
Volume 30, Issue 4
December 2008
ISSN 0164-0704
ELSEVIER
Journal of MACROECONOMICS
Editors:
DOUGLAS McMILLIN
THEODORE PALIVOS
CHRIS PAPAGEORGIOU
Administrative Editor:
TAEHEE HAN
Editorial Board:
T.M. ANDERSEN
B. BERNANKE
J. CABALLE
S. CHAKRABORTY
B.C. DANIEL
M.B. DEVEREUX
M.S. EICHENBAUM
T. EICHER
J.S. FACKLER
M.S. FELDSTEIN
M.R. GARFINKEL
J.C. HALTIWANGER
P.W. HOWITT
R.G. HUBBARD
H. HUIZINGA
B.F. INGRAM
S. KIM
I. KING
S. KOZICKI
W.D. LASTRAPES
B. LOCKWOOD
I. MIHOV
K. MINO
K. NISHIMURA
F. POSTEL-VINAY
B. RAVIKUMAR
L. REICHLIN
P.A. SAMUELSON
A.C. SIBERT
C.J. WALLER
C.E. WALSH
P. WANG
T. WHITED
C.K. YIP
CONTENTS
B. Taub and R. Zhao, Why doesn't Luxembourg send all its capital to India? 1335
F. Caliendo and K.X.D. Huang, Overconfidence and consumption over the life cycle 1347
T.E. Cone, Optimal information acquisition and monetary policy 1370
M.H. Marquis and B. Trehan, On using relative prices to measure capital-specific technological progress 1390
P.-R. Agénor, Health and infrastructure in a model of endogenous growth 1407
T. Osang and J. Sarkar, Endogenous mortality, human capital and economic growth 1423
T. Barr and U. Roy, The effect of labor market monopsony on economic growth 1446
D. Varvarigos, Sustained output growth under uncertainty: A simple model with human capital 1468
R. Censolo and C. Colombo, Public consumption composition in a growing economy 1479
M.A. Gomez Suarez, Utility and production externalities, equilibrium efficiency and leisure specification 1496
S.H.K. Tang, N. Groenewold and C.K.Y. Leung, The link between institutions, technical change and macroeconomic volatility 1520
A. Trew, Efficiency, depth and growth: Quantitative implications of finance and growth theory 1550
(Contents continued on outside back cover)
LAST ISSUE OF THIS VOLUME
Available online at
ScienceDirect
www.sciencedirect.com
In collaboration with the E.J. Ourso College of Business, Louisiana State University

Journal of Macroeconomics 30 (2008) 1446–1467

Contents lists available at ScienceDirect

Journal of Macroeconomics

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

The effect of labor market monopsony on economic growth

Tavis Barr *, Udayan Roy

Long Island University, Department of Economics, C.W. Post Campus, 720 Northern Boulevard, Brookville, NY 11548, United States

ARTICLE INFO

Article history:
Received 26 November 2007
Accepted 13 May 2008
Available online 23 May 2008

JEL Classification:
O57
J42
L13

Keywords:
Economic growth
Human capital
Monopsony
Wage share

ABSTRACT

In this endogenous growth model, a minimum efficient scale of production and workers' home-to-work travel costs combine to give firms monopsony power, and this monopsony power leads to slower growth. Monopsony drives the wage below the marginal product of labor. This lower wage leads to lower investment in human capital and thereby to a lower growth rate. This makes investment in human capital – and therefore the growth rate – suboptimal. We provide evidence from a cross-country panel to support our model: Urbanization, which we assume is determined by a country's exogenous population density and cropland area, positively impacts the wage share of output; the wage share positively impacts educational attainment; higher-income countries have higher wage shares; and within-country upticks in the wage share have a positive lagged effect on the growth rate.

1. Introduction

In this paper, we analyze and test a model of endogenous growth in which monopsony power in the labor market has a negative effect on the rate of growth. The combination of a minimum efficient scale of production in the final-good sector and workers' costs of home-to-work travel, in a location model, creates monopsony power for employers. Workers' homes are distributed uniformly throughout each labor market. Increasing returns in the final-good sector – which is modeled here through the assumption of a minimum efficient scale – and the resulting agglomeration, lead to the presence of only a finite number of firms in each labor market. As a result, the distance between any firm and its nearest competitor is non-trivial. This and the assumption that travel between home and work is costly for workers imply that a small reduction in the wage paid by a firm will not provoke all of its employees to leave. In other words, each firm faces a rising labor supply curve, which implies

* Corresponding author. Tel.: +1 516 299 2321.
E-mail address: tavis.barr@liu.edu (T. Barr).

0164-0704/$ - see front matter
doi:10.1016/j.jmacro.2008.05.001

monopsony power.[1] Any decrease in firms' monopsony power – which could be caused by a decrease either in workers' travel costs or in the minimum efficient scale or by an increase in urban density – leads to higher wages. Higher wages, in turn, imply a higher incentive to invest in human capital. In this way, a decrease in monopsony power increases the rate of human capital accumulation and, thereby, the rate of growth of per capita GDP.

We also show that monopsony in the labor market makes the equilibrium growth rate suboptimal. The representative individual's incentive to invest in human capital is determined by the wage. The fictional social planner's incentive, on the other hand, is determined by the marginal product of labor. Monopsony pulls the wage below the marginal product of labor and, thereby, makes the return to human capital appear *smaller* to the representative individual in the decentralized economy than to the social planner in the command economy. Consequently, there is a perpetual underinvestment in human capital, causing the equilibrium growth rate under monopsony to fall below the optimal rate.

We test the links leading from higher urbanization rates to higher wages to higher investment in education to higher growth rates by using a panel of 33 countries for which sufficient data are available. We find that urbanization rates – which we use as a proxy for urban density – positively impact the wage share of output; the wage share positively impacts educational attainment; higher-income countries have higher wage shares; and, while the large variation of wage shares between countries does not significantly explain their varying growth rates, an uptick in the wage share within a country has a positive effect on growth at a longer time delay than what is generally considered a business cycle period.

Models in which firms have monopsony power have appeared over the last two decades in various parts of the labor economics literature (see for example Albrecht and Jovanovic, 1986; Burdett and Mortensen, 1989; Green et al., 1996). Monopsony has become a relevant view of the labor market and we believe that the general-equilibrium effects of monopsony on growth need to be analyzed. The source of monopsony power in the recent literature ranges from bargaining power over firm-specific capital (Albrecht and Jovanovic, 1986) to limited arrival of information about job openings (Burdett and Mortensen, 1989) to workers' unwillingness to quit for an infinitesimally better-paying job (Barr, 2002). We have chosen to model monopsony power based on urban structure because cross-country proxy measures for urban structure are at least somewhat obtainable; by contrast, internationally comparable data on the nature of specific capital or job matching or quit behavior would be difficult to obtain for a large sample of countries.[2]

As to the empirical importance of monopsony power, Manning (2003, p. 80–81) finds that most available estimates of the elasticity of labor supply are between 2 and 5, which is nowhere near infinity. (Even an elasticity of 5 implies a wage that is 17% below the marginal product of labor.) Based on an extensive literature survey, Manning (2003, p. 361) argues that "monopsony can provide a much better explanation than perfect competition of a wide range of labor market phenomena" and goes on to list fifteen such phenomena including wage dispersion, the observed effect of employer size on the wage, and the fact that separation rates are lower for high-wage workers. Because of the importance of the labor market, its imperfections have a strong influence on relative prices and quantities in the rest of the economy.

The theory of economic growth has paid little attention to the effect of labor market imperfections on growth. Of course, market failure has played an important role in the theory of endogenous growth. For example, Arrow (1962) and Romer (1986) have explored knowledge spillovers in the production of the final good. Romer (1990) and Aghion and Howitt (1997) have used the standing-on-the-shoulders-of-giants externality in R&D. Some of these models – especially the ones that assume that inventions are protected by patents – incorporate monopolistic competition in final-good or intermediate-good markets. Barro (1990) has explored the effect on growth of government-provided infrastructure-type public goods that are essential to production. And yet, these models all assume perfectly competitive labor markets. Aghion and Howitt (1997, Chapter 4) describe several models – including some of their

[1] Monopsony is not to be taken literally to mean a one-buyer market. For our purposes, monopsony exists when firms face upward-rising supply curves for labor. This happens here because the marginal worker (who travels from further away) is more expensive to entice than the average worker. More accurate terms for our model of the labor market would be monopsonistic competition or oligopsony, but we use monopsony because it is a term that appears to have stuck; see Manning (2003, p. 3).

[2] Ridder and van den Berg (2003) demonstrates one means of generating cross-country estimates of monopsony power based on quit behavior, which could be analyzed to produce further evidence, however the number of countries for which such data are available is smaller than the number we are able to analyze using the current framework.

own – in which unemployment is caused by labor market rigidities that make it time-consuming to match workers and firms, and they then explore the effect on unemployment of the creative destruction wrought by rapid technological progress. However, unlike our focus in this paper, these models do not discuss how labor market imperfections can affect the growth rate.

Recent papers have also examined how increased competition in the market may have additional welfare effects beyond simply realigning prices and output toward the Pareto optimal level. For example, Aghion and Schankerman (2004) show that when firms are not all equally productive, increased substitutability between products will cause the market share of more productive firms to increase relative to that of less productive firms, which has the additional effects of encouraging cost-reducing investments by low-cost firms, encouraging entry of low-cost producers, and discouraging high-cost entrants. They show that when governments accept bribes, the initial market share of less productive firms can affect the subsequent share of such firms, because they will pay politicians not to increase substitutability. Another welfare effect of competitiveness is illustrated by Tse (2000), who shows how product market distortions can lower the wage and, through the wage, the investment in education. Tse combines monopoly in the product market with industry-specific human capital to derive a connection between product market distortions and the return to education. We add to this list of welfare effects by illustrating a means through which competitiveness can affect the rate of human capital accumulation and thereby also affect growth. Like Tse, we are interested in the idea that market distortions may affect the return to education. However, we are investigating an altogether separate (though not mutually exclusive) means by which markets may be distorted: in our model, there is monopsony in the labor market and each worker's human capital is equally productive in all firms. We show how the resulting distortion in the return to education affects the model economy's growth rate, and, finally, we use cross-country regressions to test for the dynamics predicted by the model.[3]

We follow Helsley and Strange (1990), and Kim (1990), in adapting Salop's (1979) model of spatial monopolistic competition in product markets to model monopsonistic competition (or, oligopsony) in labor markets.[4] Workers and firms are distributed uniformly along the unit circle. We assume that there is a minimum efficient scale of production, possibly due to transaction costs or due to complementarity of worker effort, that implies self-employment will not be as lucrative as working for a firm. In other words, even under free entry not every point on the unit circle will have a firm located at it. As a result, workers must incur a cost in getting from home to work and back. Therefore, a firm that wishes to hire more people than it currently employs will have to offer a higher wage so as to make it worthwhile for potential employees who live farther away to accept its offer. In other words, each firm faces a rising labor supply curve as in monopsony or monopsonistic competition. And, as was explained earlier, this leads to the "monopsony wedge" between the wage and the marginal product of labor.

This divergence between the wage and the marginal product of labor is the key mechanism by which labor market distortions affect growth in our paper. While some of the parameters underlying monopsony power (such as travel costs and urban structure) may potentially have other general-equilibrium price effects (for example on consumer demand), what is crucial to note is that monopsony power, in whatever form it appears, drives a wedge between the wage and the marginal product of labor and thereby lowers the return to human capital. Thus, while we cannot guarantee that models of monopsony based on other sources (such as specific capital or worker job-switching costs) would not generate other general-equilibrium price effects, we can predict that, to the extent that they create a wedge between the wage and the marginal product of labor, they can be adapted without very much contortion to generate our endogenous growth mechanism.

Our model also shows how cross-country differences in the ease with which workers reach existing jobs lead to cross-country differences in monopsony power and, therefore, growth. Although we will

[3] Tse's model could also potentially be encapsulated into a growth model, and tested using cross-country regressions as our model is. The empirical predictions of Tse's model would be difficult to distinguish from ours, though such an exercise would be feasible if one has plausible cross-country proxy variables for monopoly power. Of course, the two models are not mutually exclusive. In any event, the calibration type of exercise that Tse uses to test his model empirically has value separate from the regression exercise pursued here; our model might benefit from a calibration test – though we do not pursue such an investigation here – and a model like Tse's might benefit from regression evidence.

[4] See Duranton and Puga (2003) for a survey of urban locality models with spatially differentiated producers.

briefly discuss comparative statics, our focus is on the means by which differences in labor market mobility generate comparative growth dynamics. This may happen through several avenues. First, taking our travel cost literally as the difficulty of commuting, countries with either denser labor markets or better infrastructure will have greater labor mobility. Aghion and Schankerman (2004) interpret the unit cost of travel as a proxy for transport and telecommunications infrastructure, and for institutional barriers to economic activity. So, a reduction in the unit cost of travel is seen broadly as an increase in competition. Aghion and Schankerman also generalize Salop (1979) by distinguishing between high-cost and low-cost incumbents and between high-cost and low-cost potential entrants, and by allowing cost-reducing investments by incumbents. Helsley and Strange (1990), Kim (1990), and Duranton and Puga (2003) also see the unit cost of travel as the economic loss (or, retraining cost) when there is a mismatch between the needs of an employer and the abilities of the employee.

Labor market mobility itself may be influenced by several factors. We think of the overall population density of a country as an exogenous factor, since it is the result of agricultural productivity effects that work at a lower frequency than economic development, and is not known to be highly correlated with the current states of the world's economies. Infrastructure, on the other hand, is a public good that is probably endogenous to the level of development. As a country's output improves, we would expect its infrastructure to move from a primitive steady state to an advanced steady state, and the resultant transport costs to fall from a high steady state to a low steady state; however, we expect that, unlike human capital, transportation infrastructure will not produce steady-state growth dynamics.[5]

Worker mobility is also potentially affected by social and institutional factors. More figuratively, we can think of the travel cost as a cost to cover either a psychological distance or a social distance. Thought of as a psychological cost, it reflects the idea that people take jobs that are far from their preferred occupation; societies in which people have more weak ties (i.e., friendships with people who are culturally different) might find them able to traverse this distance more easily. The social distance would reflect the cost of finding out about a job; if social networks are sparse, or are wired such that people's connections are highly local, then finding out about a job far away (or making the contacts necessary to receive an offer) could be costly; on the other hand, if the social network has more random connections, then it may have a "small world" property whereby information from far away is easy to obtain (see Carayol and Roux, 2003). In any event, firms will have more monopsony power when social networks are weak. Labor mobility will also be affected by political and institutional factors such as the strength of labor laws (minimum wage laws, severance restrictions, etc.), and in the extreme case, political authoritarianism and civil conflict. While we think such factors are interesting, we do not have available data on these factors that we think are exogenous to the overall growth problem. Our strategy will therefore be to model labor market mobility as exogenously determined by factors of population density and available farmland, and, in regressions where growth is the dependent variable, output per person will be included as a right-hand-side variable, so that any variation in labor mobility will be thought of as conditional on the level of development.

The model therefore has several testable predictions. First, we use the population density and available farmland as exogenous variables affecting the density of labor markets through their impact on the urbanization rate; the model predicts that denser and less agricultural countries will be more urban, and monopsony power in turn should be weaker, and the share of wages in total output higher, in more urban countries. Second, a lower wage share should lead to less accumulation of human capital. It needs to be emphasized that this second prediction is the most unusual prediction of our model, and is not typical of endogenous growth models; while more advanced economies will clearly have a higher wage *level* due to accumulation of physical and human capital, the *share* of wages in output is

[5] We have taken population density as exogenous to the economy's growth problem. Other papers take city *size* as endogenously determined by technology, and have generally focussed on studying the types of arbitrage that would cause individuals to distribute themselves between cities *within* a country in a way that conforms to Zipf's Law (see Black and Henderson, 1999; Cordoba, 2004; Rossi-Hansberg and Wright, 2007). The assumptions behind these models are generally that city size is limited either in a monocentric city by commuting costs (since a larger city requires a larger maximum commuting distance) or in specialized cities by the total demand for goods produced by that particular city's industry. Perpetual productivity growth, whether exogenous or endogenous, causes migration to a particular city until the marginal worker is indifferent between migrating to two particular cities. By contrast, the determinants of urban density *between* countries are more likely found outside of the cities: The overall level of agricultural productivity and the available farmland per person. We take these factors as exogenous to our model.

generally determined by Cobb–Douglas coefficients, and is therefore invariant to the overall level of economic development. Third, as has been extensively studied (see Barro and Lee, 1993), lower investment in human capital should lead to a lower rate of growth.[6]

The rest of the paper is as follows: Section 2 describes and then analyzes the equilibrium growth path of our model. Section 3 compares the equilibrium and optimal growth paths. Section 4 discusses what the data have to say about the predictions of our model. Section 5 concludes the paper.

2. Work-related travel and monopsony

In this section, we develop our model of monopsony in the labor market. We show that monopsony leads to lower wages and that lower wages in turn lead to lower returns to the accumulation of human capital and, therefore, to slower growth.

2.1. Workers and firms

We assume an economy that is segmented into various *localities*, each of which may be thought of as a self-contained labor market. Each locality resembles a unit circle and is populated by $\bar{L}$ individuals who live uniformly distributed over the unit circle. Any increase in the economy's population leads, we assume, to a proportionate increase in the number of localities but not to any change in $\bar{L}$, which is the population density of each locality.[7]

At any time t, there are N_t firms in each locality, indexed $n = 1, 2, \ldots, N_t$. The firms that exist at t did not exist at $s < t$ and will not exist at $s' > t$; each instant in time has its own set of N_t very short-lived firms. These N_t firms are evenly distributed along the unit circle at the locations x_t, $x_t + 1/N_t, x_t + 2/N_t, \ldots, x_t + (N_t - 1)/N_t$, where x_t is a sequence of i.i.d. random variables.[8]

At time t, firm n announces a wage, w_{nt}, and uses L_{nt} units of labor (supplied by the individuals who accept its job offer) to produce the final good, Y, using the following technology:

$$Y_{nt} = \begin{cases} h_t L_{nt}, & \text{if } L_{nt} \geqslant F, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where F is the minimum amount of labor at which production is possible (hereafter, minimum efficient scale), and h_t is each worker's endowment of human capital at time t.[9] As we will show later, the decreasing minimum efficient scale implied by this production function is crucial to ensuring the existence of monopsonistic power. It means that even under free entry not every point on the unit circle

---

[6] Although our model suggests that a lower wage share should lead to lower growth only through its effect on human capital accumulation, we do not wish to preclude the idea that the wage share may affect growth in other ways.

[7] We will show later that this assumption is necessary to remove the empirically dubious prediction – known as the scale effect – that more populous countries grow faster. Although we have not modeled population density, it seems plausible to us that the various costs and benefits of population density that would bear on the determination of the equilibrium population density in a model in which population density is endogenous would be independent of the overall population of a country. Also, our assumption is, in essence, consistent with a well-known regularity in the distribution of city sizes: the power law. This law states that the relative frequency of cities of a given size remains the same across countries and over time. In other words, a doubling of a country's population will double the number of cities of population x, for every x. See Fujita, Krugman, and Venables (1999, pp. 215–216). In the empirical section of this paper, we assume that a country's urbanization rate is a good proxy for the population density of a typical self-contained labor market in that country.

[8] The main effect of assuming this continuous and random relocation of firms is that no worker is systematically unlucky, in the sense of having higher travel-to-work costs than another worker. As a result, we can analyze the dynamics of the economy within a representative agent framework. If the locations of workers *and* of firms are both permanently set, a worker who happens to be located farther away from the nearest firm than another worker, would be permanently disadvantaged. The continuous random relocation of firms assumed here levels the playing field for all workers and simplifies our analysis by making each worker a representative agent. The continuous random relocation of infinitely-lived *workers* would give identical results if firms are infinitely-lived and stationary. See Section 2.10 for more on this issue.

[9] Our interpretation of this minimum efficient scale is that either there are prohibitive transaction costs when the good is produced by too small a firm, or that there are complementarities in work effort between individual workers that will not be captured unless a single agent (i.e., firm owner) is able to capture the external effects of these complementarities. As a result, workers will subject themselves to the monopsony rent of a firm because that firm is more productive than they would be if self-employed. Not wanting to decide between these interpretations, and not wanting to pollute the model with extra terms, we simply declare that employment cannot fall below some level F.

will have a firm located at it; there will always exist some space on the unit circle between neighboring firms. As home-to-work travel is costly, the typical firm will be able to hold on to some of its employees – especially those who live nearby – even if its wage is a shade lower than that paid by other firms. That is, the competition for workers that each firm faces from its neighboring firms will never be as intense as under perfect competition. In other words, the labor market will be monopsonistic.[10]

We assume that each individual is endowed with one unit of labor. An employee of firm n who lives a distance d away from the firm needs βd units of labor to travel between home and work and, therefore, earns $(1-\beta d)h_t w_{nt}$ in wage income. A worker will, naturally, accept the job where wage income is highest.[11]

2.2. The supply of workers

In any given locality (or, unit circle), firm n is located a distance $1/N_t$ away from firms $n-1$ and $n+1$, its immediate neighbors. We assume that a worker – hereafter, the *marginal worker* – lives a distance d_{nt} from firm n and, therefore, a distance $(1/N_t) - d_{nt}$ from firm $n+1$, and is indifferent between working for these two firms. Assuming that firm n pays the wage w_{nt} and all other firms pay the wage w_t, d_{nt} can be determined from $(1-\beta d_{nt})h_t w_{nt} = [1-\beta((1/N_t) - d_{nt})]h_t w_t$ to be

$$d_{nt} = \frac{w_{nt} - \left(1 - \frac{\beta}{N_t}\right)w_t}{\beta(w_{nt} + w_t)}. \tag{2}$$

We assume that it is so costly for a worker to live in one locality and go to work in another, that in equilibrium no worker will live in one locality (or, unit circle) and work in another. It then follows that the number of workers who live within a distance d_{nt} of firm n is also the number of workers working in firm n. This number is, therefore, $2\bar{L}d_{nt}$.[12]

2.3. Travel costs

To determine the amount of labor employed by a typical firm, such as firm n, we need to calculate the amount of labor spent on travel by the firm's $2\bar{L}d_{nt}$ employees; it is only the remaining labor that can be used in production.

An employee who lives a distance s from firm n spends βs units of labor in travel (and, therefore, supplies $1-\beta s$ units of labor to the firm). Therefore, the total labor spent on travel by the employees of firm n is:

$$T_{nt} = 2\bar{L}\int_0^{d_{nt}} \beta s ds = \bar{L}\beta d_{nt}^2. \tag{3}$$

2.4. The supply of labor

It follows that the labor supply facing firm n is

$$L_{nt} = 2\bar{L}d_{nt} - \bar{L}\beta d_{nt}^2 = \bar{L}(2d_{nt} - \beta d_{nt}^2), \tag{4}$$

---

[10] Our results are true even if we assume that all individuals do not necessarily have the same endowment of human capital at time t. Under such an assumption, we can show that as long as all individuals begin with the same initial endowment h_0, they will have the same human capital at all t. Therefore, the assumption of common h_t serves to simplify exposition without affecting the generality of our results.

[11] In a larger, non-literal sense, the labor needed for travel, βd, is meant to represent the loss of productivity that occurs when a worker is employed in a job (represented by the firm's address) that is not her dream job (represented by her home address). Therefore, β could have several interpretations such as those discussed in the Introduction. Also, note that w_{nt} is the real wage paid to *effective* labor.

[12] Although the expression for d_{nt} in Eq. (2) is adequate for the purpose of investigating the symmetric equilibrium of our model, it is not applicable for values of w_{nt} that are "too high". To see why, note that when $w_{nt} = w_t + \beta/N_t$, even the workers who share the same address as firm $n+1$ would be willing to work for firm n. If w_{nt} rises even infinitesimally above $w_t + \beta/N_t$, *all* employees of firm $n+1$ would switch to firm n. This implies a discontinuous increase in d_{nt} at $w_{nt} = w_t + \beta/N_t$. Similar discontinuities occur at still higher values of w_{nt} as the other firms suffer the fate of firm $n+1$. See Salop (1979) for a detailed discussion.

where d_{nt} is given by Eq. (2). We assume $2N_t > \beta$ to ensure that this labor supply is positively sloped. That is, it can be checked that $\partial L_{nt}/\partial w_{nt} > 0$ if $2N_t > \beta$.

It can also be checked that firm n's elasticity of labor supply is

$$\epsilon_{nt} \equiv \frac{w_{nt}}{L_{nt}} \cdot \frac{\partial L_{nt}}{\partial w_{nt}} = \frac{2(1-\beta d_{nt})\left(2-\frac{\beta}{N_t}\right)w_{nt}w_t}{\beta(2d_{nt}-\beta d_{nt}^2)(w_{nt}+w_t)^2} > 0. \tag{5}$$

The positive sign follows from our assumptions that $1-\beta d_{nt}$, the labor supplied by the marginal employee, is positive; that $2N_t > \beta$; and that $L_{nt} > 0$, which implies, via Eq. (4), that $2d_{nt} - \beta d_{nt}^2 > 0$.

2.5. Profit maximization under monopsony

At any time t, firm n chooses w_{nt} to maximize profit,

$$\pi_{nt} = Y_{nt} - w_{nt}h_tL_{nt}, \tag{6}$$

subject to its labor supply, Eq. (4). Ignoring the $L_{nt} < F$ outcome, we get $\pi_{nt} = (1-w_{nt})h_tL_{nt}$, and the first-order condition for its maximization – subject to Eqs. (2) and (4) – yields

$$\frac{1-w_{nt}}{w_{nt}} = \frac{1}{\epsilon_{nt}}. \tag{7}$$

Denoting the *effective labor* employed in firm n by h_tL_{nt}, we see that the left-hand side of Eq. (7) represents the *monopsony wedge* between the marginal product of effective labor – which is unity for the $L_{nt} > F$ case; see Eq. (1) – and the wage paid to effective labor, w_{nt}. Referring to the size of this wedge loosely as the extent of firm n's monopsony power, Eq. (7) implies

Lemma 1. *A firm's monopsony power is inversely related to the elasticity of its labor supply.*

2.6. Free entry

We assume that firms enter the final-good sector as long as they expect to be profitable.[13] Now, Eq. (7) implies that

$$w_{nt} = \frac{\epsilon_{nt}}{1+\epsilon_{nt}} < 1, \tag{8}$$

which in turn implies that the equilibrium level of profits must always be positive: $\pi_{nt} = (1-w_{nt})h_tL_{nt} > 0$. This provides an unending incentive for new firms to enter the final-good sector. Therefore, N_t, the number of firms on the unit circle, is limited only by the requirement that L_{nt} must not fall below F, the minimum amount of labor that a firm needs to be productive.[14] Therefore, Eq. (4) implies that, ignoring integer issues, free-entry equilibrium must satisfy

[13] Our free-entry assumption is a reminder that monopsonistic competition is a more accurate term for our model than monopsony.

[14] It may be surprising that firms in this equilibrium earn positive profits by paying $w < 1$ and earning $\pi_{nt} = (1-w_{nt})h_tL_{nt} = (1-w)h_tF > 0$, even though there is free entry. A free-entry equilibrium in which firms make strictly positive profits can exist when the outcome at each instant t is the subgame perfect equilibrium of the following two-stage game: first, an infinite number of potential firms decide whether or not to enter the final-good sector. Then, in the second stage, the entrants simultaneously choose (a) their locations on the unit circle, and (b) the wages they would pay their workers. To see the outcome of this game, consider the stage-two equilibrium first, taking the number of entrants, N_t, as given, and comparing it to N (see Eq. (11)). Clearly, for the $N_t \leqslant N$ case, the wage given by Eq. (8) and employment L_{nt} given by Eq. (4), with $d_{nt} = 0.5/N_t$, are Nash equilibria. For the $N_t \geqslant N+1$ case, it can be shown that the stage-two equilibrium is as follows: A subgroup $\mathscr{E}$ of N firms will be distributed evenly on the unit circle – exactly as in the $N_t = N$ case – such that member of subgroup $\mathscr{E}$ is a distance $1/N$ away from its two closest neighbors from within subgroup $\mathscr{E}$. These firms pay the wage $w = 1$, hire F workers each, and earn zero profits: $\pi_{nt} = (1-w)h_tL_{nt} = 0$. Each of the remaining $N_t - N$ firms may locate anywhere on the unit circle – it does not matter where. As for their chosen wages, if such a firm is located a distance x from the nearest member of the aforementioned subgroup $\mathscr{E}$ of firms, it will pay a wage that is infinitesimally less than $1-\beta x$, hire zero units of labor, and earn zero profits. (It can be checked that no unilateral deviation from this outcome can leave any firm better off.) It follows that in stage one, when potential firms must decide whether or not to enter, the number of entrants in the subgame perfect equilibrium must be exactly N. Therefore, our free-entry outcome with strictly positive profits is formally the subgame perfect equilibrium of this two-stage game.

$$L_{nt} = \bar{L}(2d_{nt} - \beta d_{nt}^2) = F. \tag{9}$$

This quadratic equation has *two* solutions for d_{nt}, which is the distance between firm n and its marginal (or, most distant) employee. But only one of these two solutions satisfies the assumption that $1 - \beta d_{nt}$, the labor supplied by the marginal employee, is positive. This solution is

$$d_{nt} = d = \frac{1}{\beta} \cdot \left(1 - \sqrt{1 - \frac{\beta F}{\bar{L}}}\right), \tag{10}$$

a constant for all firms n and for all times t. (We assume $\beta F < \bar{L}$ throughout to ensure a real solution.) This leads us to:

Lemma 2. *There is a distance d such that all workers who live within that distance of a firm – and none that live farther away – will work for that firm in equilibrium at any time t. Moreover, d is increasing in the cost of unit-distance travel, β, and in the minimum efficient scale, F, but decreasing in the density of workers on the unit circle, $\bar{L}$. Moreover, if $\beta F/\bar{L}$ tends to zero, so does d.*

To see why this is so, recall that in free-entry equilibrium each firm will employ F units of labor. It follows that if F increases, each firm will have to hire more employees – after all, $1 - \beta x$, the labor supplied by an employee who lives a distance x from the firm, does not change when F increases. To hire more employees, the firm must hire people from farther away, which implies an increase in d. As for the effect of β on d, an increase in β, the labor required for unit-distance travel, implies that each employee now provides less labor to her employer. Therefore, each employer must hire more employees to meet her unchanged need for F units of labor. Therefore, once again, d must increase. Finally, an increase in $\bar{L}$ implies a higher density of potential employees in every locality (or, unit circle). Therefore, each firm can now meet its need for F units of labor by hiring from a smaller segment of the unit circle. That is, d decreases.

2.7. Symmetric equilibrium

Given that all firms produce the same good, using the same technology, and given that firms and workers are distributed uniformly on the unit circle, it is reasonable to assume – as we will prove later – that *in equilibrium* all firms pay the same wage, w_t.

By substituting $w_{nt} = w_t$ in Eq. (2) we get $d_{nt} = d_t = 1/(2N_t)$. (This is intuitive: In a symmetric equilibrium with N_t firms on the unit circle, the marginal employee must be located exactly halfway between two neighboring firms, which are a distance $1/N_t$ apart.) Eq. (10) then yields

$$N_t = N = \frac{1}{2d} = \frac{\beta}{2} \cdot \left(1 - \sqrt{1 - \frac{\beta F}{\bar{L}}}\right)^{-1}, \tag{11}$$

a constant at all times t. This implies:

Lemma 3. *The equilibrium number of firms, N, is decreasing in β and F and increasing in $\bar{L}$: that is, $dN/d\beta < 0$, $dN/dF < 0$, and $dN/d\bar{L} > 0$. Moreover, as $\beta F/\bar{L}$ tends to zero, N tends to infinity.*

Eqs. (5) and (10) plus the symmetry condition, $w_{nt} = w_t$, then yield

$$\epsilon_{nt} = \epsilon = \frac{\bar{L}}{\beta F} - 1 > 0, \tag{12}$$

a constant for all firms and for all time.[15] We have therefore established:

Proposition 1. *The equilibrium value of the elasticity of labor supply, ϵ, is decreasing in β and F and increasing in $\bar{L}$: that is, $d\epsilon/d\beta < 0$, $d\epsilon/dF < 0$, and $d\epsilon/d\bar{L} > 0$. Moreover, as $\beta F/\bar{L}$ tends to zero, ϵ tends to infinity, reflecting the transformation of the labor market from monopsony to perfect competition.*

---

[15] Recall that we have assumed $\beta F < \bar{L}$ to ensure real solutions for d and N.

Note that $\beta F/\bar{L} > 0$ ensures monopsony in the labor market and $\beta F/\bar{L} = 0$ implies perfect competition. Therefore, the main lesson of Proposition 1 is that the existence of monopsony in our model depends on a combination of commuting costs, dispersed locations of workers and firms, and a minimum efficient scale in production.

The equilibrium wage follows from Eqs. (7) and (12):

$$w_{nt} = w = \frac{\epsilon}{1+\epsilon} = 1 - \frac{\beta F}{\bar{L}} < 1, \tag{13}$$

a constant for all firms and for all time. This implies:

Proposition 2. *The equilibrium wage is decreasing in* β *and* F *and increasing in* $\bar{L}$*:* $dw/d\beta < 0$, $dw/dF < 0$, *and* $dw/d\bar{L} > 0$. *Moreover, if* $\beta F/\bar{L}$ *decreases to zero, the labor market approaches perfect competition and* w *increases to unity, the marginal product of effective labor.*

Note that an increase in the number of workers does not cause wages to fall. Instead, a bigger workforce attracts more firms, and the resulting stiffer competition among firms for workers leads, counterintuitively, to a higher wage.[16] This counterintuitive link between the size of the workforce and the wage is used in the urbanization literature – see Helsley and Strange (1990), Kim (1990), and Duranton and Puga (2003) – to model the formation of cities. As more people migrate to a city, wages rise, thereby causing further migration and further wage increases. This attractive force is balanced by the repulsive force exerted by the costs of urban congestion.

Finally, let us consider the fraction, μ, of each locality's endowment of labor that is employed in production. The free-entry condition, Eq. (9), implies that each of the N firms in the final-good sector employs F units of labor. Therefore, $\mu = NF/\bar{L} = N\bar{L}(2d - \beta d^2)/\bar{L}$. As $N = 1/(2d)$ – see Eq. (11) – we get

$$\mu = \frac{\bar{L}(2d - \beta d^2)}{2d\bar{L}} = 1 - \beta \cdot \frac{d}{2} = \frac{1}{2}\left(1 + \sqrt{1 - \frac{\beta F}{\bar{L}}}\right) < 1, \tag{14}$$

by way of Eq. (10). This implies:

Lemma 4. *The fraction of all labor that is employed in production,* μ*, is decreasing in* β *and* F *and increasing in* $\bar{L}$*: that is,* $d\mu/d\beta < 0$, $d\mu/dF < 0$, *and* $d\mu/d\bar{L} > 0$. *Moreover, as* $\beta F/\bar{L}$ *tends to zero,* μ *tends to unity.*

Note that the effects of β, F and $\bar{L}$ on μ are the same as those on N, the number of firms. This is exactly what one would expect: When the number of firms increases, workers' commuting distances decrease and, therefore, a larger fraction of available labor can be put to productive use.

Having determined the equilibrium wage, w, and the labor utilization rate, μ, we can now focus on each individual's equilibrium wage income. We saw in the second paragraph of Section 2.1 that the distance that worker i must travel between home and work at time t – let us denote this by z_{it} – is a random number (between 0 and d). Therefore, worker i's wage income at time t, $(1 - \beta z_{it})wh_t$, is also random.

We assume, however, that a perfectly competitive insurance market is available for this uncertainty. It can then be checked that at instant t, individual i signs the following location-insurance contract: she pays $\beta \cdot ((d/2) - z_{it}) \cdot wh_t$ to the insurer when $(d/2) - z_{it}$ is positive and is paid that same amount by the insurer when $(d/2) - z_{it}$ is negative. As a result, every worker ends up with $(1 - \beta \cdot (d/2))wh_t \equiv \mu wh_t$ in income from work, irrespective of the length of her commute.[17]

[16] This is a crucial link in the chain of reasoning that leads to the motivating prediction of our empirical analysis in Section 4. Under monopsony, higher population density leads to higher wages, which provides a higher incentive for investment in human capital, which leads to faster growth. The typical intuition of competitive analysis, on the other hand, is that more people means *lower* wages, which leads to a starkly contrasting prediction that higher population density leads to *slower* growth. The contrast between the two approaches is right here, in the relation between population density and the wage.

[17] Our assumptions regarding location in the second paragraph of Section 2.1 imply that z_{it} is a sequence of i.i.d. random variables over time. Therefore, z_{it} is re-drawn at every instant, so that for any time interval of positive Lebesgue measure, there is an infinite number of independent and identically distributed draws of z_{it}. Therefore, for any non-trivial time interval, every worker will earn the same wage income. Therefore, although we have assumed insurance to simplify the exposition, the outcome would be essentially the same without insurance.

Assuming that all workers own the same number of shares in the economy's firms, it follows that the income of every individual at time t is

$$y_t = \mu w h_t + \frac{1}{\bar{L}} \sum_{n=1}^{N} \pi_{nt}, \tag{15}$$

where π_{nt} denotes the profit of firm n at time t.

2.8. Maximization of lifetime utility

Each individual must decide how much of her income to consume and how much to invest in human capital. Every unit of unconsumed income is converted into one unit of additional human capital:

$$\dot{h}_t = y_t - c_t. \tag{16}$$

Formally, we assume that each individual's maximization problem is to choose consumption, c_t, to maximize lifetime utility

$$U = \int_0^{\infty} \mathrm{e}^{-\rho t} \frac{c_t^{1-\sigma} - 1}{1-\sigma} \mathrm{d}t, \tag{17}$$

where $\rho > 0$ is the rate of time preference and $\sigma > 0$ is the reciprocal of the (constant) elasticity of marginal utility, subject to Eqs. (15) and (16), taking w_t as given. (Of course, we also assume $\sigma \neq 1$.)

The present-value Hamiltonian for this maximization problem is

$$\mathscr{H}_t = \mathrm{e}^{-\rho t} \frac{c_t^{1-\sigma} - 1}{1-\sigma} + \lambda_t \cdot \left(\mu w h_t + \frac{1}{\bar{L}} \sum_{n=1}^{N} \pi_{nt} - c_t \right), \tag{18}$$

where λ_t is the discounted shadow price of human capital. The first-order conditions are $\partial \mathscr{H}_t / \partial c_t = 0$ and $\partial \mathscr{H}_t / \partial h_t = -\dot{\lambda}_t$ and, respectively, they yield the equations $\mathrm{e}^{-\rho t} c_t^{-\sigma} = \lambda_t$ and $\lambda_t \mu w = -\dot{\lambda}_t$.

These two equations yield the Euler equation:

$$\rho + \sigma \frac{\dot{c}_t}{c_t} = -\frac{\dot{\lambda}_t}{\lambda_t} = \mu w, \tag{19}$$

which yields the constant *equilibrium growth rate* of consumption:

$$g = \frac{\dot{c}_t}{c_t} = \frac{\mu w - \rho}{\sigma}, \tag{20}$$

where both w and μ have been determined earlier – see Eqs. (13) and (14) – in terms of the parameters of our model.

Recall from Section 2.7 that μ is the fraction of the economy's endowment of labor that is actually used in production. Therefore, it follows from Eq. (1) that per capita income is

$$y_t = \frac{\mu \bar{L} h_t}{\bar{L}} = \mu h_t, \tag{21}$$

which implies that y_t and h_t must grow at the same rate in equilibrium: $\dot{y}_t / y_t = \dot{h}_t / h_t$. Further, it can be shown that the unique equilibrium of our model is also the steady state equilibrium in which h_t grows at the same rate as c_t (a formal proof is available upon request). Consequently, we conclude that human capital, output, and consumption, all grow at the rate g, which, therefore, may justifiably be called *the* equilibrium growth rate of the economy.

We saw in Section 2.7 that if there is a decrease in either β or F or an increase in $\bar{L}$, then both w and μ increase. Eq. (20) then implies:

Proposition 3. *The equilibrium growth rate of per capita consumption, per capita output, and human capital is decreasing in β and F but increasing in $\bar{L}$: that is, $\partial g / \partial \beta < 0$, $\partial g / \partial F < 0$, and $\partial g / \partial \bar{L} > 0$. Moreover, this growth rate is inversely related to both the rate of time preference and the elasticity of marginal utility:*

that is, $\partial g/\partial\rho < 0$ *and* $\partial g/\partial\sigma < 0$*. As* $\beta F/\bar{L}$ *tends to zero, g tends to* $(1-\rho)/\sigma$*, its highest, competitive level.*[18]

The effects of β, F, and $\bar{L}$ on the growth rate, g, work through two channels. To see this, consider an increase in $\bar{L}$. (The effects of β and F work analogously.) First, an increase in $\bar{L}$, the population density of a typical locality (or, self-contained labor market), leads to an increase in N, the number of firms on the unit circle. As a result, workers find jobs closer to home and, therefore, spend less labor on home-to-work travel. This leads to an increase in μ, the fraction of all labor that is used in production. Moreover, the increase in N leads to greater competition among firms over workers, which leads to an increase in w, the wage.[19] And, as is clear from Eq. (20), these increases in μ and w boost the growth rate.

Note also that while $\partial g/\partial\bar{L} > 0$ does imply that countries with higher worker densities in labor markets will grow faster, it does not imply that countries with higher populations will grow faster. Recall from Section 2.1 that we have assumed that each country is composed of multiple localities that function as self-contained and island-like labor markets. We have also assumed that increases in population lead to proportionate increases in the number of localities and, therefore, have no effect on population density. In this way, we avoid all empirically dubious scale effects. Moreover, our model of the labor market is not intended to describe rural or agricultural labor markets. Therefore, $\bar{L}$ is best defined as the density of workers in urban labor markets, and it is this density that is positively related to growth in our model.

Now, the *share of output that is invested in education* in equilibrium is $s_e \equiv \dot{h}_t/y_t \equiv (\dot{h}_t/h_t)\cdot(h_t/y_t)$. By Eqs. (20) and (21), we get

$$s_e = \frac{g}{\mu} = \frac{1}{\sigma}\cdot\left[w - \frac{\rho}{\mu}\right]. \tag{22}$$

As a result, Proposition 2 and Lemma 4 then lead us to the following:

Proposition 4. *The share of income that is invested in education is decreasing in* β *and F but increasing in* $\bar{L}$*: that is,* $ds_e/d\beta < 0$, $ds_e/dF < 0$*, and* $ds_e/d\bar{L} > 0$*.*

Note that the effects of β, F, and $\bar{L}$ on investment in education (s_e) have the same sign as their effects on the number of firms (N), the elasticity of each firm's labor supply (ϵ), the wage rate (w), the utilization of labor (μ), and the growth rate (g).

Also, recall, from Eq. (15), that per capita wage income at time t is $\mu w h_t$ and, from Eq. (21), that per capita output is μh_t. It follows that the *share of wage income in output* is $\mu w h_t/\mu h_t$, which is nothing but the wage, w. Our model therefore predicts that *the wage share of output, the share of output that is invested in education, and the economy's growth rate, must all move in the same direction, either up or down*. We test this prediction in Section 4.

The Euler equation, Eq. (19), represents the equalization of $\rho + \sigma g$, the permanent instantaneous cost of the deferred gratification that is necessary for the accumulation of an additional unit of human capital, and μw, the permanent instantaneous benefit from an additional unit of human capital. The only reward for thrift in our model is μw, which is the additional wage income associated with additional human capital. Monopsony, by reducing the wage, reduces the only incentive that people have to accumulate human capital. This reduces the growth rate.

2.9. Physical capital

Although we have excluded physical capital from our model to make its logic more transparent, we wish to emphasize that the inclusion of physical capital does not affect our results in any substantial way (though it does introduce transitional dynamics, whereas the economy above is always in steady state). Formally, we state the following proposition, whose proof is available upon request:

[18] *This is because, as we saw from Lemma 4 and Proposition 2,* μ *and w both tend to unity when* $\beta F/\bar{L}$ *tends to zero.*

[19] See Duranton and Puga (2003, pp. 24–25).

Proposition 5. *Consider an economy that is identical to the one discussed so far, except that (a) the production function,* Eq. (1), *is replaced by*

$$Y_{nt} = \begin{cases} AK_{nt}^{\alpha}(h_t L_{nt})^{1-\alpha}, & \text{if } K_{nt} \geqslant 0 \text{ and } L_{nt} \geqslant F, \\ 0, & \text{otherwise,} \end{cases}$$

where K_{nt} is the amount of physical capital rented by firm n at time t at the competitive real rental rate, r_t, and (b) the accumulation equation – Eq. (16)*: $\dot{h}_t = y_t - c_t$ – is replaced by its natural counterpart, $\dot{h}_t + \dot{k}_t = y_t - c_t$, where k_t is the per capita stock of physical capital at time t. In this economy, the effect of β, F, and $\bar{L}$ on the steady state equilibrium growth rate is unchanged: $dg/d\beta < 0$, $dg/dF < 0$, and $dg/d\bar{L} > 0$, exactly as in Proposition 3.*

In this case, the new Euler equation retains Eq. (19) but further requires that $\rho + \sigma g = r$ must also hold, where r is the steady state real rental rate of physical capital.

2.10. Short-lived firms

We now return to our assumption that the firms in any locality (or, unit circle) are continuously and randomly relocated; see Section 2.1 and Footnote 8. As this assumption implies that a firm stays at a given location for just one instant, it may seem unduly restrictive. However, we have developed a separate model (available upon request) with infinitely-lived workers *and* infinitely-lived firms with permanent addresses on the unit circle, and that model predicts, just as this section's model does, that an increase in population density ($\bar{L}$) leads to higher wages, to greater investment in education, and to faster growth. In other words, the predictions that motivate our empirical analysis in Section 4 are not dependent on our assumption of short-lived (or continuously and randomly relocated) firms. On the plus side, the model in the present paper is a lot clearer than the model with infinitely-lived firms.

Also, it is well-known that the assumption of infinitely-lived households in neoclassical growth theory is not meant to be taken literally. In the interpretation of Barro and Sala-i-Martin (2005; p. 86), "Each household contains one or more adult, working members of the current generation. In making plans, these adults take account of the welfare and resources of their prospective descendants. We model this intergenerational interaction by imagining that the current generation maximizes utility and incorporates a budget constraint over an infinite horizon. ...The immortal family corresponds to finite-lived individuals who are connected through a pattern of operative intergenerational transfers based on altruism." As no lower bound is postulated for the longevity of a typical generation, each instant in the life of the immortal family can represent one generation, and a firm that lives for an instant can be thought of as a firm that lives as long as a typical generation.[20]

Using infinitely-lived firms does require a simplification: In Section 2.3 we suppose that workers pay for commuting with a shorter workday. If worker and firm positions were fixed, this would imply that workers further away from the firm would permanently work fewer hours, giving them a lower perceived return to human capital and therefore a slower rate of productivity growth. Since this outcome may seem extreme, our model with infinitely-lived firms assumes that the time spent by workers on home-to-work travel cuts into one's leisure but does not reduce one's remunerated labor. In other words, a worker close to the firm and one far from the firm work the same hours, earn the same incomes, and have the same return from education. The only difference is that the worker located far from the firm enjoys less leisure. Any two workers are indistinguishable in terms of their economic behavior. This sameness makes the model a representative agent model that can be analyzed in the usual manner.

To sum up, the links between urban structure, monopsony, wages, return to education, investment in education, and economic growth can be established with or without the assumption of short-lived firms.

[20] Moreover, as was mentioned in Footnote 8, the results of this section are also true in a model with stationary *firms* and continuously and randomly relocated *workers*, both infinitely-lived. Assuming that a moment in time represents a generation, our results, therefore, apply to a world in which firms are infinitely-lived and stationary but each generation of a family finds its own new place to live.

3. Equilibrium and optimum

We now compare the decentralized equilibrium outcome discussed above with the Pareto optimal solution to the social planner's problem.

3.1. The optimum outcome

At every instant t, the social planner must decide: (1) how many firms there should be in the final-good sector in each locality (or, unit circle), (2) where should each firm be located on the unit circle, and (3) of the $\bar{L}$ workers on each unit circle, which worker should work for which firm. In order to minimize the labor spent on travel, the social planner would want N_t, the number of firms, to be large. However, N_t must not be so large that some firms end up with less than F units of labor, which is the minimum labor a firm needs to be productive; see Eq. (1).

These considerations – it is straightforward to show – imply (i) that for each locality the social planner will place all N_t firms uniformly on the unit circle, each a distance $1/N_t$ away from its nearest neighbors; (ii) all workers that live within a distance $1/2N_t$ of a firm will be employed in that firm; and (iii) N_t will be such that each firm will employ exactly F units of labor, the minimum labor needed for production.

Note, however, that these are *exactly* the properties of the decentralized market outcome of Section 2. It follows that the social planner will choose $N_t = N^* = N$, where N is given by Eq. (11). Consequently, the fraction of the economy's endowment of labor that is actually used to produce the final good will be $\mu_t = \mu^* = NF/\bar{L} = \mu$, where μ is exactly as it was in Eq. (14) for the decentralized market economy.

Therefore, the per capita output in the command economy is once again given by Eq. (21). The *social planner's optimization problem* is then to choose c_t to maximize the lifetime utility of the representative individual – U in Eq. (17) – subject to Eqs. (16) and (21). By setting up the Hamiltonian as before – $\mathscr{H}_t = e^{-\rho t}(c_t^{1-\sigma} - 1)/(1-\sigma) + \lambda_t \cdot (\mu h_t - c_t)$ – and working out the first-order conditions, we get the social planner's Euler equation:

$$\rho + \sigma\frac{\dot{c}_t}{c_t} = -\frac{\dot{\lambda}_t}{\lambda_t} = \mu. \tag{23}$$

As before, it can be shown that in the optimum outcome c_t, h_t, and y_t all grow at the same rate. This growth rate is, therefore,

$$g^* = \frac{\dot{c}_t}{c_t} = \frac{\mu - \rho}{\sigma}. \tag{24}$$

The comparative static effects of β, F, $\bar{L}$, ρ, and σ on $N^*(=N)$, $\mu^*(=\mu)$, and g^* follow in a straightforward manner from Eqs. (11), (14), and (24).

3.2. Optimum and equilibrium compared

As the equilibrium wage is $w < 1$ – see Eq. (13)– we see that *the equilibrium growth rate under monopsony is lower than the optimal growth rate*: $g = (\mu w - \rho)/\sigma < (\mu - \rho)/\sigma = g^*$. Monopsony – which, recall, exists when $\beta F/\bar{L} > 0$ – keeps the equilibrium wage below unity, which is the marginal product of (effective) labor; see Eq. (1). The representative individual in the decentralized economy earns $\mu w h_t$, in wage income, plus profits, which are taken as given. Therefore, every additional unit of human capital at time t is perceived to yield an additional income of μw at every subsequent moment. The social planner, on the other hand, sees per capita output as given by Eq. (21) and understands that the return to human capital is actually $\mu > \mu w$. In this way, *monopsony, by keeping the wage below the marginal product, causes the representative individual in the decentralized economy to have a suboptimal incentive to accumulate human capital.* As a result, *an economy characterized by a monopsonistic labor market grows at a suboptimal rate.*

It is also straightforward to show – from Eqs. (21) to (24)– that the *optimal share of education spending in total output* is $s_e^* \equiv \dot{h}_t/y_t \equiv (\dot{h}_t/h_t) \cdot (h_t/y_t) = g^*/\mu$, and that another consequence of monopsony is that *the equilibrium level of investment in education under monopsony is suboptimal*: $s_e^* = g^*/\mu > g/\mu = s_e$, because $g^* > g$.

Summing up this section's results, we state the following

Proposition 6. *Under monopsony, the equilibrium growth rate is lower than the optimal growth rate* $(g < g^*)$*, and the equilibrium share of education spending in output is lower than the optimal share* $(s_e < s_e^*)$*. However, the equilibrium number of firms is equal to the optimal number of firms* $(N = N^*)$.

Finally, recall from Proposition 2 that as $\beta F/\bar{L}$ tends to zero, the wage, w, tends to unity, which is the marginal product of effective labor. Moreover, as is clear from Eq. (14), μ approaches unity as $\beta F/\bar{L}$ tends to zero. Therefore, *as monopsonistic competition in the labor market turns into perfect competition, the equilibrium rate of investment in education and the equilibrium rate of growth both increase to their respective optimal values.*

4. Empirical results

Our theory suggests three testable hypotheses. We will list them in the order that they appear in our model, starting with the labor market and moving through to educational attainment, and finally the growth rate. First, monopsony power should be weaker and hence the wage share of output should be larger in economies with denser labor markets, both because of lower travel costs and potentially also because of more extensive social networks. Second, investment in education should be higher in economies where there is a larger wage share of output. Our model also suggests that the rate of growth should be higher when there is more investment in education; this particular insight is well-known (see Lucas, 1988), and has been substantiated empirically by Barro and Lee (1993); however, we will add to such results by placing them in the context of how the wage share affects the level of investment in human capital.

Finally, our model suggests that countries with a higher wage share of output should grow faster. It is well-known that real wages are higher during the peaks of business cycles (Dunlop, 1938; Tarshis, 1939), and to the extent that growth is generated by business cycle peaks, estimates of the relationship between wages and growth will pick up this effect. However, we think that estimating the relationship between the wage share and growth is of a broader interest for two reasons. First, the Dunlop–Tarshis result is an empirical one, not a theoretical one. It has been most frequently explained as a result of imperfect product market competition (see Rotemberg and Woodford, 1990); however, to the extent that short bouts of growth may produce cyclical peaks, our model may also provide an explanation for this phenomenon. Second, between-country cyclical effects should be less pronounced if we compare observations of countries over several years, during which they pass through both peaks and troughs of the business cycle.

None of the above hypotheses, if confirmed, would provide definitive evidence in favor of our model, as any of them could potentially result from other influences on the economy (e.g., the demand for education could be higher in more urbanized settings simply because the skills provided by education are more productive in urban jobs). Nevertheless, we would like to at least observe whether the model is generally consistent with some basic empirical observations; hopefully these observations will also have at least some implications for theory. We offer simple tests of our hypotheses that could likely be refined if more elaborate data were available, but which we think are at least sufficient to suggest whether or not the model is plausible. To conduct the tests, we formed a panel of country-level data from four sources. We use data on the investment and government shares of output, GDP per capita, and growth of GDP per capita from Version 6.1 of Heston et al. (2002) [21]; data on the fraction of the population over 25 that had attained a secondary school education, broken down by gender, from Barro

[21] The variables are *ci*, *cg*, *rgdpch*, and *grgdpch*, respectively; our model would be most ideally tested by examining the impact of wages and schooling on growth per worker; however, doing so would cut our sample size in half. We have therefore stuck with the conventional per capita measure of GDP growth.

and Lee (2000)[22]; data on population density and the percent of a country's surface that is permanent cropland from World Bank (2004); and data on the employee compensation, operating surplus of private unincorporated enterprises (OSPUE), indirect taxes, and current GDP from United Nations (1985, 1990, 1995, and 1997).[23]

While there was little difficulty obtaining data for a given country and year on GDP, and its consumption, investment, and wage shares, other variables did substantially restrict the size of our sample. First, our data on the wage share of output needed to be adjusted to account for the self-employed (see Gollin, 2002 and Bernanke and Gürkaynak, 1993). The measure for the wage share that we use is

$$\frac{\text{Employee compensation}}{\text{GDP} - \text{Indirect taxes} - \text{OSPUE}}.$$

In other words, we measure corporate employee compensation as a fraction of corporate income. Conventionally, OSPUE is considered at least partly labor income, since it reflects the income of the self-employed, which, especially in developing countries with large informal sectors, largely results from their labor expenditure. However, monopsony power only directly affects the wage share through its impact on the wages of corporate employees; the appropriate measure of its effect, therefore, should be *the wage share of output within the corporate sector*, which we have measured above. In this way our measure is also immune to the problem that the corporate sector comprises a smaller share of output in poorer countries. Data on OSPUE were available for a much smaller sample of countries than those for which national income statistics were available; most observations of OSPUE date between 1986 and 1996, although we have observations in our sample dating as far back as 1975. Because the year-to-year variation in the data is quite noisy (especially the measure of indirect taxes), we used a five-year unweighted moving average of the variable. The second variable that heavily restricted our sample was the fraction of the labor force with a primary and/or secondary education, which was only available for every fifth year; this time, we imputed intermediate values using a moving average.[24]

Finally, we restricted our sample to observations in which GDP, GDP growth, education, and urban population variables were non-missing, so that differences in sample selectivity would not affect the results when variables were added to our regressions. Additionally, because the level of net indirect taxes has some outlier observations, it is possible for the computed wage share to be less than zero or greater than one. Not wishing to have our results driven by outliers, we restricted our sample to wage shares between 0.4 and 0.8. We are left with a sample of 369 observations from 32 countries.[25] Table 1 gives means for all of the variables used for each sample.

Our model suggests that the variables under consideration are determined simultaneously. We take a country's population density and cropland area as exogenous; some may object that fertility and migration are determined by economic circumstances, however our view is that these variables were largely determined before the process of industrialization began. We have attempted to mitigate any endogeneity problems by considering the effects of these variables at a lag, and, in some specifications, by using differences-in-differences estimates. Given the exogenous variables, we have the fol-

[22] The percent attaining at least a secondary education was that attaining a secondary education plus that attaining a tertiary education. Because these data were only available for five-year intervals, we used a moving average to interpolate values for other years.

[23] Observations for 1986 through 1996 were taken from the 1997 Yearbook; for 1984–85 from the 1995 Yearbook; for 1980–1984 from the 1990 Yearbook; and for 1975–1979 from the 1985 Yearbook. Data on employee compensation, indirect taxes, and GDP were taken from Table 1.3. This measure of current GDP was used only to calculate the wage share measures. Data on OSPUE were taken from Table 1.3 where available. For some countries, OSPUE was not available in Table 1.3 but was available for resident households in Table 1.6. Examination of countries where OSPUE data were available in both tables suggested that the differences between these two measures (essentially the income of unincorporated travelling salesmen and international consultants) were negligible, and therefore it was legitimate to substitute the OSPUE data from Table 1.6 where they were not available in Table 1.3.

[24] Data on education were also available from World Bank (2004), but contained even fewer observations in the relevant time period than that from Barro and Lee (2000).

[25] The countries in our final sample are Australia, Belgium, Botswana, Bulgaria, Canada, Colombia, Congo (Republic), Czech Republic, Denmark, Ecuador, Finland, France, Hungary, Italy, Jamaica, Japan, Korea (South), Mauritius, Netherlands, New Zealand, Peru, Philippines, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Thailand, Tunisia, the United Kingdom, and the United States.

Table 1
Means of selected variables

Variable	Mean	Std. Dev.	Min.	Max.	No. Obs.
GDP/k	9.3446	0.7540	7.4434	10.4131	369
Per capita growth	2.2028	3.5790	−16.2883	11.2552	369
Wage share (Lag 5)	0.6629	0.0827	0.4020	0.7944	369
Women's education (Lag 5)	31.0810	19.9674	2.4200	93.0000	369
Men's education (Lag 5)	48.7022	21.3142	5.5340	92.3633	369
I/Y (Lag 5)	22.3639	7.2161	3.0337	42.6151	369
G/Y (Lag 5)	14.1701	7.1364	3.0603	46.6381	369
Pct perm cropland (Lag 10)	2.9734	3.8828	0.0035	14.7567	369
ln Popn density (Lag 10)	4.0038	1.6412	0.3775	6.2552	369
Percent urban (Lag 10)	64.4608	18.2434	10.2000	96.4000	369

Note: Education, wage share, investment, and government spending variables reported at a five-year lag from GDP variables. Population density and cropland are reported at a 10-year lag.

lowing simultaneous hypotheses: (1) a denser population with fewer agricultural prospects, will be more urban; (2) a more urban country, because of lower travel costs and higher social network density, will have a higher wage share of output; (3) the wage share of output, in turn, influences educational attainment; (4) educational attainment, in turn, influences the rate of growth. We therefore have a four-equation system:

$$\text{Urban Density} = \alpha_1 + \beta_1 \cdot \text{Topography} \tag{25}$$
$$\text{Wage Share} = \alpha_2 + \beta_2 \cdot \text{Urban Density} \tag{26}$$
$$\text{Education} = \alpha_3 + \beta_3 \cdot \text{Wageshare} \tag{27}$$
$$\text{Growth} = \alpha_4 + \beta_4 \cdot \text{Education} \tag{28}$$

If each of the variables in the system above is taken to represent one variable in the dataset, then the system is just identified. In practice, we wish to overidentify the system a bit so that we can add the wage share, e.g., to the fourth equation and see what happens. We have therefore used two variables (population density and cropland density) as instruments for urban density. Finally, we have measured education (specifically, the fraction of those over 25 with at least a secondary education) separately for men and women. The third equation above therefore in practice consists of two equations.

Table 2 estimates the system above using three-stage least squares, except that in the last equation we merely look at GDP per capita instead of growth as a beginning benchmark. All regressions in all of our tables include year dummies, though we do not report their coefficients. The first three equations of our system (which comprise four equations in practice since we estimate men's and women's education separately) are as predicted by the model above: population density strongly has a positive and statistically significant effect on the percentage of a country that is urban, while cropland density has a negative and statistically significant effect; more urban countries have a higher wage share, and the wage share, in turn, has a positive and significant effect on the level of education (both at the 1% level). The final three columns show regressions of GDP per capita on the wage share and education. As can be seen, higher-income countries have a larger wage share of output; note that we are just looking at the wage share within the corporate sector, so that this controls for the fact that the wage share is higher in countries with larger corporate sectors. This suggests that the wage share of higher-income countries is larger for more than just the reasons discussed by Gollin (2002).

Note also that the relationship between the wage share and education is not predicted by most standard growth models, either of the neoclassical or endogenous variety. A standard Cobb–Douglas production function would predict that more educated workforces will have a higher wage *level*, but will still have the same wage *share*. This point is worth emphasizing a bit, since it would suggest rethinking the cross-country determinants of human capital accumulation if widely replicated. Consider, e.g., an endogenous growth model (of the Uzawa-Lucas type) with both physical and human capital, such as

Table 2
3SLS Regressions of Eqs. (25)–(28) with per capita income as dependent variable in Eq. (27)

	Pct Urban (Lag 10)	Wage share (Lag 5)	Men's education (Lag 5)	Women's education (Lag 5)	(1)-GDP/k	(2)-GDP/k	(3)-GDP/k	(4)-GDP/k
Wage share (Lag 5)					7.1771***			6.0698***
					(0.4078)			(0.4618)
Men's education (Lag 5)						0.0541***	0.0424***	0.0186***
						(0.0085)	(0.0029)	(0.0070)
Women's education (Lag 5)						−0.0152		−0.0083
						(0.0112)		(0.0086)
I/Y (Lag 5)								
G/Y (Lag 5)								
Wage share (Lag 10)			32.0157***	41.9870***				
			(11.0632)	(11.8623)				
Pct urban (Lag 10)		0.0066***						
		(0.0003)						
ln Popn density (Lag 10)	1.3919***		−0.3265	−2.3282***				
	(0.4891)		(0.5353)	(0.5666)				
Pct perm cropland (Lag 10)	−1.3694***		−1.7040***	−0.8527***				
	(0.2152)		(0.2317)	(0.2446)				
Obs	369					369	369	369

Note: Regressions contain year dummies. Standard errors in parentheses. Asymptotic two-tailed significance levels: * 10%; ** 5%; *** 1%. Equations with wage share and education as dependent variables are estimated for all systems but only reported for the first.

$$y_t = Ak_t^{\alpha}[l_t h_t(1-u_t)]^{1-\alpha}$$

where

$$\dot{k}_t = y_t - \delta k_t - c_t$$
$$\dot{h}_t = b h_t u_t$$

and consumers maximize the utility function of Eq. (17).[26] The point is that while it is well known that such a model features endogenous growth due in part to constant returns in the accumulation of human capital, the marginal product of labor, which is equal to the wage under perfectly competitive labor markets, will equal $\partial y_t/\partial l_t = A(1-\alpha)k_t^{\alpha} l_t^{-\alpha}[(1-u_t)h_t]^{1-\alpha}$. The total wage bill is then $l_t \partial y_t/\partial l_t$, which is increasing in h_t but is nevertheless a constant *share* of output: $(l_t \partial y_t/\partial l_t)/y_t = 1-\alpha$. Such a model can only explain a higher wage share in countries with more education if the cross-country wage share differences are driven by differences in α. While this is not implausible *per se*, it is not our perception that economies that specialize in activities with a relatively lower return to physical capital tend to accumulate more human capital. In our model, on the other hand, *the wage share can vary due to monopsony power even when α does not*.

As shown by the last three columns of the table, either men's education or women's education has a positive and significant effect when included by itself (only men's education is shown), but when

[26] Here, k_t and l_t are capital and labor, respectively; h_t is human capital, δ is the depreciation rate of capital, b is the productivity of education per unit of teacher human capital, and u_t is the fraction of time the representative worker spends in school rather than working. See equations (5.20) and (5.21) in Barro and Sala-i-Martin (2004).

Table 3
3SLS Regressions of Eqs. (25)–(28)

	Pct urban (Lag 10)	Wage share (Lag 5)	Men's education (Lag 5)	Women's education (Lag 5)	(1) – Growth/k	(2) – Growth/k	(3) – Growth/k
GDP/k (Lag 5)					−0.9938**	0.0919	−1.1638
					(0.4804)	(0.7272)	(1.3728)
Wage share (Lag 5)					7.5124*		7.3121
					(4.4451)		(6.3375)
Men's education (Lag 5)						0.0430	0.0646
						(0.0532)	(0.0573)
Women's education (Lag 5)						−0.0961*	−0.0806
						(0.0568)	(0.0575)
I/Y (Lag 5)					0.0796**	0.0387	0.0647*
					(0.0331)	(0.0288)	(0.0362)
G/Y (Lag 5)					−0.0882***	−0.0795***	−0.0780***
					(0.0297)	(0.0298)	(0.0299)
Wage share (Lag 10)			30.3566***	38.4432***			
			(11.2254)	(12.0131)			
Pct urban (Lag 10)		0.0056***					
		(0.0002)					
ln Popn density (Lag 10)	1.1054**		−0.1529	−2.1318***			
	(0.5227)		(0.5880)	(0.6087)			
Pct perm cropland (Lag 10)	−1.6305***		−2.0329***	−1.0879***			
	(0.2247)		(0.2545)	(0.2630)			
Obs.	369					369	369

Note: Regressions contain year dummies. Equations with wage share and education as dependent variables are estimated for all systems but only reported for the first. Standard errors in parentheses. Asymptotic two-tailed significance levels: *10%; **5%; ***1%.

included together, the coefficients take opposite signs. One might read the coefficients as saying that women's education is the better proxy for the country's overall level of human capital, and that given this level, countries with higher (i.e., more unequal) levels of men's education do worse. The final column of Table 2 shows that the wage share and education variables become insignificant when included together, suggesting that they are proxies for the same effect. While the standard errors are too large for us to draw any strong conclusions, this does suggest that education levels may play a large part in the higher wage shares of industrialized economies, which is worth separate investigation.[27]

Table 3 repeats the system of equations in Table 2, but this time with the real per capita growth rate as the dependent variable in the final equation. We have attempted to follow the overall form of Barro and Lee (1993), Table 5, as a benchmark case, with a couple of modifications. Barro and Lee regress the rate of growth on 10-year lagged values of GDP per capita, male and female secondary school completion rates, the investment and government shares of output, life expectancy, the black market premium on foreign exchange, and anticipated political instability. They find that higher past GDP predicts lower growth (which they consider as showing convergence conditional on the level of

[27] Note that this is a comparative static result, and in our framework, this could result from more developed countries having more transportation and communication infrastructure, which could make labor more mobile and therefore reduce (but not eliminate) monopsony power.

Table 4
3SLS Regressions of Eqs. (25)–(28) with Fixed Effects

	Pct urban	Wage share	Men's education	Women's education	(1) – Growth/k	(2) – Growth/k	(3) – Growth/k
	(Lag 10)	(Lag 5)	(Lag 5)	(Lag 5)			
GDP/k (Lag 5)					−1.9748***	−0.7502	−3.4306
					(0.5963)	(0.8596)	(11.6710)
Wage share (Lag 5)					16.7017***		18.9930
					(5.1498)		(78.9913)
Men's education (Lag 5)						0.1730*	0.2315
						(0.0887)	(0.3281)
Women's education (Lag 5)						−0.5895***	−0.2917
						(0.1672)	(1.1228)
I/Y (Lag 5)					0.0541	0.1966*	0.0063
					(0.0356)	(0.1087)	(0.8018)
G/Y (Lag 5)					−0.1135***	−0.2382***	−0.0689
					(0.0337)	(0.0859)	(0.7168)
Wage Share (Lag 10)			−36.4570***	−11.8547			
			(10.6936)	(11.9663)			
Pct Urban (Lag 10)		0.0052***					
		(0.0002)					
ln Popn density (Lag 10)	1.8837***		1.5513***	0.4015			
	(0.4921)		(0.5121)	(0.5459)			
Pct perm cropland (Lag 10)	−1.8236***		−1.8549***	−0.6684***			
	(0.2113)		(0.2227)	(0.2364)			
Obs.	369					369	369

Note: Regressions contain year dummies and country dummies. Equations with wage share and education as dependent variables are estimated for all systems but only reported for the first. Standard errors in parentheses. Asymptotic two-tailed significance levels: * 10%; ** 5%; *** 1%.

human capital); that male secondary school attainment predicts higher growth but, controlling for this, female secondary school attainment predicts lower growth (which they attribute to an effect similar to the one we discussed about Table 2 above); that past investment spending (I/Y) increases growth, but past government spending (G/Y) decreases it; that life expectancy increases growth; and that black market premia and political instability decrease it.

While attempting to preserve this overall framework for comparability, we have modified it in a few ways. First, because our wage share variable is available for a narrower time frame than the other variables, we consider the independent variables at five and ten year lags rather than ten and twenty year lags. Second, Barro and Lee use lagged endogenous variables as instruments, whereas we use other exogenous variables (specifically, population and cropland density) as motivated by our model above. Finally, we do think it important to include the investment, government spending, and lagged GDP variables to control for business cycle effects. But because our model has a narrower focus, we do not consider the effects of political instability, life expectancy, and the black market premium on foreign exchange rates.

We are able to replicate all of Barro and Lee's coefficients for the relevant variables when we use a full sample of countries (this result is not shown but is available on request); however, the availability of the wage share data substantially reduces the size of our sample, and so the coefficient on male education is generally insignificant in this subsample and not even positive for all of our specifications. The overall effect of the wage share is significant, and loses its significance (though only about a third of its magnitude) when education is included. Curiously, the sample only displays the conditional convergence found by Barro and Lee when the wage share is included.[28] Note also that the effect of the wage share is controlling for the level of GDP per capita, meaning that it implies that countries with a wage share that is uncharacteristically high for their level of development will grow faster. The positive coefficient for lagged investment share and the negative coefficient for lagged government spending share are as found in the literature.

Table 4 includes country fixed effects, making it a differences-in-differences estimate, which allow us to ask the question of which factors might move a country from a low-growth to a high-growth economy. If there is not much within-country movement of this sort, then adding fixed effects will make the data more noisy, and indeed we can see that the effect of the wage share on education loses its significance, as do the effects of past government and investment spending. In any event, the results do not change much from Table 3, though they are slightly more in line with standard predictions. This time, men's education is a significant predictor of economic growth when included by itself, and the effect of the wage share not only becomes an insignificant predictor of growth when education is added, but substantially falls in magnitude (the education variables, on the other hand, lose significance in predicting growth when the wage share is added to the right-hand side, but do not change their magnitude much).

Overall, then, we think the data do not suggest any serious inconsistencies between our model and the basic observed relationships between population density, wages, human capital accumulation, and growth. Population density increases the wage share. A higher wage share significantly increases education of both men and women. High income countries have larger wage shares, and this effect is partially collinear with the effect of education on income. Finally, the wage share has a positive and significant effect on the growth rate, and this effect overlaps heavily with the effect of education on growth.

5. Conclusion

Factor markets, just like product markets, have imperfections. To the extent that these imperfections alter the rate of investment, whether in human capital or physical capital, they may alter the rate of growth. We have identified a very specific mechanism whereby monopsony power drives the wage below the marginal product of labor and thereby distorts the worker's incentive to invest in human capital so that it does not reflect the true productivity effects of such investment. We have provided evidence that this mechanism at least plausibly exists: First, the wage share of output is positively and significantly impacted by population density, which, as we interpret it, will lower workers' travel costs and undercut firms' monopsony power. Second, the level of education is positively and significantly impacted by a change in the wage share of output; this is a result that is predicted by our model and inconsistent with many others, so while we do not view it as definitive supporting evidence, it does suggest an empirical puzzle that ought to be studied more. Third, high income countries have higher wage shares of output, and this relationship is a proxy for the effect of education on output (or vice versa). Finally, the wage share of output positively predicts growth, and differences-in-differences estimates show that countries that increase their wage share grow faster subsequently.

Both the theory and the evidence beg other questions that would be useful for future research. On a theoretical level, it would be interesting to know the growth dynamics of more complex labor markets. For example, we have only examined the effect of monopsony on an economy with a single monopsonistic sector; on the other hand, monopsony power may also affect the rate of employment in non-monopsonistic sectors (for example, the rate of self-employment or participation in partnerships), which may have altogether different effects on growth. Tse (2000) has shown the way in which a sector with a monopolistic product market may impact output in a two-sector model, and Aghion and Schankerman (2004) have shown the importance of uneven firm productivity in determining the rate of investment. While the simplicity of our one-sector model has its appeal, it would also be

[28] Barro and Lee find a negative lagged coefficient for GDP per capita, which they interpret as showing that a country with high output relative to its human capital stock should subsequently converge to the output level consistent with its human capital stock, i.e., it should grow slowly.

useful to know how monopsony power affects growth when its extent is uneven. Moreover, Tse's model assumes that the self-employment sector, where there is no market power, has low human capital requirements, while the endogenous growth models of Romer (1990) and others assume that self-employment is entrepreneurial and inventive. These two views need to be reconciled in a multi-sector model.

Empirically, our estimates were meant to test the overall plausibility of our model; more exhaustive estimates would be useful. Our principle difficulty is in the small available sample size, which mainly results from the small number of country-year cells for which careful estimates of the wage share of output are available. Because these are historical data, the only real option to expand the dataset is to use a more crude estimate of the wage share, which many would find unappealing. It would also be helpful if we had a more direct cross-country estimate of monopsony power; one possible solution would be to use the estimator suggested by Ridder and van den Berg (2003), but the variables required for this estimator would further restrict our data set substantially, and the resulting data would not overlap well with the available measures of the wage share.

In future empirical work, the effect of monopsony power on the accumulation of skills ought to be examined in more detail. For example, workers of different skill levels may be subject to different amounts of monopsony, and the extent to which their rate of investment in human capital is distorted will therefore also be differently affected. This, in turn, may not only affect the overall rate of growth, but the types of technologies that are developed by an economy. A second, related point is that monopsony power may be stronger or have different effects in different industries, and economies with larger wage shares may grow faster in some industries but not in others.

More research is also needed into the cross-country determinants of monopsony power. Population density has a demonstrable effect on the wage share of output, which in itself is interesting, but begs the question exactly why and how. Is it that workers in more densely populated areas simply have lower travel costs, or are they better able to share information with each other about alternative employment, or are there perhaps skill interactions between urban workers, or possibly cultural norms, that somehow give them more bargaining power with their employers? The data at our disposal do not enable us to answer these questions.

Finally, our mechanism needs to be disentangled from others that might produce similar effects. Imperfect competition in product markets, for example, may have an effect on the wage share of output. Moreover, the dynamics of economic growth are entangled both theoretically and empirically with the dynamics of the business cycle. The degree to which fast-growing economies are experiencing a cyclical peak, which may (for reasons unspecified) exhibit a high wage share, needs to be distinguished from a mechanism that runs in the reverse direction, whereby a high wage share causes economies to grow faster and may even produce short-term cyclical effects. While our data use lags that are longer than a typical business cycle, there is insufficient theoretical distinction between a long-term growth spurt and an extra-long business cycle for us to know how to distinguish the effects empirically. We think that these questions together provide ample material for future research, and we hope that the present work has provided sufficient evidence that they are questions worth asking.

Acknowledgement

We would like to thank James Schmitz, Caroline Hoxby, Kevin Bowman, and an anonymous referee for valuable comments and suggestions. The usual caveat applies.

References

Aghion, P., Howitt, P., 1997. Endogenous Growth Theory. The MIT Press, Cambridge, MA.
Aghion, P., Schankerman, M., 2004. On the welfare effects and political economy of competition-enhancing policies. Economic Journal 114, 800–824.
Albrecht, J.W., Jovanovic, B., 1986. The efficiency of search under competition and monopsony. Journal of Political Economy 94, 1246–1257.
Arrow, K.J., 1962. The economic implications of learning by doing. Review of Economic Studies 75, 155–173.
Barr, T., 2002. Monopsony Power and the Returns to Skill. Columbia University Ph.D. Dissertation.
Barro, R.J., 1990. Government spending in a simple model of endogenous growth. Journal of Political Economy 98, S103–S125.

Barro, R.J., Lee, J.-W., 1993. Losers and Winners in Economic Growth. NBER Working Paper No. 4341. National Bureau of Economic Research, Cambridge, Massachusetts; 1990; <http://www.nber.org/papers/w4341>.
Barro, R.J., Lee, J.-W., 2000. International Data on Educational Attainment: Updates and Implications. CID Working Paper No. 42, April 2000; <http://www.cid.harvard.edu/ciddata/ciddata.html>.
Barro, R.J., Sala-i-Martin, X., 2004. Economic Growth. The MIT Press, Cambridge, MA.
Bernanke, B.S., Gürkaynak, R., 1993. Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously. NBER Working Paper No. 8365; National Bureau of Economic Research, Cambridge, Massachusetts. <http://www.nber.org/papers/w8365>.
Black, D., Henderson, V., 1999. A theory of urban growth. Journal of Political Economy 107, 252–284.
Burdett, K., Mortensen, D.T., 1989. Equilibrium Wage Differentials and Employer Size. Discussion Paper No. 860; Northwestern Center for Mathematical Studies, 1989.
Carayol, N., Roux, P., 2003. Self-Organizing Innovation Networks: When Do Small Worlds Emerge? Cahiers du GRES No. 2003-8, Toulouse.
Cordoba, J., 2004. On the Distribution of City Sizes. Working Paper, Rice University.
Dunlop, J., 1938. The movement of real and money wage rates. Economic Journal 48, 413–434.
Duranton, G., Puga, D., 2003. Micro-foundations of Urban Agglomeration Economies. NBER Working Paper No. 9931; National Bureau of Economic Research, Cambridge, Massachusetts. <http://www.nber.org/papers/w9931>.
Fujita, M., Krugman, P., Venables, A.J., 1999. The Spatial Economy: Cities, Regions, and International Trade. The MIT Press, Cambridge, MA.
Gollin, D., 2002. Getting income shares right. Journal of Political Economy 110, 458–474.
Green, F., Machin, S., Manning, A., 1996. The employer-size wage effect: can dynamic monopsony provide an explanation? Oxford Economic Papers 48, 433–455.
Helsley, R.W., Strange, W.C., 1990. Matching and agglomeration economies in a system of cities. Regional Science and Urban Economics 20, 189–212.
Heston, A., Summers, R., Aten, B., 2002. Penn World Tables Version 6.1. Center for International Comparisons at the University of Pennsylvania.
Kim, S., 1990. Labor heterogeneity, wage bargaining, and agglomeration economies. Journal of Urban Economics 28, 160–177.
Lucas, R.E., 1988. On the mechanics of economic development. Journal of Monetary Economics 22, 3–42.
Manning, A., 2003. Monopsony in Motion: Imperfect Competition in Labor Markets. Princeton University Press, Princeton, NJ.
Ridder, G., van den Berg, G., 2003. Measuring labor market friction: a cross-country comparison. Journal of the European Economic Association 1, 224–244.
Romer, P.M., 1986. Increasing returns and long-run growth. Journal of Political Economy 94, 1002–1037.
Romer, P.M., 1990. Endogenous technological change. Journal of Political Economy 98, S71–S102.
Rossi-Hansberg, E., Wright, M.L.J., 2007. Urban structure and growth. Review of Economic Studies 74, 597–624.
Rotemberg, J.J., Woodford, M., 1990. Cyclical Markups: Theories and Evidence. National Bureau of Economic Research, Cambridge, Massachusetts; NBER Working Paper No. 3534; <http://www.nber.org/papers/w3534>.
Salop, S., 1979. Monopolistic competition with outside goods. Bell Journal of Economics 10, 141–156.
Tarshis, L., 1939. Changes in real and money wages. Economic Journal 49, 150–154.
Tse, C.Y., 2000. Monopoly, human capital accumulation and development. Journal of Development Economics 61, 137–174.
United Nations 1985, 1990, 1995, 1997. National Accounts Statistics: Main Aggregates and Detailed Tables. Publishing Division, United Nations, New York; 1985, 1990, 1995, 1997.
World Bank, 2004. World Development Indicators 2004. <http://publications.worldbank.org/ecommerce/catalog/product?item_id=631625>.