TABLE 1
 ROBUSTNESS OF KEY PARAMETERS TO EXCLUSION OF YOUNG MALES: FOCUS
 ON INTRAGENERATIONAL MOBILITY

	Estimate	Std Er- ror	90% Confidence Interval	
Selection of Urban Migrants	0.1552	0.0589	0.0514	0.2421
Selection of Rural Persisters	-0.0411	0.0714	-0.1749	0.0522
Treatment Effect, Urban Migrants	0.1745	0.2387	-0.0819	0.8338
Treatment Effect, Rural Persisters	0.0601	0.2254	-0.2131	0.6928
Effect of ($y_{1i}^* - y_{0i}^*$) on Migration Probability	0.1943	0.1418	0.0481	0.5080
Effect of Own Class in 1851 on 1881 Class, Movers				
1,2 - Professional, Intermediate	2.2861	0.3233	1.7544	2.8179
3 - Skilled	0.5045	0.2227	0.1382	0.8709
4 - Partly Skilled	0.3541	0.2649	-0.0817	0.7899
Effect of Own Class in 1851 on 1881 Class, Stayers				
1,2 - Professional, Intermediate	1.2882	0.4005	0.6295	1.9469
3 - Skilled	0.4352	0.2737	-0.0150	0.8854
4 - Partly Skilled	0.2283	0.2007	-0.1017	0.5584
Effect of Father's Class on 1881 Class, Movers				
1 - Professional	1.1952	0.1837	0.8930	1.4973
2 - Intermediate	0.3565	0.1151	0.1672	0.5459
3 - Skilled	-0.0131	0.1395	-0.2426	0.2165
4 - Partly Skilled	0.7234	0.1675	0.4479	0.9990
Effect of Father's Class on 1881 Class, Stayers				
1 - Professional	1.4815	0.2682	1.0404	1.9226
2 - Intermediate	0.3572	0.1508	0.1091	0.6053
3 - Skilled	0.3535	0.0928	0.2008	0.5061
4 - Partly Skilled	-0.0115	0.0902	-0.1599	0.1369

Notes: $N = 2,211$ males aged 14–29 in 1851. Standard errors and confidence intervals of the first five parameters are calculated by bootstrapping via data resampling with 1,000 repetitions.

on migration probability is slightly smaller, though still economically and statistically significant. Perhaps the most significant change is in the effect of own class in 1851 and father's class on 1881 class attainment. For both migrants and persisters, the same essential patterns appear; however, the effect of own class in 1851 is stronger with the youngest males excluded from the analysis. With older males the 1851 class variable is more representative of their true starting point in the labor market, so it should be expected to be a better predictor of later adult socioeconomic status.

REFERENCES

- Amemiya, Takeshi. *Advanced Econometrics*. Oxford: Basil Blackwell, 1986.
- Anderson, Michael. *National Sample from the 1851 Census of Great Britain: Introductory User Guide*. University of Edinburgh, Dept. Of Economic and Social History, September 1987. [<http://www.data-archive.ac.uk/doc/1316/mrdoc/pdf/a1316uab.pdf>]
- Armstrong, W. A. "The Use of Information about Occupation." In *Nineteenth-Century Society*, edited by E. A. Wrigley, 191–310. Cambridge: Cambridge University Press, 1972.

- Baines, Dudley. *Migration in a Mature Economy: Emigration and Internal Migration in England and Wales, 1861–1900*. Cambridge: Cambridge University Press, 1985.
- _____. “Population, Migration and Regional Development, 1870–1939.” In *The Economic History of Britain Since 1700*, vol 2, edited by R. Floud and D. McCloskey, 29–61. Cambridge: Cambridge University Press, 1994.
- Black, Duncan, and Vernon Henderson. “A Theory of Urban Growth.” *Journal of Political Economy* 107, no. 2 (1999): 252–84.
- Borjas, George J. “Self-Selection and the Earnings of Immigrants.” *American Economic Review* 77, no. 1 (1987): 531–53.
- Booth, Charles, Ed. *Life and Labour of the People in London*. 9 vols. London: Macmillan and Co., 1892–97.
- Boyer, George, and Timothy J. Hatton. “Migration and Labour Market Integration in Late Nineteenth-Century England and Wales.” *Economic History Review* 50, no. 4 (1997): 697–734.
- Crafts, N. F. R. *British Economic Growth during the Industrial Revolution*. Oxford: Oxford University Press, 1985.
- Ferrie, Joseph P. “How Ya Gonna Keep ’Em Down on the Farm [When They’ve Seen Schenectady]? Rural to Urban Migration in the U.S., 1850–70.” Prepared for the Cliometrics Conference, May 1999. [<http://www.faculty.econ.nwu.edu/faculty/ferrie/papers/urban.pdf>].
- _____. *Yankeys Now: Immigrants in the Antebellum United States, 1840–1860*. New York: Oxford University Press, 1999.
- _____. “The Poor and the Dead: Socioeconomic Status and Mortality in the U.S., 1850–1860.” NBER Working Paper No. h0135, Cambridge, MA, August 2001.
- Great Britain. General Register Office. *The Registrar-General's Decennial Supplement. England and Wales*. London: H.M.S.O., 1921.
- _____. *Census of England and Wales, 1921. Classification of Occupations*. London: H.M.S.O., 1924.
- Hatton, Timothy J., and Jeffrey G. Williamson. “What Drove the Mass Migrations from Europe in the Late Nineteenth Century?” *Population and Development Review* 20, no. 3 (September 1994): 533–59.
- Heckman, James. “Varieties of Selection Bias.” *The American Economic Review* 80, no. 2 (1990): 313–18.
- Higgs, Edward. *A Clearer Sense of the Census*. London: HMSO, 1986.
- Hunt, E. H. *Regional Wage Variations in Britain, 1850–1914*. Oxford: Oxford University Press, 1973.
- _____. “Industrialization and Regional Inequality: Wages in Britain, 1760–1914.” *This JOURNAL* 46, no. 4 (1986): 935–66.
- Lee, C. H. *British Regional Employment Statistics, 1841–1971*. Cambridge: Cambridge University Press, 1971.
- Long, Jason. “The Economic Return to Primary Schooling in Victorian England.” Unpublished Manuscript. [<http://www.colby.edu/economics/faculty/jmlong/research/schooling.pdf>].
- Long, Jason, and Ferrie, Joseph P. “A Tale of Two Labor Markets: Career Mobility in the U.K. (1851–81) and U.S. (1850–80).” Prepared for the Economic History Society Annual Conference, Durham, England, April 2003. [<http://www.colby.edu/economics/faculty/jmlong/research/usukmobility.pdf>].
- Lucas, Robert E., Jr. “On the Mechanics of Economic Development.” *Journal of Monetary Economics* 22, no. 1 (July 1988): 3–42.

- Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.
- Manski, Charles. "Nonparametric Bounds on Treatment Effects." *The American Economic Review* 80, no. 2 (1990): 319–23.
- Mitchell, B. R., with Phyllis Deane. *Abstract of British Historical Statistics*. Cambridge: Cambridge University Press, 1962.
- Mokyr, Joel. *Why Ireland Starved*. London: George Allen & Unwin, 1983.
- Newey, Whitney K., James L. Powell, and James R. Walker. "Semiparametric Estimation of Selection Models: Some Empirical Results." *The American Economic Review* 80, no. 2 (1990): 324–28.
- Pollard, Sidney. "Labour in Great Britain." In *The Cambridge Economic History of Europe, vol 7: The Industrial Economies: Capital, Labour, and Enterprise*, edited by P. Mathias and M. M. Postan, 97–179. Cambridge: Cambridge University Press, 1978.
- Pooley, Colin, and Jean Turnbull. *Migration and Mobility in Britain Since the Eighteenth Century*. London: UCL Press, 1998.
- Ravenstein, E. G. "The Laws of Migration." *Journal of the Royal Statistical Society* 48, no. 2 (1885): 167–227.
- Redford, Arthur. *Labour Migration in England, 1800–1850*, 2nd ed. Manchester: Manchester University Press, 1964 (1926).
- Royle, Stephen A. "Social Stratification from Early Census Returns: A New Approach." *AREA* 9 (1977): 215–19.
- Sjaastad, Larry A. "The Costs and Returns of Human Migration." *Journal of Political Economy* 70, no. 5 (1962): 80–93.
- Southall, Humphrey. "Regional Unemployment Patterns among Skilled Engineers in Britain, 1851-1914." *Journal of Historical Geography* 12, no. 3 (1986): 268–86.
- Southall, Humphrey, and David M. Gilbert. "A Good Time to Wed?: Marriage and Economic Distress in England and Wales, 1839 to 1914." *Economic History Review* 49, no. 1 (1996): 35–57.
- Steckel, Richard H. "Census Matching and Migration: A Research Strategy." *Historical Methods* 21, no. 2 (1988): 52–60.
- _____. "The Health and Mortality of Women and Children, 1850–1860." *This Journal* 48, no. 2 (1988): 333–45.
- Tuttle, Carolyn. "The Role of Children in the Industrial Revolution." Ph.D. diss., Northwestern University, 1985.
- Williamson, Jeffrey G. *Coping with City Growth during the British Industrial Revolution*. Cambridge: Cambridge University Press, 1990.
- Woods, Robert. *The Demography of Victorian England and Wales*. Cambridge: Cambridge University Press, 2000.
- Wrigley, E. A., and R. S. Schofield, *The Population History of England, 1541–1871*. Cambridge, MA: Harvard University Press, 1981.