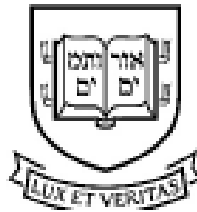


ECONOMIC ACTIVITY ACROSS SPACE:  
A SUPPLY AND DEMAND APPROACH

By

Treb Allen and Costas Arkolakis

COWLES FOUNDATION PAPER NO. 1856



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS  
YALE UNIVERSITY  
Box 208281  
New Haven, Connecticut 06520-8281

2023

<http://cowles.yale.edu/>



# John Maynard Keynes Narrates the Great Depression: His Reports to the Philips Electronics Firm

Robert W. Dimand & Bradley W. Bateman

To cite this article: Robert W. Dimand & Bradley W. Bateman (26 Oct 2023): John Maynard Keynes Narrates the Great Depression: His Reports to the Philips Electronics Firm, Challenge, DOI: [10.1080/05775132.2023.2272544](https://doi.org/10.1080/05775132.2023.2272544)

To link to this article: <https://doi.org/10.1080/05775132.2023.2272544>



Published online: 26 Oct 2023.



Submit your article to this journal [↗](#)



Article views: 26



View related articles [↗](#)



View Crossmark data [↗](#)



# John Maynard Keynes Narrates the Great Depression: His Reports to the Philips Electronics Firm<sup>†</sup>

Robert W. Dimand and Bradley W. Bateman

## ABSTRACT

In October 1929, the Dutch electronics firm Philips approached John Maynard Keynes to write confidential reports on the state of the British and world economies, which he did from January 1930 to November 1934, at first monthly and then quarterly. These substantial reports (Keynes's November 1931 report was twelve typed pages) show Keynes narrating the Great Depression in real time, as the world went through the US slowdown after the Wall Street crash, the Credit-Anstalt collapse in Austria, the German banking crisis (summer 1931), Britain's departure from the gold exchange standard in August and September 1931, the US banking crisis leading to the Bank Holiday of March 1933, the London Economic Conference of 1933, and the coming of the New Deal. This series of reports has not been discussed in the literature, though the reports and surrounding correspondence are in the Chadwyck-Healey microfilm edition of the Keynes Papers. We examine Keynes's account of the unfolding events of the early 1930s, his insistence that the crisis would be more severe and long-lasting than most observers predicted, and his changing position on whether monetary policy would be sufficient to promote recovery and relate his reading of contemporary events to his theoretical development.

## Introduction

On October 23, 1929, just as Wall Street began to crash<sup>1</sup> and the world economy moved into exceptionally interesting times, Dr. H. F. van Walsem, counsel and secretary to the Dutch electronics firm N. V. Philips Gloeilampenfabrieken<sup>2</sup>, wrote to “J. M. Keynes, Esq., C.B. Cambridge” asking him to write a monthly letter to the firm's Economic Intelligence Service about the state of the British economy and the world economy. John Maynard Keynes's letters to Philips, monthly from January 1930 to November 1931 and then, because of budget cuts to Philips's Economic Intelligence Service, quarterly from February 1932 to November 1934, show Keynes narrating the events of the Great Depression as they occurred, and reveal his perception of the convulsions of the

**CONTACT** Robert W. Dimand  [rdimand@brocku.ca](mailto:rdimand@brocku.ca), [robert.dimand@yale.edu](mailto:robert.dimand@yale.edu)  Department of Economics, Brock University, St. Catharines, Ontario L2S 3A1, Canada.

<sup>†</sup>Unpublished writings of JOHN MAYNARD KEYNES copyright The Provost and Scholars of King's College Cambridge 2023.

(Fall 2023: Visiting Professor of Economics, Yale University)

Robert W. Dimand is a professor of economics at Brock University, Canada, and a visiting professor at Yale University, and author of *The Origins of the Keynesian* and an editor of *The Elgar Companion to John Maynard Keynes*.

Bradley W. Bateman is president emeritus of Randolph College, Virginia, coauthor of *Capitalist Revolutionary: John Maynard Keynes* and an editor of *The Cambridge Companion to Keynes*.

world economy as he wrote his *General Theory of Employment, Interest and Money* (1936). This substantial body of Keynes's commentary on economic fluctuations (the November 1931 letter alone is twelve typed, double-spaced pages) has hitherto been neglected in the literature on Keynes. Keynes's reports and the associated correspondence, preserved in the Keynes Papers at King's College, Cambridge, are included in the 1993 Chadwyck-Healey microfilm edition of the Keynes Papers (section BM/5 Memoranda Exchanged with Business Houses), but the expense of this edition (which was sold only as a complete set of 170 reels of microfilm, priced at £9,700 or \$17,000, plus \$175 for a hardcover catalogue, Cox 1993) meant that only a few copies were sold. According to the WorldCat catalogue, there are five sets in libraries in the United States (Library of Congress, Harvard, Yale, Ohio State, and University of Texas at El Paso), two in Great Britain (Universities of Oxford and Sheffield), one in Canada (Victoria University in the University of Toronto) and a few in Germany (Göttingen), Italy and elsewhere but surprisingly little use has been made even of these copies of Keynes's letters to N. V. Philips. Neither Moggridge (1992) nor Skidelsky (1983–2000, 2003), major biographies of Keynes by the authors who know the Keynes Papers best, mentions Keynes's reports to Philips (but Backhouse and Bateman 2011, 129, have a paragraph about Keynes's July 1930 report). As Jacqueline Cox (1995, 171) notes, the thirty volumes of Keynes's *Collected Writings* (1971–1989) include “only a third of the bulk classified as economic” in the Keynes Papers at King's and do not include Keynes's philosophical papers there, while “the personal papers were barely touched.” Donald Moggridge (2006, 136–137) observes that “There has, inevitably, been heavier use of the Keynes Papers in King's College Cambridge, which have the advantage of being available elsewhere on microfilm, than, say, his papers in the National Archives or his correspondence with his publishers, the last of which reveals the risks of depending on the Cambridge collection alone.” A vast amount of research has been done about Keynes and his economics, yet not all the relevant material has been explored (see Backhouse and Bateman 2006, Dimand and Hagemann 2019).

These reports reveal Keynes's reading of what was happening in the British and world economies through the first four years of the Great Depression, and provide the empirical counterpart to the record of Keynes's theoretical development in this period given by notes taken by students at Keynes's lectures from 1932 to 1935 (Rymes 1987, 1989, Dimand 1988, Dimand and Hagemann 2019). After the success of *The Economic Consequences of the Peace* (1919), Keynes no longer needed to be paid for lecturing, and so gave a single series of eight lectures each year, on the subject of whatever book he was writing at the time, so his lectures from 1932 to 1935 are in effect annual drafts of the book that became *The General Theory*. These lectures at Cambridge and the reports to N. V. Philips on what was happening in the economy provide theoretical and empirical supplements to Keynes's *Collected Writings* (1971–1989), respectively, in following Keynes's intellectual development in the Great Depression, from *A Treatise on Money* (1930) to *The General Theory* (1936). In Keynes's workload, his reports to Philips from 1930 to 1934 took the place of the London and Cambridge Economics Service Special Memoranda on commodity markets that he wrote from 1923 to 1930 (Keynes [1923–30] 1983, 267–647), which provided an empirical counterpart to his normal backward-ation theory of futures contracts ([1923] 1983, 1930, Chapter 29).

Replying on October 31 to von Walsem's letter inviting him to write the monthly letter to the firm's Economic Intelligence Service, Keynes was "quite ready to discuss this proposal with one of your representatives" but wished to clarify "that there will be no question of the publication of the letters and that they will be purely for the information of your own people" – and that "it would not be practicable to me to undertake such work except in return for a somewhat substantial fee which might be higher than you would be willing to offer." On November 4, von Walsem assured him that the letters would not be published and "There are only two persons who, though not in our service, are closely related to our firm, who also receive a copy of our Intelligence Service which they, however, are bound to consider as absolutely confidential." He suggested £100 a year. On November 13, Keynes, having "considered your kind proposal in relation to the fees which I have received on previous occasions for somewhat analogous work," offered to undertake the task for an initial six months, for £150 a year<sup>3</sup>. Although Van Walsem had initially asked for the suggestion of other authors if Keynes preferred not take on the task at the suggested £100 a year, and Keynes equally pointedly offered to suggest such alternative authors if Philips did not care to pay £150 a year, Van Walsem accepted Keynes's terms for Philips on November 22: "We think it desirable that one of our gentlemen will see you in order to discuss some details in the first half of December next."

In the event two representatives of Philips (Messrs. Sannes and du Pré) met with Keynes for a discussion summarized "for good order's sake" by van Walsem on December 21, 1929 (by which time van Walsem had already received a December 18 note by Keynes on the Australian exchange position). He recorded agreement that Keynes's monthly letter would treat "some important factor in the development of the British economic situation and give your opinion as to its effects on trade in general and on our business in particular. Also you will draw our attention to important events in the domains especially interesting us, in so far as these come to your knowledge ... Whenever you think it necessary you will give us your views on the situation in different parts of the British Empire or eventually of other countries. If possible we shall suggest [to] you special points to be considered in your letters." Von Walsem wrote again on June 21, 1930 to confirm "that the arrangement has given us full satisfaction so that we are willing to continue on the same terms" and enclosed a cheque for 75 pounds. The arrangement also satisfied Keynes; he wrote on January 1, 1931, that "I have enjoyed preparing the letters." Keynes's letters balanced opinions about trade in general with observations about matters affecting Philips more specifically. Thus on January 11, 1930, Keynes stated that "The Factory capacity for Radio Sets seems to have become quite appalling during 1929" before proceeding more generally "to take this opportunity of emphasizing the anxiety which is felt here about the Australian position ... I think that Australia may have more difficulties with her balance of trade during the coming year than the Argentine."<sup>4</sup>

### **The Slump of 1930: Investment, Debts and Deflation**

Keynes's April 1930 letter suggested that, although a general improvement had not yet arrived, "there are a fair number of indications that we may be somewhere in the

neighborhood of the bottom point.” In particular, “the continuance of cheap money, and even more the expectation of such continuance, is bound to be effective in the situation in the course of a few months,” but the effect on employment would be slower than on business feeling and the Stock Exchange and “it would not be surprising to see British unemployment figures go on mounting even to the neighborhood of 2,000,000 up to the end of this calendar year. ... The effect of many rationalization schemes now in train will be for some time to come to improve profits rather than employment.” With a large amount of Australian gold en route to the Bank of England, “there is less anxiety about the British exchange position than there has been for a very considerable time past” and Keynes expected the creation of the Bank for International Settlements to have a positive effect on confidence, a foreshadowing of his emphasis at Bretton Woods on the importance of designing appropriate international monetary institutions. Keynes doubted that the Federal Reserve Board would reverse its cheap money policy “until business and employment in the United States is a great deal better than it is now.” This emphasis on expectations would be characteristic of Keynes’s *General Theory* (although equally in line with Irving Fisher’s quantity theoretic concern with expected inflation), as is the measurement of the ease of monetary policy by the cheapness of money, that is, by low nominal interest rates. Because nominal interest rates (especially short-term rates such as the Treasury Bill rate) were very low in a period of deflation, the Federal Reserve Board continued to view monetary conditions as easy throughout what Milton Friedman and Anna Schwartz (1963) later termed the “Great Contraction” of the US money supply (during which the monetary base increased, but not by enough to offset the rise in currency/deposit and reserve/deposit ratios), despite Fisher drawing the attention of his former student, Federal Reserve Governor Eugene Meyer, to the statistics on the shrinkage of the money supply, the sum of currency and demand deposits (Cargill 1992, Dimand 2019).

On June 24, 1930, H. du Pré emphasized that, “In reply to your remarks about the character of your monthly letters, we assure you that we leave it entirely to you to judge in each case which are the topics which are most worth being discussed by you.” Nonetheless, “There is one question upon which we particularly should like to have your opinion.” Keynes’s monthly letters had repeatedly stated that recovery depended on the bond market becoming more active, with new loans being used not just for the refunding of floating debt but for new productive investment. “But on the other hand these last months many articles in the economic press” saw excessive capacity in many industries; “in other words that the world has first to grow into a productive apparatus which is too big for immediate needs. If this should be true, can a renewed investment-activity soon be hoped for, and if it soon comes, would it really do good? Of course there would be less unemployment in a number of industries; but would not prices of consumptive commodities, and so cost of living, rise? And especially it might turn out after some time, that the new activity has only added to the – supposed – actual over-investment, so that the disequilibrium would only be greater. It may of course be that entirely new industries are going to take the lead, but we do not yet see any that are very likely to do so. We should be much obliged if you would solve this puzzle for us or at least give your views on the pretended overcapacity and its probable effects on future developments in your next letter.” This letter sheds light on the audience for Keynes’s reports in the secretariat of N. V. Philips: not just salesmen looking for tips

about the market for radio sets in Great Britain or elsewhere, but thoughtful businessmen pondering sophisticated economic issues such as the dual nature of productive investment in creating demand while increasing capacity (a problem to which the warranted growth rate of Harrod 1939 was an attempted solution).

In his July 1930 letter (seven typed pages, plus a six-page note on the bond market), Keynes warned that “it is now fully clear the world is in the middle of an international cyclical depression of unusual severity ... a depression and a crisis of major dimensions ... I believe that the prevailing opinion in the United States is still not pessimistic enough and is relying too much on a recovery in the early autumn, an event which is, in my opinion, most improbable. Nothing is more difficult than to predict the date of recovery. But all previous experience would show that a depression on this scale is not something from which the recovery comes suddenly or quickly.” He felt that “The optimism of Wall Street and the hoarding tendencies of France may prevent any real recovery of the International Loan Market this year” and considered whether this might lead to “a psychological atmosphere in which really drastic scientific measures will be taken by Great Britain and the United States in conjunction to do what is humanly possible to cause a turn of the tide next spring. But one is traveling here into the realm of the altogether uncertain and unpredictable.” In contrast, the Harvard Economic Society (founded by Harvard economics professors Charles J. Bullock and Warren Persons) stated in its weekly letter on June 28, 1930, that “irregular and conflicting movements of business should soon give way to sustained recovery” and on July 19 that “untoward elements have operated to delay recovery but the evidence nevertheless points to substantial improvement” (quoted by Galbraith 1961, 150, see also Walter Friedman 2014).

Responding to du Pré’s query, Keynes reiterated that recovery would be preceded by “a substantial fall in the long-period rate of interest ... leading in due course to the recovery of investment.” But now he explained that he was not thinking of investment in manufacturing industry, “the world’s capacity for which is probably quite ample for the present.” Even at the highest estimate, the total cost of bringing Britain’s industrial plant up to date “would not use up the country’s savings for more than, say, three months. Moreover, when expected profits are satisfactory the rate of expenditure by manufacturing industry in fixed plant is not very sensitive to the rate of interest.”

“On the other hand,” in contrast to manufacturing, “the borrowing requirements for building, transport and public utilities are not only on a far greater scale, but are decidedly sensitive to the rate of interest. If I were to put my finger on the prime trouble to-day, I should call attention to the very high rate of interest for long-term borrowers ... the long-term rate of interest is higher to-day than it has been in time of peace for a very long time past. When, at the same time, there is a big business depression and prices are falling, it is not surprising that new enterprise is kept back at the present level of interest.” He drew attention to “those who might be called distress borrowers, that is say countries which have an urgent need for borrowing to pay off existing debts, and are consequently ready to pay a very high rate of interest,” citing prospective Austrian, Hungarian and Australian loans on the London bond market, and remarked that “the effect of the German Loan has been to supply the French Treasury with funds, which it has withdrawn from the French market and is keeping unemployed in the

Bank of France.” Keynes’s July 1930 letter (discussed briefly by Backhouse and Bateman 2011, 129) illuminates both his analysis of the present situation and the role of investment in his economics. His distinction between investment in manufacturing, responsive to expected profit rather than interest rates, and interest-sensitive investment in construction, transport and public utilities clarifies his theory of investment. Increased investment was crucial for recovery of the world economy, and low long-term interest rates were necessary for high levels of investment in construction, transport and public utilities, the largest part of investment (even if manufacturing investment depended more on expected profits). In regard to the current situation, Keynes explained the forces getting long-term interest rates high even when prices were falling and short-term interest rates were low, but felt that “progress has been made toward getting the necessitous borrowers out of the way.” On the immediate practical level, Keynes’s distinction between the determinants of the two categories of investment dealt with du Pré’s question of how low long-term interest rates could stimulate investment given excess productive capacity in manufacturing. And yet, unlike Harrod (1939), Keynes’s July 1930 letter did not come to grips with the theoretical point raised by du Pré, the dual character of investment in creating both demand and productive capacity.

Keynes’s August 1930 letter dissented from the view widely held in the United States “even in responsible quarters, that we may expect an autumn recovery with some confidence ... a good deal of the American optimism is based on analogies drawn from the date of recovery after the 1920-21 slump” (compare the Harvard Economic Society’s statement on August 30 that “the present depression has about spent its force,” quoted by Galbraith 1961, 150). He argued that “Too much emphasis cannot be laid on the really catastrophic character of the price falls of some of the principal raw materials since a year ago” (even larger than appeared from published index numbers, because those included a number of commodities subject to price controls), which “must profoundly affect the purchasing power of all overseas markets.” Long-term interest rates remained high, reducing new capital investment. In contrast, Keynes considered general opinion about the British position to be “perhaps a little too pessimistic.” Britain was already in a difficult position before the slump of 1929 and 1930, because of the 1925 return to the gold exchange standard at the prewar parity (over the eloquent protests of Keynes 1925). But the heavy unemployment in the slump was limited to textiles and heavy industry (iron and steel, coal, and shipbuilding), export-based sectors already hit by the return to gold at an overvalued exchange rate (in his December 1930 letter, Keynes stated that if textiles, iron and steel, and coal were omitted, there was practically no decline in the Index of Production from a year before and an improvement from two years before). Keynes explained that British unemployment statistics, when used in international comparisons, “probably overstate the case” since the British statistics included “a great many workers in definite employment, but working short time ... It is even the case that workers taking their normal summer holidays are now included in the figures of the unemployed.” According to *The Economist*, the aggregate profits of all British joint stock companies reporting their earnings in the first half of 1930 “were not only greater than in the previous year, but were larger than in any previous year. This was partly due to the prosperity of British Oil Companies operating abroad, but by no means wholly.” Nor did Keynes share the worries of financial opinion in London (and so some extent his own previous letter to Philips) about “the constant dribble of gold to France.”



In Keynes's September 1930 letter to Philips, he was "still of the opinion that real recovery is a long way off. But at the same time it seems to me not unlikely that we are at, or near, the lowest point ... It is time, therefore, to cease to be a 'bear', even if it is not yet time to be a 'bull'." His February 1931 letter began, "Glancing through the letters of previous months, I find that they were all extremely pessimistic (with a brief lapse into modified optimism in September, corrected in October). Nevertheless, in the light of the actual course of events they were scarcely pessimistic enough. Nor do I see any reason for expecting any appreciable alleviation in the coming months." His September 1930 letter reported that "An extraordinary example of the way in which a situation can suddenly turn round, when a tendency has been greatly overdone, has been seen on the London Stock Exchange in the last two weeks. There has been no recovery of business in Great Britain to account for it. The real facts are much as they were a month ago. But market pessimism, aided by bear operations, had brought security prices down to an absurdly low level not justified by the circumstances ... everyone knew in his heart that prices were falling to foolish levels. The result was that within a few days the prices of many leading securities had risen from 10 to 20 per cent." The stock market had diverged from any level that could be construed as reflecting underlying fundamentals, but then abruptly bounced back. Keynes again stressed that Britain was not doing as badly as the United States in the slump: the fall in the British index of production from the previous year "is certainly less than 10 per cent" whereas the US index of industrial production for July 1930 was 37% below that for July 1929.

Keynes's 1930 "October Letter" warned that, "The catastrophic increase in the value of money has raised the burden of indebtedness of many countries beyond what they can bear ... in many parts of the world the fall of prices has now reached a point where it is straining the social system at its foundations. Agriculturists and other producers of primary materials are being threatened with ruin and bankruptcy all over the world. It is useless to expect a recovery of markets in such conditions" (and in his February 1931 letter he again warned that "The prospect of a long series of defaults [by debtor countries exporting raw materials] during 1931 is not to be excluded"). All of the gains that Germany had received in the Young Plan for reparations compared to the Dawes Plan were obliterated because "the clause in the Dawes Plan by which her [Germany's] liabilities in terms of gold were to be modified in the event of a change in prices was not included in the Young Plan." Keynes declared himself "rather more pessimistic ... than a month ago." He remarked that in Britain, "Very slight steps have been taken, as yet, in the direction of reducing wages, which is probably inevitable, but will not get anyone much further if all countries alike embark on wage-cutting policies."

These themes of Keynes's October 1930 letter to Philips, the danger of ruin and bankruptcy from price deflation in a world where debts are fixed in money terms and the futility of wage-cutting, appeared publically in his December article in *The Nation and Atheneum* on "The Great Slump of 1930" (reprinted in his *Essays in Persuasion*, 1931). There Keynes (1931, 138–139) warned that, since wage and price deflation increases the real burden of debt and wage cuts reduce purchasing power, "neither the restriction of output nor the reduction of wages serves in itself to restore equilibrium" and went on to emphasize that "Moreover, even if we were to succeed eventually in reestablishing output at the lower level of money-wages appropriate to (say) the pre-war

level of prices, our troubles would not be at an end. For since 1914 an immense burden of bonded debt, both national and international, has been contracted, which is fixed in terms of money. Thus every fall of prices increases the value of the money in which it is fixed. For example, if we were to settle down to the pre-war level of prices, the British National Debt would be nearly 40% greater than it was in 1924 and double what it was in 1920; ... the obligations of such debtor countries as those of South America and Australia would become insupportable without a reduction of their standard of life for the benefit of their creditors; agriculturalists and householders throughout the world, who have borrowed on mortgage, would find themselves the victims of their creditors. In such a situation it must be doubtful whether the necessary adjustments could be made in time to prevent a series of bankruptcies, defaults, and repudiations which would shake the capitalist order to its foundations” (see also Dimand 2011). Here, before Fisher (1932, 1933, see Dimand 2019), was the concern with the effect of deflation on the real value of nominal deflation that reappeared in Chapter 19, “Changes in Money Wages,” of *The General Theory*, where Keynes (1936, 264) warned that “if the fall of wages and prices goes far, the embarrassment of those entrepreneurs who are heavily indebted may soon reach the point of insolvency – with severely adverse effects on investment.”

### **Contested Budgets, Trade Balance and the Banking and Exchange Crises of 1931**

In 1930, Keynes’s “November Letter” argued that foreign opinion underestimated the financial strength that accompanied Britain’s industrial weakness: “it is forgotten that the adverse tendencies of the foreign exchanges, until recently, have been due, not to the absence of a favorable foreign trade balance, but to the eagerness of British investors to take advantage of the high profits or high rates of interest obtainable abroad. In 1929 the British favorable balance available for new foreign investment was greater than that for any other country, greater even than that for the United States. The Bank of England’s difficulties were due to the fact that the pressure of savers to take advantage of opportunities abroad was even greater.” Subsequent events in Wall Street and elsewhere had made overseas investment less appealing to British savers, so that the Bank of England was holding twenty million pounds sterling more of gold than a year before. In his December 1930 letter, Keynes reported that, even though “The perpetual drain of gold to France provides a source of nervousness and irritation in the money market” and although thirty million pounds sterling of gold had moved from Britain to France in the previous three months, the Bank of England held twenty-two million pounds sterling more in gold than a year before (but Keynes’s March 1931 letter reported that a drain of twenty million pounds sterling of gold from the Bank of England in the previous three months “causing nervous talk to prevail in London”). Despite Keynes’s repeated insistence on the financial strength of sterling and the growing gold reserves of the Bank of England (less than a year before the crisis of August and September 1931 that forced Britain off the gold exchange standard), the underlying message was that capital mobility under fixed exchange rates would constrain even the Bank of England from trying to lower long-term interest rates to stimulate investment. Until Britain left

the gold standard and allowed sterling to float, Keynes's letters to Philips monitored the strength of protectionist sentiment in the British Government, but he lost interest in tariff proposals once the exchange rate was no longer pegged (see Keynes 1931). But there was one bright spot for Britain: Keynes's February 1931 letter stressed that "It must not be overlooked that England is gaining enormously by the tremendous drop in the price of her imports as compared with that of her exports."

Keynes's April 1931 letter to Philips is notable for explaining that Britain's apparent budget deficit of £23.5 million for the fiscal year ending March 31 "is not as bad as it sounds, since this figure is reached after allowing for the repayment of £67,000,000 of debt. So that, apart from debt repayments, there was a surplus on the year's workings of £43,500,000. It must be doubtful whether any other country is showing so favorable a result. Even if the sum borrowed for the unemployment fund, which lies outside the budget<sup>5</sup>, were to be deducted, there would still have on the year a net reduction of debt." The next year's was expected to be larger, but "If no debt were to be repaid, there would probably be no deficit, even for the forthcoming year." Keynes's May 1931 letter, reporting on the budget presented by Labor Chancellor of the Exchequer Phillip Snowden, noted that "there will still be some reduction of debt during the forthcoming year, though not on as large a scale as formerly." A few months later, when Snowden and Prime Minister Ramsay MacDonald broke with their party to join the Conservatives in a National Government to deal with a budget and exchange crisis, Snowden found it convenient to overlook that the apparent budget deficit was an artifact of budgeting for a reduction in the national debt, and to denounce his former Labor Cabinet colleagues for endangering the savings of small depositors by having the Post Office Savings Bank lend to the Unemployment Insurance Fund, without mentioning that such loans were guaranteed by the Treasury or that he had neglected to inform his Cabinet colleagues of the borrowing (as Keynes indignantly explained in two paragraphs in the draft of his November 1931 letter, deleted from the final version).

Keynes's May 1931 letter is also notable, in light of the subsequent exchange crisis that forced Britain off gold in September, for insisting that "The improvement in the sterling exchanges and the better gold position of the Bank of England, as it appears in the public returns, are not deceptive and may be assessed at even more than their face value." He held that "When there is no longer serious pressure on the Bank of England's gold, the stage will be set for really cheap money throughout the world ... It will not mean a recovery, but it will pave the way for the recovery of investment which must precede the recovery of prices and profits." Keynes again emphasized that "the fall in the prices of the commodities imported by Great Britain has been so much greater than the fall in the prices of her exports. On the visible trade balance Great Britain was £5,000,000 better off in the first quarter of 1931 than in either of the preceding years ... Thus the main burden of the present crisis falls on the raw-material-producing countries, and Great Britain is likely to gain gold in spite of the immense decline of her exports."

By the next month, as the Credit-Anstalt collapsed in Vienna (see Schubert 1991), as French and American capital then took flight from Germany (see Balderston 1994), and as share prices slumped in London, Wall Street and on most European bourses, Keynes felt "that we are now entering the crisis, or panic, phase of the slump. I am inclined to think that we look back on this particular slump we shall feel that this phase has been

reached in the summer months of 1931, rather than at any earlier date.” He warned that “the consequences of a change in the value of money, as reflected in the prices of leading commodities, so violent as that which has occurred in the last eighteen months, cannot be regarded too gravely. Until prices show a material rise the whole fabric of economic society will be shaken. Each decline of commodity prices and each further collapse on the Stock Exchanges of the world brings a further group of individuals or institutions into a position where their assets doubtfully exceed their liabilities.”

### **Looking across the Atlantic: The American Slump**

Keynes’s July 1931 letter focused on the United States, where 21% of the industrial population was unemployed with perhaps another 20% working only two or three days a week: “it is quite out of the question that there should be anything which could be called a true recovery of trade at any time within, say, the next nine months. The necessary foundations for such a recover simply do not exist.” Many of the loans of small banks to farmers or secured by real estate “are non-liquid and probably impaired. Thus there is a strong desire for the utmost liquidity while obtainable on the part of the ordinary Bank; and general unwillingness to take any unnecessary risks or to embark on speculative enterprise, even where the risk may be actuarially a sound one. The nervousness on the part of the Bankers is accompanied by a nervousness of the part of their depositors ... So there is quite a common tendency to withdraw money from the banks and keep resources hoarded in actual cash ... It was estimated that in the country as a whole as much as \$500,000,000 was hoarded in actual cash in this way” (see Fisher 1933, Friedman and Schwartz 1963, Bernanke 2000). Keynes stressed that, “The American financial structure is more able than the financial structure of the European countries to support the strain of so great a change in the value of money. The very great development of Bank deposit and of bondage indebtedness in the United States means that a money contract has been interposed between the real estate on the one hand and the ultimate owner of the wealth on the other. The depreciation in the money value of the real estate sufficient to cause margins to run off, necessarily tends therefore to threaten the solidity of the structure.”

Keynes reported in his July 1931 letter that although US agricultural wages had fallen by 20 to 25%, and there had also been large cuts to wages in small-scale industrial enterprises, hourly wages were practically unchanged for two thirds of the workers in large-scale industrial enterprises while the hourly wages of the other third had been reduced by some 10%. In October 1934, however, Keynes stated in his Cambridge lectures that “Labor will and has accepted reductions in money wages, in the USA in 1932, and it will not serve to reduce unemployment” with one student’s notes calling the money-wage reductions “catastrophic” (Rymes 1987, 131).

### **Germany Defaults, Britain Abandons the Gold Parity**

Turning from the United States, Keynes remarked near the end of his July letter that, “At the moment of writing there are heavy gold drains from London; but I do not think that this need be regarded with any undue alarm,” a judgment that proved too sanguine.

More presciently, he added “The real danger in the situation comes from the possibility of the declaration of a general moratorium in Germany and the collapse of the mark [Germany defaulted on July 15]. The repercussion of such events on the solvency of the banking and money market systems of the world would be most serious.” The next month, in his August 1931 newsletter (dated August 4), Keynes reported that “the bulk of the remaining short-term German debt is due to British and American banks and accepting houses; many accepting houses being landed with what are certainly frozen and may prove doubtful debts. Their own credit has suffered with the inevitable result, since they were the holders of large foreign balances, of a drain of gold from London ... it would seem to be only ordinary prudence to act on the assumption that, while worse developments in Germany are doubtless possible, even apart from this the general underlying position is worse than the ordinary reader of newspapers believes it to be.” While “Great Britain is suffering from the temporary shock to confidence due to the difficulties of the accepting houses,”<sup>6</sup> the situation of the world economy as a whole was more serious: “We are certainly standing in the midst of the greatest economic crisis of the modern world. Important though the German developments have been I would emphasize that these have been essentially consequences of deeper causes which are affecting all countries alike ... For there is no financial structure which can withstand the strain of so violent a disturbance of values.” A handwritten postscript at the end of the typed August 1931 letter warns Keynes’s readers “not to be encouraged even by the appearance of apparently good news. The world financial structure is shaken and is rotten in many directions. Patching arrangements will be attempted, but they will not do much good, and it would be a mistake to place reliance on them.” The next day, August 5, Keynes, writing to Prime Minister J. Ramsay MacDonald to urge rejection of the May Report, stated that “it is now virtually *certain* that we shall go off the existing parity at no distant date ... when doubts, as to the prosperity of a currency, such as now exist about sterling, have come into existence, the game’s up” (Keynes 1971–1989, Vol. XX, 591–593; Skidelsky 2003, 446), but he did not say so in print or to Philips – and he rejected, on patriotic grounds, a suggestion by O. T. Falk that the Independent Investment Trust, of which Keynes and Falk were directors, should replace a dollar loan with a sterling loan, which Keynes condemned as “a frank bear speculation against sterling.” The Independent Investment Trust lost £40,000 by not switching its financing (Keynes 1971–1989, Vol. XX, 611–612; Moggridge 1992, 528–529; Skidelsky 2003, 447).

It was not only the world financial structure that was shaken; so was the Secretary Department of N. V. Philips. On August 6, 1931, H. du Pré wrote plaintively to Keynes, “Though we could hardly expect otherwise from your former letters, we note that you are not at all optimistic about the developments in the latter part of this year. These last weeks we read in the papers some statements from several Americans (among them people of authority), which hold a somewhat more cheerful view for the coming months. Must we infer from your letter that they are still, or again, too optimistic or is it possible that since your return from America<sup>7</sup> there have been some improvements, which may lead one to expect some improvement at least for the autumn?” Even Roger Babson, who had made his reputation by being bearish about the stock market in September 1929 (as he had been since 1926), was bullish by early 1931 (see W. Friedman 2014).

Keynes's reply on August 12 crushed any hopes: "In response to your enquiry, nothing has happened to make me more optimistic. As regards America, I consider that recovery this autumn is altogether out of the question. But the minds of all of us are of course dominated by the European and indeed the world situation. This still seems to me to be, as I have already described it, more serious than the general public know. I should recommend as complete inaction as is possible until further crises, or further striking events of some kind or another have occurred to clear up the situation."

Keynes's September letter (dated September 10, 1931), after the Conservative-dominated National Government displaced Labor, warned that "the hysterical concentration on Budgeting economy, which has also spread to the curtailment of expenditure by Local Authorities is calculated to produce unfavorable developments. For the widespread curtailment of expenditure is certain to reduce business profits and increase unemployment and lower the receipts of the Treasury, whilst it will do very little to tackle what is the fundamental problem, namely the improvement of the British Trade Balance. We seem likely to be faced by a period during which the balance of trade will not be sufficient to give confidence to foreign depositors."

It turned out, however, that one part of the cuts in government spending, the reduction in pay of the armed services, did indirectly dispose of the balance of payments problem. Since the government's version of equal sacrifice was that a vice-admiral earning £5 10s a day would lose 10 shillings a day (a reduction of 1/11), while naval lieutenants earning £1 7s a day and able-bodied seamen earning 5 shillings a day should each lose a shilling a day, reductions of 1/27 and 1/5, respectively (Muggeridge 1940, 109n), a naval mutiny erupted at Invergordon on September 16 (the first British naval mutiny since 1797), leading to abandonment of a fixed exchange rate on September 21 and a prompt 20% depreciation of sterling. Once the gold parity was abandoned, interest rates could be lowered without any balance of payments crisis. Commander Stephen King-Hall remarked "the strange combination of circumstances which caused the Royal Navy to be used by a far-seeing Providence as the unconscious means of ... releasing the nation from the onerous terms of the contract of 1925 when the pound was restored to gold at pre-war parity ... In 1805 the Navy saved the nation at Trafalgar; it may be that at Invergordon it achieved a like feat" (quoted by Muggeridge 1940, 111n). As for the budget deficit, Chancellor Snowden, who in the preceding Labor government had steadfastly blocked any reduction in the Sinking Fund contributions for paying down the national debt, now presented a budget reducing the annual Sinking Fund contribution by £20 million. Keynes declared in his October 1931 letter to Philips, "Great Britain's inevitable departure from the gold standard having occurred, it has been received with almost universal relief and in industrial circles a spirit of optimism is now abroad ... Since the City and the Bank of England did their utmost to avoid the change, they feel that honor is satisfied. In other quarters the effect is to relieve a tension which was becoming almost unbearable ... I have no doubt at all as to the reality of the stimulus which British business has obtained." Fisher (1935), assembling data on twenty-nine countries, found that recovery began only once a country abandoned the gold parity and was able to pursue a looser monetary policy (see Dimand 2003).

Keynes concluded his October 1931 letter, “The general passion for liquidity is bringing the value of cash in terms of everything else to so high a level as to be very near breaking point. This does not apply to Great Britain since her crisis was a balance of payments crisis rather than a banking crisis strictly so called. Thus the possibility of a general European and American banking crisis is the main risk, the possibility of which has now to be borne in mind.” The US banking crisis culminated in the “Bank Holiday” of March 1933, while all the major German and Italian banks passed into government ownership.

On November 3, 1931, Dr. du Pré was “very sorry to say that the necessity for the strictest economy which makes itself felt in all departments of our concern at present, impels us to an important curtailment of the budget of our Economic Intelligence Service” which would now issue bulletins every three months, instead of monthly. He asked Keynes for quarterly letters for £50 per annum, instead of monthly letters for £150 per annum. Keynes replied on November 9 that he read the letter “without any great surprise. I had been rather hesitating in my mind as to whether it is worth while to continue the arrangement on the new basis. But on the whole I feel that I should not like to break the friendly relations which have arisen between us, merely because times are bad.” He accepted the offer<sup>8</sup>, asking to be reminded when each quarterly report was due, and enclosed his November letter stating that Britain was “to a considerable extent getting the best of both worlds since broadly speaking the countries from which we buy our food and raw materials have followed us off gold, whilst our manufacturing competitors have remained on the old gold parity.”<sup>9</sup> He felt that Continental observers were mistaken to think that Britain would want to return to gold: “Foreigners always underestimate the slow infiltration of what I have sometimes called ‘inside opinion’, whilst ‘outside opinion’ remains ostensibly unchanged. Then quite suddenly what ‘inside opinion’ becomes ‘outside opinion’. Foreigners are quite taken by surprise, but the change is really one which had been long prepared. In the later months of the old gold standard there was a hardly a soul in this country who really believed in it. But it was considered that it was our duty for fairly obvious reasons to do everything we possibly could to keep where we were.”

Keynes’s May 1932 quarterly letter stressed that, “The most important development, if one is thinking not so much of the moment but of laying the foundations for future improvement, is to be found in the return to cheap money, which was interrupted by the financial crisis of last summer and the departure from gold. I am more and more convinced in the belief, which I have held for some time, that an ultra-cheap money phase in the principal financial centers is an indispensable preliminary to recovery ... Nevertheless it would be imprudent to expect too much at any early date from the stimulus of cheap money. The courage of enterprise is now so completely broken, that the effect on prices of money however cheap will be very slow. I consider it likely, therefore, that the cheap money phase may be extremely prolonged and that it may proceed to unprecedented lengths before it produces its effect.” He concluded, “For the time being the world is marking time, – waiting for it does not quite know what. I emphasize again the fact that the position in Great Britain, and in some of her Dominions, is relatively good. But for the time being, I see no light anywhere else ... It would certainly be much too soon to take any steps whatever to be ready for a possible revival.”

## Looking across the Atlantic: Hope from the New Deal

Keynes's August 1932 memorandum was notable for its explanation of why US stock prices had risen sharply and why that need not signal an end to the industrial crisis: the financial crisis had driven down stock prices until "the securities of many famous and successful companies were standing at little more than the equivalent of the net cash and liquid resources owned by those companies ... the assets in question would either be worth nothing as a result of the general breakdown of contract, or must, in any circumstances apart from that, be worth a very great deal more than their quotations. Consequently, it is logical and right that the fear of their being worth nothing having been brought to an end, there should be a rapid recovery of the quotations on a very striking scale. It does not need a termination of the industrial crisis, or even an expectation of its early cessation, in order to justify the new levels."

In his February 1933 memorandum, commenting on the likely futility of the projected World Economic Conference, Keynes recalled that "I have myself put forward more drastic proposals for an international fiduciary currency, which would be the legal equivalent of gold. If this were agreed to, the position would be so much eased that various other desirable measures would also become practicable. I do not despair of converting British opinion to such a plan, but I am told that continental opinion would be almost unanimously opposed it." Keynes had contemplated such proposals long before Bretton Woods.

Keynes's August 1933 memorandum (actually mailed July 20, before Keynes left for holidays) held that "My own view is that President's Roosevelt's programme is to be taken most seriously as a means not only of American, but of world recovery. He will suffer set-backs and no one can predict the end of the story. But it does seem fairly safe to say that his drastic policies have had the result of turning the tide in the direction of better security not only in the United States, but elsewhere ... Perhaps in the end President Roosevelt will devalue the dollar in terms of gold by 30 or 40 per cent." His November 1933 memorandum regretted "the failure of the President during his first six months to act inflation as well as talk it. In actual fact Governmental loan expenditure in the United States up to the end of September was on quite a trifling scale" but since then it seemed to be increasing: "if during the next six months the President is at last successful in putting into circulation a large volume of loan expenditure, I should expect a correspondingly rapid improvement in the industrial prosperity of America. This, if it occurs, would have a great influence on the rest of the world and especially on Great Britain ... it might pave the way for a rate of improvement sufficiently rapid to deserve the name of real recovery." Keynes's February 1934 memorandum reported that in the United States "everything is moving strongly upwards. This is to be largely attributed to the fact that Governmental loan expenditure is now at last occurring on a large scale ... the disbursement by the American Treasury of new money against borrowing has reached or is approaching \$50,000,000 weekly and should maintain this rate for a few months to come." In his August 1934 memorandum, having visited the United States since his May memorandum, he found there "a recession which is somewhat more than seasonal," aggravated since his visit by a "failure of the corn crop ... so acute as to be little short of a national disaster" but the actual and prospective level of US Government loan-



financed expenditure made him optimistic about prospects for the US economy in the autumn and winter. He also reported that “the view is generally held in Great Britain that the gold block countries – including Holland not less than the others – cannot permanently maintain their present parity with gold without a disaster. Now or later it seems to us certain that the necessity for devaluation will be admitted.” The reports end with Keynes’s November 1934 memorandum, with no correspondence in the Keynes Papers concerning the end of his relationship with the Philips firm.

### Conclusion: The Message of Keynes’s Reports to Philips

Keynes’s letters to the Philips electronics firm reveal he perceived events in the British and world economies from the beginning of 1930 through November 1934, and provide pungent and insightful commentary. These reports high-light the importance to Keynes of cheap money as a stimulus to investment – he was not just concerned with fiscal policy as the means to recovery, however much he placed emphasis from 1933 onward on the loan-financed expenditure of the Roosevelt Administration in the US. Keynes’s response to a query from du Pré is particularly interesting about Keynes’s distinction between those investment expenditures that are sensitive to interest rates and those that are not. The reports stress a theme discussed more briefly in Keynes’s 1931 Harris Foundation lectures in Chicago (in Wright, ed., 1931) and in Chapter 19 of *The General Theory*, and at greater length by Irving Fisher (1932, 1933) (and later by Hyman Minsky 1975): since debt are contracted in nominal terms, a rise in the purchasing power of money increases the risk of bankruptcy, repudiation and default – and it is not just actual defaults that are costly, but also the perception of increased riskiness. Keynes recognized the exceptional seriousness of the Depression, dissenting firmly from predictions of an early recovery, and he saw clearly how defending overvalued gold parities forced central banks to keep interest rates high, instead of pursuing ultra-cheap money to restore investment. This hitherto-neglected body of evidence allows one to watch the unfolding of the world economic crisis of the early 1930s through Keynes’s eyes, extraordinary events as viewed and narrated by an extraordinary economist. At £12 10s per report (by no means a trivial sum at the time), N. V. Philips certainly got their money’s worth.

### Notes

1. “Thursday, October 24, is the first of the days which history – such as it is on the subject – identifies with the panic of 1929” (Galbraith 1961, 103–104), but already on Monday, October 21, Irving Fisher had characterized the fall in stock prices as just the “shaking out of the lunatic fringe” and on Tuesday, Charles Mitchell of the National City Bank declared that “the decline has gone too far” (Galbraith 1961, 102).
2. Philips Incandescent Lamp Works, later Philips Electronics, successor to a firm founded by Lion Philips (originally Presburg), maternal uncle of Karl Marx (Gabriel 2011, 44, 110, 291–93, 295, 299, 315, 334, 366). Although relations between uncle and nephew were “strained by politics” (Gabriel 2011, 291), Mary Gabriel (2011, 299) refers to Marx’s “fund of last resort, his uncle ... He had sold himself to this pragmatic businessman as a successful writer only temporarily short of cash.” Gabriel (2011, 642) remarks that “Marx’s dabbling in the stock market has been questioned by some scholars, who believe he may simply have wanted his uncle to believe he was engaged in ‘capital’ transactions, not *Capital*.” After the death of Lion

Philips, his sons did not reply to Marx's letter asking for help with his daughter Laura's wedding (Gabriel 2011, 364). Anthony Sampson (1968, 95) reported that the firm's chairman Frits Philips was "a keen Moral Rearmer and a fervent anti-communist, embarrassed by the fact that his grandfather was a cousin of Karl Marx."

3. For a sense of what £150 a year might have meant to Keynes: Moggridge (1992, 508, 585) and Skidelsky (2003, 417–418, 519, 565) report that Keynes's net worth fluctuated from £44,000 at the end of 1927 to £7,815 at the end of 1929, then rising to over £506,222 at the end of 1936, dropping again to £181,244 at the end of 1938. The offer from Philips came at a particularly low point in his finances. According to Skidelsky (2003, 265) "investment, directorship and consultancy income" accounted for more than 70% of Keynes's income between 1923-24 and 1928-29 (including £1,000 a year as chairman of National Mutual Life Assurance), books and articles for another 20%, leaving no more than a tenth of income from such academic sources as teaching, examining, being secretary of the Royal Economic Society and editor of its journal, and being Bursar and a Fellow of King's College.
4. However, writing to Keynes on January 21, H. du Pré was moved "to remark that the latest figures from the Argentine which, according to the handwritten note at the bottom of your letter, you intended to enclose, were not received here, so that we cannot give you an opinion about their importance for us."
5. When the majority report of the May Committee on National Expenditure projected on July 31, 1931, that the budget deficit for 1931-32 would be £120 million, necessitating £96 million of cuts to unemployment benefits, road construction, and government and armed forces pay, it counted all borrowing by the Unemployment and Road funds as "public expenditure on current account" as well as "the usual provision for the redemption of debt" of £50 million (Winch 1969, 126–130). Keynes accused the majority on the May Committee of not "having given a moment's thought to the possible repercussions of their programme, either on the volume of unemployment or on the receipts of taxation" – he estimated it would add 250,000 to 400,000 to the unemployed, and reduce tax receipts by £70 million (*New Statesman and Nation*, August 15, 1931; Keynes 1971-89, Vol. IX, 141–145; Winch 1969, 130, Skidelsky 2003, 446).
6. With regard to Britain, Keynes noted that "There is, however, tremendous pressure of public opinion towards the Government Economy, which means in the main a reduction in the salaries of Government employees and of the allowances of the unemployed. It is equally difficult for the present [Labour] Government either to refuse or concede concessions to this trend of opinion. But if a movement in this direction takes place, which is still most doubtful, it remains exceedingly open to argument whether the result on the actual level of unemployment will be favourable."
7. Keynes had given three Harris Foundation Lectures on "An Economic Analysis of Unemployment" at the University of Chicago in June and July 1931, published in Quincy Wright, ed. (1931), and reprinted in Keynes (1971-89), Vol. XIII. These lectures mostly expounded the analysis of Keynes's *Treatise*, but the third lecture also examined the debt-deflation process, the undermining of the financial structure by an increase in the real value of debts and fall in the nominal value of collateral (Keynes 1971-89, Vol. XIII, 359–361, see Dimand 2011).
8. He also raised a "small personal matter", asking for advice on buying a new wireless set that would "have a thoroughly good loud speaker, both for voice and music reproduction and should be able to pick up distant stations such as Moscow."
9. A passage crossed-out in the draft of Keynes's November 1931 letter, in the section discussing the general election, stated that, "As has been the case in the last three or four General Elections, it is that old wretch Lord Rothermere [publisher of the *Daily Mail*] who has been dead right. It is said that he has made a profit on the crisis of £100,000, buying majorities on the Stock Exchange." Skidelsky (2003, 472) relates that Keynes "consistently lost money (his own and his friends') on the results of general elections."

## References

- Backhouse, Roger E., and Bradley W. Bateman, eds. 2006. *The Cambridge Companion to Keynes*. Cambridge, UK: Cambridge University Press.
- Backhouse, Roger E., and Bradley W. Bateman. 2011. *Capitalist Revolutionary: John Maynard Keynes*. Cambridge, MA: Harvard University Press.
- Balderston, Theo. 1994. "The Banks and the Gold Standard in the German Financial Crisis of 1931." *Financial History Review* 1 (1):43–68. <https://doi.org/10.1017/S0968565000001554>
- Bernanke, Ben S. 2000. *Essays on the Great Depression*. Princeton, NJ: Princeton University Press.
- Cargill, Thomas F. 1992. "Miscellany: Irving Fisher Comments on Benjamin Strong and the Federal Reserve in the 1930s." *Journal of Political Economy* 100 (6):1273–7. <https://doi.org/10.1086/261861>
- Chadwyck-Healey. 1993. *The John Maynard Keynes Papers in King's College, Cambridge, 170 Reels of Microfilm*. Cambridge, UK: Chadwyck-Healey Ltd.
- Cox, Jacqueline. 1993. *A Catalogue of the John Maynard Keynes Papers in King's College, Cambridge*. Cambridge, UK: Chadwyck-Healey Ltd.
- Cox, Jacqueline. 1995. "Keynes: An Archivist's View." In *New Perspectives on Keynes*, Annual Supplement to History of Political Economy. Vol. 27, edited by Allin F. Cottrell and Michael S. Lawlor, 163–75. Durham, NC: Duke University Press.
- Dimand, Robert W. 1988. *The Origins of the Keynesian Revolution*. Aldershot, UK: Edward Elgar Publishing.
- Dimand, Robert W. 2003. "Irving Fisher on the International Transmission of Booms and Depressions through Monetary Standards." *Journal of Money, Credit and Banking* 35 (1):49–90. <https://doi.org/10.1353/mcb.2003.0002>
- Dimand, Robert W. 2011. "The Consequences to the Banks of the Collapse of Money Values, 1931 and 2009." In *Perspectives on Keynesian Economics*, edited by Arie Arnon, Jimmy Weinblatt, and Warren Young, 233–45. Heidelberg: Springer Verlag.
- Dimand, Robert W. 2019. *Irving Fisher*. London: Palgrave Macmillan, Great Thinkers in Economics.
- Dimand, Robert W., and Harald Hagemann, eds. 2019. *The Elgar Companion to John Maynard Keynes*. Cheltenham, UK.
- Fisher, Irving. 1932. *Booms and Depressions*. New York: Adelphi.
- Fisher, Irving. 1933. "The Debt-Deflation Theory of Great Depressions." *Econometrica* 1 (4):337–57. <https://doi.org/10.2307/1907327>
- Fisher, Irving. 1935. "Are Booms and Depressions Transmitted Internationally through Monetary Standards?" *Bulletin de L'Institut Internationale de Statistique* 28 (1):1–29. Reprinted in Dimand (2003).
- Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States, 1867–1960*. Princeton, NJ: Princeton University Press for the National Bureau of Economic Research.
- Friedman, Walter A. 2014. *Fortune Tellers: The Story of America's First Economic Forecasters*. Princeton, NJ: Princeton University Press.
- Gabriel, Mary. 2011. *Love and Capital: Karl and Jenny Marx and the Birth of a Revolution*. New York: Little, Brown.
- Galbraith, John Kenneth. 1961. *The Great Crash 1929, with New Introduction*. Boston, MA: Houghton Mifflin.
- Harrod, Roy F. 1939. "An Essay in Dynamic Theory." *The Economic Journal* 49 (193):14–33. <https://doi.org/10.2307/2225181>
- Keynes, John Maynard. 1983 [1923]. "Some Aspects of Commodity Markets." *Manchester Guardian Commercial, Reconstruction in Europe* 29 March 1923; reprinted in Keynes (1971–89), XII:255–66.
- Keynes, John Maynard. 1983 [1923–30]. *London and Cambridge Economic Service Special Memoranda*, as reprinted in Keynes (1971–89), XII: 267–647.
- Keynes, John Maynard. 1919. *The Economic Consequences of the Peace*. London: Macmillan.

- Keynes, John Maynard. 1925. *The Economic Consequences of Mr. Churchill*. London: Leonard and Virginia Woolf, at the Hogarth Press.
- Keynes, John Maynard. 1930. *A Treatise on Money*, 2 Volumes. London: Macmillan.
- Keynes, John Maynard. 1931. *Essays in Persuasion*. London: Macmillan.
- Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.
- Keynes, John Maynard. 1971–89. *The Collected Writings of John Maynard Keynes*, 30 volumes, general editors Donald E. Moggridge and E. A. G. Robinson, volume editors Elizabeth S. Johnson and Donald E. Moggridge. London: Cambridge University Press, for the Royal Economic Society.
- Minsky, Hyman P. 1975. *John Maynard Keynes*. New York: Columbia University Press.
- Moggridge, Donald E. 1992. *Maynard Keynes: An Economist's Biography*. London: Routledge.
- Moggridge, Donald E. 2006. "Keynes and His Correspondence." In *The Cambridge Companion to Keynes*, edited by Roger E. Backhouse and Bradley W. Bateman, 136–59. Cambridge, UK: Cambridge University Press.
- Moggridge, Malcolm. 1940. *The Thirties: 1930-40 in Great Britain*. London: Hamish Hamilton.
- Rymes, Thomas K., ed. 1987. *Keynes's Lectures: Notes of Students*. Ottawa: Department of Economics, Carleton University.
- Rymes, Thomas K., ed. 1989. *Keynes's Lectures, 1932–35: Notes of a Representative Student*. London: University of Michigan Press.
- Sampson, Anthony. 1968. *The New Europeans*. London: Hodder and Stoughton.
- Schubert, A. 1991. *The Credit-Anstalt Crisis of 1931*. Cambridge, UK: Cambridge University Press.
- Skidelsky, Robert. 1983–2000. *John Maynard Keynes*, 3 volumes. London: Viking.
- Skidelsky, Robert. 2003. *John Maynard Keynes, 1883-1946: Economist, Philosopher, Statesman*. London: Macmillan.
- Winch, Donald. 1969. *Economics and Policy: A Historical Study*. New York: Walker and Company.
- Wright, Quincy, ed. 1931. *Unemployment as a World Problem: The Harris Foundation Lectures*. Chicago, IL: University of Chicago Press.

# Economic Activity across Space: A Supply and Demand Approach

Treb Allen and Costas Arkolakis

**T**he spatial distribution of people is incredibly concentrated: 8 percent of the US population lives in the ten largest US cities, but those cities take up less than 0.1 percent of total US land area. Why this concentration? More generally, what determines the distribution of people and economic activity across space? And how can economic policies affect the spatial distribution of economic activity? This essay will show that these questions can be answered through the familiar lens of supply and demand curves.

We begin by applying this intuition to the well-known Rosen-Roback framework (Rosen 1979; Roback 1982). But as we will discuss, the distribution of economic activity in this early spatial model depends only on local geography, not on what happens to other regions. For example, a change in one location—say, a large infrastructure investment that improves its productivity—is predicted to have an identical impact on all other locations, regardless of where they are. Thus, intuitive spatial features like where a location is located on a map and who its neighbors are entirely absent: it is a spatial model where space does not matter.

In reality, spatial linkages create rich interactions between locations. One implication of these interactions is that a large infrastructure investment that improves the productivity in one location will have greater impacts on close-by locations than locations further away. To account for such spatial linkages, we extend the intuition of the Rosen-Roback model to modern economic geography frameworks

■ *Treb Allen is Distinguished Associate Professor of Economics and Globalization, Dartmouth College, Hanover, New Hampshire. Costas Arkolakis is Professor of Economics, Yale University, New Haven, Connecticut. Their email addresses are [treb.allen@dartmouth.edu](mailto:treb.allen@dartmouth.edu) and [costas.arkolakis@yale.edu](mailto:costas.arkolakis@yale.edu).*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at <https://doi.org/10.1257/jep.37.2.3>.

where locations are connected through the flow of goods, based on our earlier work in Allen and Arkolakis (2014). In this framework, the economic fate of a location depends not only on its own “local” geography but also on the local geography of its neighbors, the effect of which is mediated by the strength of the economic ties, creating a “global geography.” Despite this added complexity, we show the same tools based on supply and demand used to understand predictions of the earlier Rosen-Roback framework extend readily to a globally integrated world.

This globally integrated framework can be applied to understand both the direct and indirect impacts of real world economic policies that change either the local or global geography. We discuss how the framework can be applied to spatial data, while also highlighting the most common pitfalls and offering strategies for traversing them. Finally, we provide a brief overview of the many ways in which this framework has been applied thus to understanding the spatial distribution of economic activity, as well as pointing out several interesting and still unexplored questions for future researchers. To keep the discussion as straightforward and accessible as possible, we relegate all mathematical details and derivations to the Appendix, where we also provide a companion Matlab toolkit to help researchers apply these techniques on their own.

## **Understanding the Spatial Distribution of Economic Activity through the Lens of Supply and Demand**

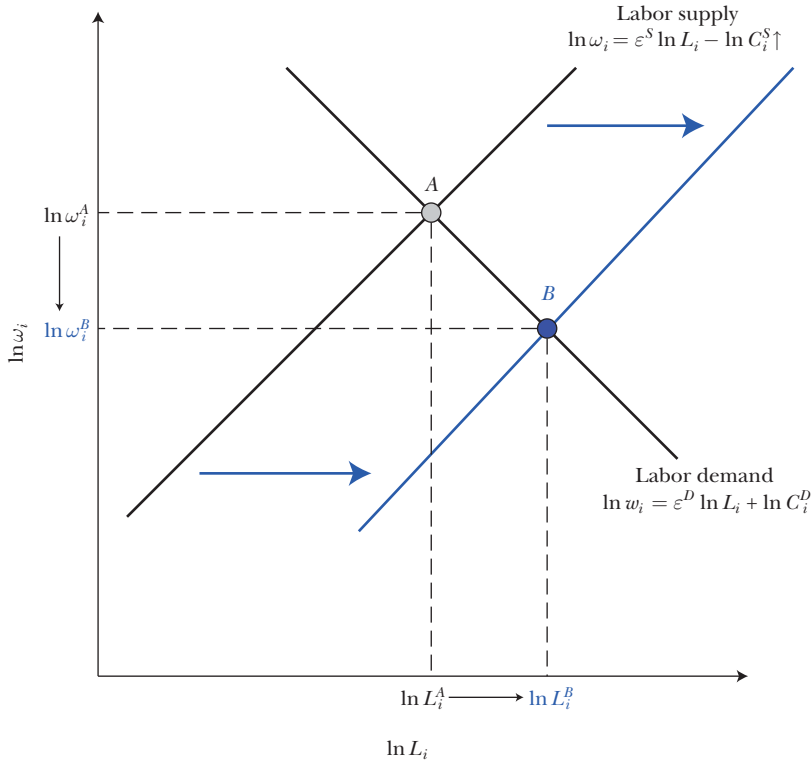
We now discuss the Rosen-Roback framework. Consider a world comprising many different locations. These locations each have their own “local” geography. The “local” geography of a location includes a whole host of things, from natural, geographic features like the climate, elevation, and natural beauty, to other less tangible characteristics of a location like the quality of its political institutions. Local geography can affect the spatial distribution of economic activity in two ways. First, it can affect the desire of people to live in a location and hence labor supply (we will call such factors “amenities”). Second, it can affect how productive people are in a location and hence labor demand (we will call such factors “productivities”).

Figure 1 illustrates the spatial equilibrium that results for labor demand and labor supply in this market. The labor market envisaged in this Rosen-Roback approach is one defined by location, rather than by the specific skills or sectors of workers: we think about the supply and demand for all workers in Detroit rather than the supply and demand for nurses or auto mechanics.

Let us first examine the labor demand curve more closely. Wherever people choose to live, they earn a wage from producing a good and then use that wage to buy goods and services. Let us assume that the wage they earn in any location  $i$  depends on two things: (1) the number of people living in that location; and (2) the productivities of that location. The result is a labor demand curve:

$$\ln w_i = \varepsilon^D \ln L_i + \ln C_i^D.$$

Figure 1  
A Supply Shock in the Local Spatial Equilibrium



Source: Authors' creation.

Note: This figure illustrates the effect of an increase in the labor supply shifter on the equilibrium population and wages in a local spatial economy.

In this relationship, the terms for wages and quantity of labor are expressed in log terms, and so  $\varepsilon^D$  is the demand elasticity.  $C_i^D$  is the local productivity in region  $i$  that arises from its local geography. The local productivity may capture, for example, how productive the factors are in location  $i$  or the relative cost of capital in a location.

The elasticity of demand is typically assumed to be negative, such that the labor demand function in Figure 1 is downward sloping. The economic intuition behind this slope is often based on assuming decreasing returns to scale in production of the good or simply the presence of a fixed factor such as capital (for example, see Kline and Moretti 2014; Donaldson and Hornbeck 2016). In other words, there is diminishing marginal product for each additional unit of labor added in the location. Thus, as the population of a location increases, each additional worker is less and less productive, causing the wage to fall. But other scenarios are possible. For example, the presence of external economies also can affect the slope of the demand

function. If more workers in a location result in everyone being more productive, the labor demand curve can become more elastic; if these external economies are sufficiently strong, the demand curve may even slope upwards. This situation can lead to outcomes like multiple equilibria or like “black hole” equilibria, where everyone lives in one location (for discussion, see Krugman 1991; Fujita, Krugman, and Venables 1999). While such scenarios have academic interest, in what follows we will stick with the more common (and, arguably, more empirically relevant) case of a downward-sloping demand curve.

If people each choose their place of residence to be as happy as possible, what makes people happy in this framework? Two things: higher consumption (so, all else equal, workers prefer higher real wages) and living somewhere nice (that is, a place with high amenities). In a model where everyone is identical, all inhabited locations must make people equally as happy. If prices are the same everywhere (so that the real wage is the nominal wage) and the amenity value of a location depends in part on how many other people live there, then workers’ indifference across all inhabited locations generates this labor supply curve:

$$\ln w_i = \varepsilon^S \ln L_i - \ln C_i^S.$$

Again, the left-hand side of the equation is the wage for each worker in the region  $i$ , and  $L_i$  is the number of workers in the region. Because wages and the quantity of workers are expressed in logs,  $\varepsilon^S$  is the elasticity of labor supply.  $C_i^S$  is the local amenity in region  $i$ —for example, better parks or planetariums.<sup>1</sup>

Economists usually think of a supply curve as sloping upward, as the labor supply curve is shown in Figure 1. A common underlying assumption in this setting is that the supply curve will slope up as long as more people in a location make each individual less happy; for example, the presence of a housing market where a higher population drives up housing prices and rents or the existence of idiosyncratic preferences where a higher population means the marginal resident’s match quality is worse can also lead to upward sloping labor supply curves.<sup>2</sup>

It is theoretically possible for the labor supply curve to slope downward (and issues of multiplicity and black holes to arise) if the amenity value of a location is increasing in its population, perhaps because of greater investments in public goods or greater variety in consumables in that specific location, but again, we will set that possibility aside here.

In this model, the equilibrium of economic activity—that is, the population and wage in a specific location—arises from combining the labor demand and labor supply curves. The spatial equilibrium is highlighted at point A in Figure 1.

<sup>1</sup>See the online Appendix A.1 for a particular microfoundation that delivers the specific labor demand and labor supply functions shown here.

<sup>2</sup>For a discussion of heterogeneous preferences and housing market, see Helpman (1998), Allen and Arkolakis (2014), Redding (2016), and Ahlfeldt et al. (2015).



To see how the “local” geography shapes the spatial equilibrium, consider a simple counterfactual scenario where the amenity value of residing in a location improved. For example, suppose the advent of air conditioning technology made the hot climate of the US Southwest less oppressive. An improvement in amenities shifts outward the labor supply curve, moving the equilibrium from point A to point B in Figure 1. The population in the location increases, but its wage declines. The US Southwest is now a better place to live, but the influx of workers depresses the wages.

The fact that we can analyze each location separately, depending on the amenity shock it receives, illustrates the somewhat paradoxical nature of the Rosen-Roback framework. It is a spatial model, but the distribution of economic activity depends only on local geography, not on what happens to other regions. Intuitive spatial features like where a location is located on a map and who its neighbors are entirely absent: it is a spatial model where space does not matter. By looking at one location at a time, it does not consider economic linkages between those locations.<sup>3</sup> Taking such linkages into account will create the concept of “global” geography which we introduce and analyze next.

## **The Role of Global Geography in the Spatial Distribution of Economic Activity**

Different locations can be linked with each other in many ways: people may live in one location and work in another; people may migrate from one location; people may talk with each other, leading to the spatial diffusion of ideas; and so on. But perhaps the most obvious spatial linkage occurs through the flow of goods. Much of what an individual consumes is produced in another location: according to the 2017 United States Commodity Flow Survey (CFS 2017),<sup>4</sup> most freight shipments crossed state boundaries, with only 22 percent of the value of freight destined for a state also originating in the same state. Moreover, the pattern of trade flows are far from uniform. As panel A of Figure 2 highlights using the same data, nearby states trade more with each other while the total volume of trade increases with the size of the trading partners, a phenomenon originally observed in international trade flows and oftentimes referred to as “gravity” (Anderson 2010; Head and Mayer 2013).

How does incorporating such spatial linkages affect the spatial equilibrium? It turns out that much of the basic intuition above remains; in particular, we can still analyze the spatial equilibrium using the familiar techniques of supply and demand, albeit now augmented with a concept of both “local” and “global” geographies.

<sup>3</sup>In the Rosen-Roback framework, a change in the local geography in one location can have aggregate general equilibrium effects on, say, the price of capital. But such general equilibrium effects affect all locations equally and hence do not affect the spatial distribution of economic activity.

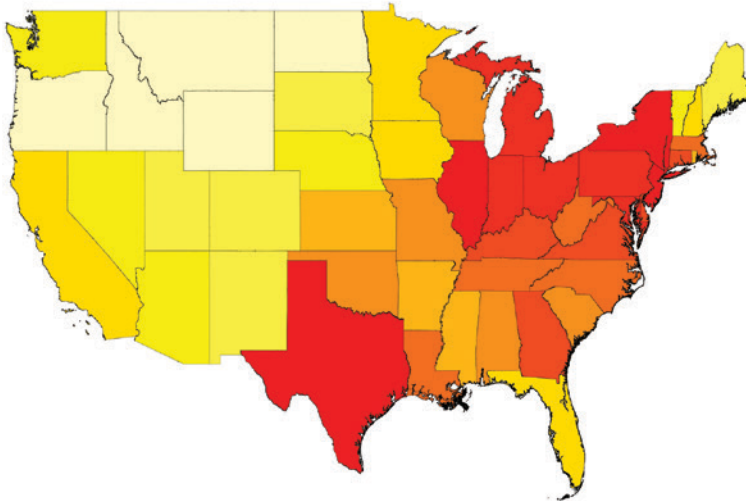
<sup>4</sup>The Commodity Flow Survey is conducted by the US government and is the primary source of data on within-US trade flows. In general, it is difficult to measure intracountry trade flows, making analysis of within-country trade difficult, although notable exceptions include work in Canada (Anderson and Van Wincoop 2003; McCallum 1995), India (Donaldson 2018), and the Philippines (Allen 2014).

*Figure 2*  
**Spatial Linkages and Market Access**

Panel A. Interstate trade flows



Panel B. Market access



*Source:* Authors calculations based on data from CFS (2017).

*Notes:* This figure illustrates the spatial linkages across US states arising from trade flows. Panel A depicts the relative size of state-to-state bilateral trade flows, with thicker red lines indicating larger values and thinner yellow lines indicating smaller values. Panel B indicates the resulting (outward) market access of each state assuming trade costs  $T_{ij}$  are inversely proportional to distance, with the darker red states indicating greater outward market access and the lighter yellow states having lesser outward market access.

### **The Global Geography**

The model discussed below is based on prior work (Allen and Arkolakis 2014), but variations of this spatial framework with equivalent or similar mathematical

formulations have recently been used in a variety of frameworks.<sup>5</sup> The setup retains the same features as above, but now we introduce a key distinction: goods are no longer costlessly traded. There are trade relationships between different locations, governed by the presence of spatial frictions.

These spatial frictions can be described as the economic distance between regions  $i$  and  $j$ . Conceptually, economic distance is proportional to the value of trade flows between two locations (conditional on origin and destination fixed effects). There are many possible factors that influence the economic distance between locations—whether they speak the same language, share the same legal systems, share similar cultural heritages, and so on. But one of the most important contributors to economic distance is simply the geographic distance between any two locations. Indeed, one of the most robust empirical relationships in all of economics is that trade flows between locations are roughly inversely proportional to the geographic distance between them (for discussion, see Disdier and Head 2008; Chaney 2018). Put another way, a very good start to measuring “economic distance” is simply with geographic distance.

When spatial frictions exist and goods are no longer costlessly traded, two things change. First, the price of goods produced by workers in a location depends in part on how nearby the consumers of those products are. The closer the consumers are, the more demand for their products and the higher the price (and hence the higher the wage) that the workers can obtain. This outward market access affects the labor demand curve of a location. Second, the price of goods purchased by consumers in a location depends in part on how nearby the producers of those products are. The closer the producers, the lower the price for those products and the higher the real wage of the consumers. This inward market access acts as a shifter to the labor supply curve of a location.

Together, the outward and inward market accesses comprise the global geography of a location. Following Anderson and Van Wincoop (2003) and Redding and Venables (2004), the outward market access ( $MA_i^{out}$ ) can be expressed algebraically as:

$$MA_i^{out} = \sum_j T_{ij} \times \frac{Y_j}{MA_j^{in}},$$

where  $T_{ij}$  is the inverse of economic distance between two locations and  $Y_j = w_j L_j$  is the total income of location  $j$ . Intuitively, outward market access summarizes the selling potential of a market, indicating how well a region is connected to other locations. For example, New Jersey has a high outward market access because there are lots of potential consumers of its products in its neighboring states of New York and Pennsylvania. Outward market access is greater when neighboring locations are closer (that is, when the inverse economic distance  $T_{ij}$  is greater), which is especially

<sup>5</sup>See for example Redding (2016), Donaldson and Hornbeck (2016), Allen, Arkolakis, and Takahashi (2020), Faber and Gaubert (2019), and Eckert and Peters (2022). Redding and Rossi-Hansberg (2017) offer a comprehensive review of the quantitative spatial framework.

beneficial when those neighboring locations are richer (have higher  $Y_j$ ) or have worse alternatives for buying their own goods ( $MA_j^{in}$ ).

“Inward market access” is similarly defined as the capacity of locations to buy from other locations:

$$MA_j^{in} = \sum_i T_{ij} \times \frac{Y_i}{MA_i^{out}}.$$

For example, New Jersey also has high inward market access because it is able to purchase its goods from nearby large producers. Like outward market access, inward market access is greater the smaller the economic distance to other locations, and again this matters more when nearby locations either produce a lot higher ( $Y_i$ ) or have poor alternatives for selling their goods (that is, have a lower  $MA_i^{out}$ ).

Outward and inward market accesses are obviously quite closely related and, indeed, will be proportional to each other in the special case when economic distances are the same in both directions. Note, however, that the economic distance that matters for inward market access is the one in which a location is the destination, whereas for outward market access, the economic distance that matters is the one in which the location is the origin. As a result, when economic distances are not the same in both directions, the inward and outward market accesses will generally be different.

The global geography summarizes how each location depends on economic activity in all other locations, where closer locations are given greater weights. These algebraic formulations highlight that inward and outward market accesses are intertwined, with each dependent in part on the other. Despite this interdependence, it is straightforward to solve for both the market access measures as long one observes the income in each location and the economic distances between locations. The companion Matlab code available as an appendix to this paper provides a convenient algorithm for doing so.

Panel B of Figure 2 depicts the (outward) market access for each US state, where we proxy the inverse economic distance  $T_{ij}$  with the inverse of geographic distance, measured as the distance (as the crow flies) between the geographic center of each state. States with high economic output that are close to other states with high output such as those in the Northeast have good market access; states with less economic output that are far away from states with higher economic output such as Montana have poor market access. As we will discuss in the next main section, an appealing feature of this framework is that the inverse economic distance can also be measured more explicitly with a combination of observed bilateral trade flows and observed bilateral geographic characteristics such as distance or time of travel.

### **The Global Spatial Equilibrium**

It turns out the global spatial equilibrium with spatial linkages can be analyzed using labor supply and demand curves, just as in the local spatial equilibrium above. Now, however, supply and demand will not only depend on local geography, but also

on global geography. In particular, the labor demand now also depends on outward market access  $MA_i^{out}$ , becoming:

$$\ln w_i = \varepsilon_{local}^D \ln L_i + \varepsilon_{global}^D \ln MA_i^{out} + \ln C_i^D.$$

Better outward market access acts analogously to better local productivities,  $C_i^D$ , shifting the demand curve for local labor outwards with an elasticity  $\varepsilon_{global}^D$ . That elasticity is greater the less substitutable the goods produced in  $i$  are with goods produced elsewhere in the world.

Similarly, labor supply now depends on inward market access  $MA_i^{in}$ , becoming:

$$\ln w_i = \varepsilon_{local}^S \ln L_i + \varepsilon_{global}^S \ln MA_i^{in} + \ln C_i^S.$$

Better inward market access acts analogously to better local amenities  $C_i^S$ , shifting the supply curve for labor outwards with an elasticity  $\varepsilon_{global}^S$  which again is larger the less substitutable goods produced in different locations are with each other.

The two limiting cases deserve special mention. When  $\varepsilon_{local}^S$  grows very large and approaches infinity, the local population becomes invariant to changes in economic conditions, whereas when  $\varepsilon_{local}^S$  becomes very small and approaches zero, labor supply is infinitely elastic to local economic conditions. These special cases correspond to important cases in the literature, as we will discuss below.

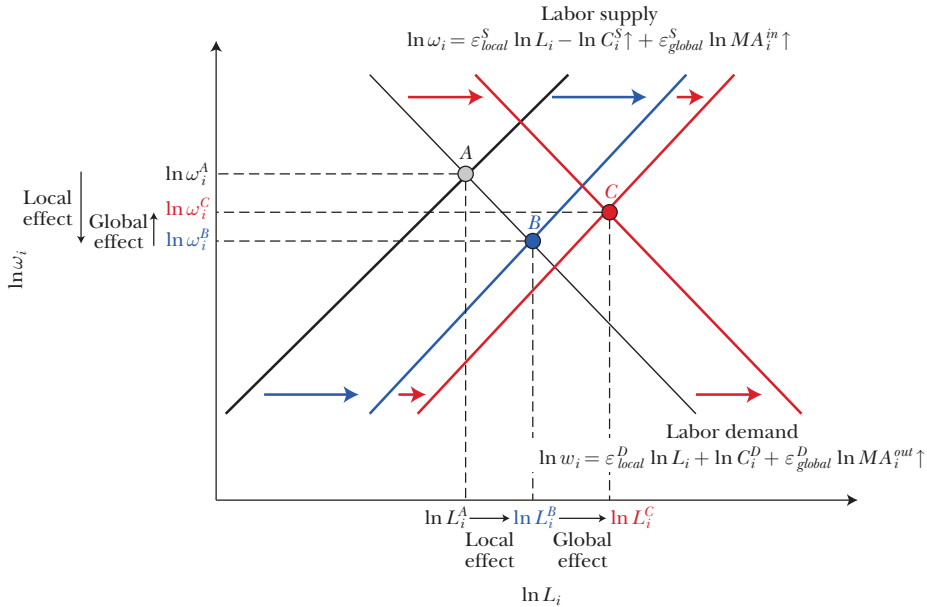
Given the global geography, the global spatial equilibrium is determined just as in the local spatial equilibrium above: find the wage and population in each location that equates supply with demand; point A on panel A of Figure 3 depicts such an equilibrium.

So what has changed in the global spatial equilibrium? The crucial insight is that the global geography in one location depends on the spatial equilibria in all other locations. If something changes about the local geography anywhere in the world, it will affect the global geography everywhere in the world, although it will affect nearby locations more than locations far away. Hence, the global geography puts space back into the spatial economy.

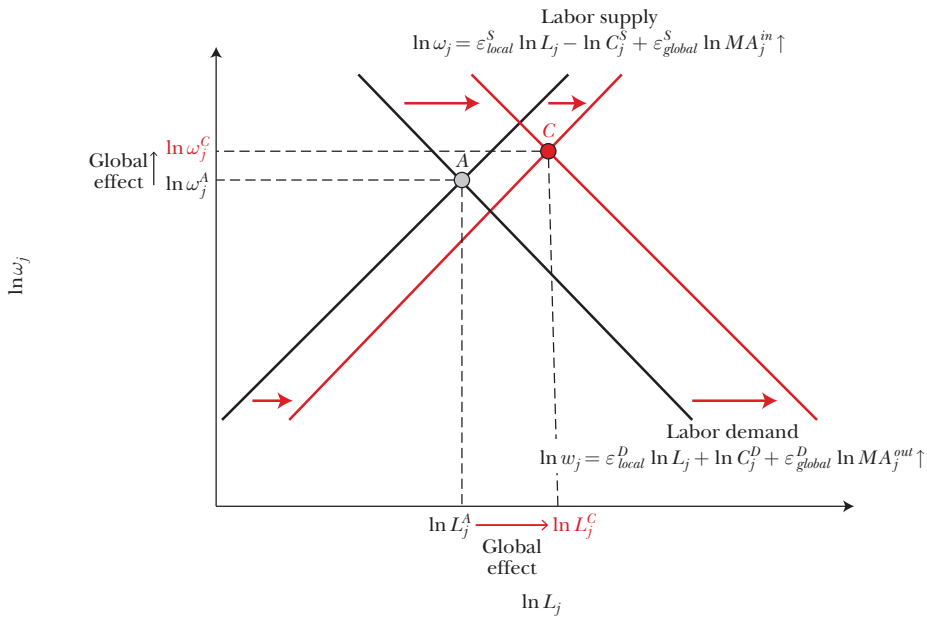
To illustrate this global spatial equilibrium, let us return to the example above. Suppose that air conditioning is invented, which makes some hot and previously inhospitable location  $i$  much more hospitable, raising the amenity of living there. Again, this innovation will shift outward labor supply curve in location  $i$  to point B in panel A of Figure 3, increasing the population in location  $i$  and reducing the wages. But the story does not end here, as this change in population and wages will affect the global geography. As long as the elasticity of local demand is greater than  $-1$ , the income  $Y_i$  of location  $i$  will increase, raising both the inward and outward market access and resulting in an additional shift outward to both the labor demand and labor supply curves. This additional global effect further increases the population in location  $i$  and mitigates the downward fall in wages, as illustrated in point C in panel A of Figure 3.

Figure 3  
**A Supply Shock in the Global Spatial Equilibrium**

Panel A. The directly affected location



Panel B. An indirectly affected location



Source: Authors' creation.

Note: This figure illustrates the effect of an increase in the labor supply shifter in one location its own equilibrium population and wages (panel A) and another neighboring location (panel B).

At the same time, changes in the economic activity in location  $i$  affect the global geography of other locations. Consider a neighboring location  $j$  initially in equilibrium, as illustrated by point A in panel B of Figure 3. Because the income of location  $j$  has improved, both its supply and demand curves will shift outwards as well. Intuitively, the greater nearby economic activity both increases the demand for the goods produced in  $j$  and increases the supply of goods consumed in  $j$ . As a result, the population in  $j$  increases too (and its wages rise), changing its equilibrium to point C in panel B of Figure 3, despite there being no change in its own local geography.<sup>6</sup>

But will changes in the economic activity in location  $j$  not have subsequent impacts on the global geography in all other locations? And will those changes not have even further impacts on the global geography, *ad infinitum*? Yes and yes: indeed, this infinite feedback loop between the global geography in every location is part of what makes the global spatial equilibrium so interesting to study. In reality, point C in panels A and B of Figure 3 represents the limit of the infinite sequence of these adjustments of each location's global geography to adjustments made in the global geography everywhere else. Indeed, this iterative process is what both the algorithm for calculating the equilibrium change in market accesses in the companion Matlab code and many tools for studying the mathematical properties of the equilibrium system are based upon.<sup>7</sup>

Having shown how one can determine the global spatial equilibrium through the use of supply and demand curves, we now turn to describing the process through which this framework can be combined with spatial data to assess the impact of changes in geography on the real world spatial distribution of economic activity.

## Estimating Labor Supply and Demand

In the previous section, we saw how a supply and demand framework can be used to understand how changes in the geography affect the distribution of economic activity across spatially connected locations. One of the most attractive

<sup>6</sup>Whether nominal wages rise or fall—that is, whether outward or inward market access increases more—depends on the choice of the numeraire. Here we set mean wages equal to one as the numeraire, so falling wages in location  $i$  must be offset by rising wages elsewhere.

<sup>7</sup>In the special case where the augmented labor supply curve is infinitely elastic, the local and global demand elasticities are equal in magnitude, and the inverse economic distances are symmetric, the equilibrium global economy is one in which the wages and populations of each location are (log) proportional to the eigenvector centrality of a location in the network defined by the world geography (that is, by the combination of the economic distances, productivities, and amenities). Higher eigenvector centrality means that a node in a network is nearby to other nodes with high eigenvector centralities. Here, it means that locations are more populated (and wealthier) the closer they are to other more populated (and wealthy) locations. Moreover, the eigenvalue of the system corresponding to this eigenvector turns out to be the welfare of the global economy (which is characterized by a single scalar because the infinitely elastic labor supply ensures welfare is equalized across all locations). In the more general case, the equilibrium of the spatial economy constitutes a network system of nonlinear equations. The properties of such systems remains an active field of research: Allen, Arkolakis, and Li (2020) offer a starting point.

aspects of the global spatial frame work described above is its ability to integrate seamlessly with readily available spatial data. In this section, we describe this interplay between theory and data.

### **Spatial Economic Data: Local and Linkages**

We focus here on two types of spatial data: data on the local economic activity of a location and data on the strength of economics linkages between locations across space.

Suppose that a researcher can observe in the data how many people reside in a certain location  $L_i$  and the total income of a location  $Y_i$ . Indeed, such data are readily available; for example, in the United States, population data and income data at the county level can be constructed from the decennial Census going back to the year 1840. The IPUMS (Integrated Public Use Microdata Series) National Historical Geographic Information Systems (Manson 2020) has provided an enormous public good in assembling these data and making them publicly available. Even in parts of the globe where spatially disaggregated income data are not readily available, one can proxy for economic activity using satellite data on the intensity of lights at nighttime, a practice pioneered by Henderson, Storeygard, and Weil (2012) and summarized in this journal in Donaldson and Storeygard (2016). Furthermore, databases that assemble information from various sources provide disaggregated information on economic activity at a granular geographic level, such as the G-econ database (Nordhaus and Chen 2006) that provides proxies of global income and population at the one-arc degree.

We furthermore assume that all income accrues to labor, which allows us to recover average wages for a location given knowledge of income and population. This strong assumption clearly abstracts from sources of income like capital, landholdings, firm profits, and others. One could argue that all these sources of income eventually accrue to individuals as well; indeed, as long as the income remains in a particular location, the predictions of the global spatial framework does not change by incorporating these other sources of income. (For example, as long as individuals in a location own their own homes, a model where individuals spend money on housing is no different—we say it is “isomorphic”—to the framework described above.) But in reality, not all income earned in a location accrues to the labor in that location, and such spatial flows of income would present another linkage between locations that we abstract from here.

Next consider data on economic linkages across space. As noted earlier, geographic distance is offers a convenient proxy for economic distance. But recently, researchers have begun to improve upon the distance proxy with measures of actual travel costs between locations. For example, Donaldson (2018) estimates the relative cost of traveling between locations by means of road, rail, and waterways by calculating the lowest cost route using Dijkstra’s (1959) algorithm—the same algorithm used by, for example, Google Maps. Allen and Arkolakis (2014) use a continuous space extension of the Dijkstra algorithm known as the Fast Marching Method (Tsitsiklis 1995; Sethian 1999) to calculate travel times along



the optimal route between locations. Allen and Arkolakis (2022) offer an analytical solution for the inverse economic distance as a function of the underlying transportation network.

Intuitively, these related approaches all share two advantages. First, they provide more precise estimates of the economic distance between two locations than distance alone would provide. (For example, Milwaukee, Wisconsin, and Grand Rapids, Michigan, are about 115 miles apart as the crow flies, but travel between the two around Lake Michigan more than doubles the distance). Second, accounting for the underlying transportation network allows researchers to assess how changes in transportation infrastructure (for example, improving the interstates I-90 and I-94 that connect Milwaukee and Grand Rapids) affect the spatial distribution of economic activity.

For any observed measure(s) of the economic linkages, the inverse economic distance ( $T_{ij}$ ) can then be constructed by regressing the observed (log) value of trade flows on those measures, conditioning on the origin and destination fixed effects. The predicted values of this gravity-model regression (excluding the estimated fixed effects) are the implied inverse economic distance.<sup>8</sup> For example, if one uses travel times as a measure of economic linkages, the inverse economic distance would be the product of travel time and its estimated coefficient from such a regression.

### Estimating Supply and Demand

Given measures of income  $Y_i$  in each location and a measure of the strength of the linkages  $T_{ij}$  between locations, we can calculate the global geography of every location—that is, the inward and outward market accesses  $MA_i^{in}$  and  $MA_i^{out}$ .<sup>9</sup> We provide an iterative algorithm for solving that nonlinear system of equations in the companion Matlab code.

Now let us return to our augmented supply and demand equations for the global case. We observe the left-hand-side price variable, the wage for each location  $w_i$ , and the right-hand-side quantity variable, the population  $L_i$ . We also observe the data needed to calculate the market access variables ( $MA_i^{in}$  and  $MA_j^{out}$ ).

We would like to estimate the coefficients on the right hand side variables, which represent the local and global elasticities of supply and demand. In doing so, the residual terms will be equal to measures of local productivity and local amenities

<sup>8</sup>An alternative procedure would be to calibrate the inverse economic distance to exactly match the observed bilateral trade flows by including the regression residual in its construction. Such a procedure—which is closely related to the “exact hat algebra” pioneered by Dekle, Eaton, and Kortum (2008) and discussed in Costinot and Rodríguez-Clare (2014)—can result in an over-fitting problem when conducting counterfactuals (Dingel and Tintelnot 2020).

<sup>9</sup>Recovering the global geography from the observed income and economic distances is a well-behaved problem. One can show using tools from Allen, Arkolakis, and Li (2020) that there exists unique (to-scale) inward and outward market accesses  $MA_j^{in}$  and  $MA_i^{out}$  that solve the equations for any set of incomes  $Y_i$  and inverse economic distances  $T_{ij}$ .

(that is,  $\ln C_i^D$  and  $\ln C_i^S$ ).<sup>10</sup> Or put another way, we would like to estimate a system of supply and demand curves where we observe data on equilibrium outcomes of price and quantity at different times, which poses problems that are all-too-well understood!

How do we go about estimating our supply and demand curves? It might perhaps be more informative to start with what not to do. Following in the footsteps of Baldwin and Taglioni (2006), let us award medals for different types of errors that can arise, ranking them from most to least obvious.

### **The Bronze Medal Error**

One glaring mistake in estimating supply and demand equations and—our “bronze medal” error—would be to use ordinary least squares regression. This approach is clearly not appropriate due to familiar simultaneity issues: what appears in data on wages and workers are the intersections of supply and demand curves, which do not trace out the shape of either a supply or a demand curve, but rather a series of movements in both of them. (To put it another way, because the right-hand-side population variable is determined in equilibrium from equating supply and demand, it will be correlated with both the productivity and amenity shifters.) As a result, the coefficient from such an ordinary least squares regression will not recover either the supply or demand elasticity.

One strategy for overcoming this bronze medal error would be to employ instrumental variables; for example, using variation in the amenity  $\ln C_i^S$  as an instrument for the equilibrium population to estimate the labor demand elasticity and using variation in the productivity as an instrument for the equilibrium population to estimate the supply elasticity. Conceptually, this involves looking at a source of shifts in labor supply (in this case, local amenities) to trace out a labor demand curve, and a source of shifts in labor demand (in this case, changes in local productivity) to trace out a labor supply curve. As long as the chosen instrumental variation in the amenities and productivities are uncorrelated, this will yield consistent estimates of the demand and supply elasticities.

What are examples of such instruments? One example comes from Glaeser and Gottlieb (2009), who argue that the advent of air conditioning improved the amenity of locations with warm climates. Under the assumption that the climate of a location is not also correlated with the change in the productivity of a location, the climate of a location can be used as an instrument for change in population to identify the demand elasticity (for example, Allen and Donaldson 2020).

Conversely, Allen and Donaldson (2020), following Bustos, Caprettini, and Ponticelli (2016), argue that increased global demand for soy improved the productivity of locations particularly well-suited for the production of soy. Under the assumption that the potential yield of soy in a location (say, relative to its potential

<sup>10</sup>This approach of recovering the underlying geography based on the supply and demand residuals is equivalent (but perhaps easier to digest) to an approach that directly inverts the equilibrium market clearing conditions, as in Allen and Arkolakis (2014) and Redding (2016).

yield for corn) did not also change the amenity of a location, the potential relative yield of soy to corn can be used as an instrument to identify the supply elasticity. Of course, the climate or agroclimatic properties are likely correlated with myriad characteristics of a location, making it unlikely these assumptions hold when comparing wages and populations across locations in cross section at a point in time. As such, it is preferable to rely on panel variation, looking at changes in wages and populations across locations over time (or, equivalently, including location fixed effects in the estimation of the supply and demand equations).

### **The Silver Medal Error**

Somewhat less obviously, our “silver medal” error would be to ignore the spatial linkages between locations and simply estimate supply and demand using the local supply and demand equations based on the Rosen-Roback model. However, doing so ignores the variation in inward and outward market access across locations, relegating that variation to the residual term.

The instrumental variable strategy just described to address simultaneity bias is insufficient to address this bias. To see this, suppose you are estimating the labor demand equation, while using an amenity shifter like the arrival of air conditioning as an instrumental variable for population. Even if that amenity shifter is uncorrelated with productivities, it will be correlated with outward market access, biasing the estimate of the demand elasticity. Indeed, the only situation where this bias does not arise is in the special case when all locations share the same market access (as in the local spatial equilibrium).<sup>11</sup>

Fortunately, avoiding this mistake is straightforward: from the discussion above, one can construct measures of inward and outward market access from readily available spatial economic data. Including these market access measures in the supply and demand equations is a simple remedy to avoid the silver medal error.

### **The Gold Medal Error**

An even more subtle concern is that outward and inward market access measures are themselves almost surely correlated with the productivity and amenity of a location. After all, the market access of a location depends in part on its own economic activity, which of course depends in equilibrium on its productivity and amenity. As a result, just including the market access measures in the supply and demand equations as controls will result not only in biased estimates of both the local and the global elasticities of supply and demand.

To address this concern, one can again use an instrumental variables strategy, instrumenting for both the population in a location and for the market access of

<sup>11</sup> Our “silver medal” error is similar in spirit to Baldwin and Taglioni’s (2006) “gold medal” error of failing to control for variation in market access in gravity equations. The two errors are distinct because unlike a gravity regression, the supply and demand regressions are not estimated using bilateral flows. As a result, their proposed solution of controlling for market access with origin and destination fixed effects does not apply here.

that location. We discussed above possible instruments for the population; what about for market access? An appropriate instrument would be correlated with market access, but uncorrelated with local productivities or amenities.

In conceptual terms, think of market access as a type of inverse economic distance-weighted average of economic activity near a location. For an appropriate instrumental variable, suppose you use the measures of local productivities and amenities along with plausible values of the model elasticities to calculate the local equilibrium of a hypothetical economy using the basic local-area Rosen and Roback supply and demand equations. In this hypothetical economy, spatial linkages do not matter and the only heterogeneity in productivities and amenities across locations arise from observables. Next, combine the implied equilibrium income in each location from this hypothetical economy with the observed economic distance and use the market access expressions above to calculate what the market access would be in such a hypothetical economy. This hypothetical market access measures how well connected each location is to the rest of the world, if the income in each location depended only on its observed productivities and amenities.

The hypothetical market access is a valid instrument for the actual market access under the assumption that observed productivities and amenities elsewhere in the world are uncorrelated with a location's own unobserved productivities and amenities. Using the hypothetical market access as an instrument then isolates the impact of market access on the supply and demand curves using this variation in productivities and amenities elsewhere through the spatial structure of the model.<sup>12</sup> Examples of such "model implied" instruments can be found in Monte, Redding, and Rossi-Hansberg (2018), Allen, Arkolakis, and Takahashi (2020), and Adão, Arkolakis, and Esposito (2019).

### **Taking Stock**

Suppose you have successfully avoided the bronze, silver, and gold medal errors by estimating the labor supply and demand curves while appropriately using instrumental variables for the observed population and the market access terms. Now what?

You are now armed with estimates of the model elasticities, data on wages, populations, and market access terms, and with residuals terms from the supply and demand equations that correspond to the productivities and amenities in each location. Put another way, if you know the supply and demand elasticities, you can always find the local geography such that the observed distribution of economic activity—combined with the inverse economic distances you have constructed—is the global spatial equilibrium of the model.

<sup>12</sup>Another possibility would be to construct an instrument based on the augmented global supply and demand equations but excluding the own location (and perhaps also nearby locations) from the sum. Even if there is no spatial correlation in the productivity and amenity of locations, however, the equilibrium economic activity elsewhere depends in part on the economic activity of the own location (and hence the own productivity and amenity shifters), so such an instrument is unlikely to satisfy the exclusion restrictions.

Because you have recovered the geography that is consistent with the observed economic activity and you know the model elasticities, you are now able to assess how changes to the geography will affect the global spatial equilibrium. In the next section, we will discuss ways in which this approach can inform understanding concerning the effects of various events and policy decisions.

## **Understanding the Spatial Impact of Economic Policies**

We have seen how the global and local geographies interact through supply and demand to shape the spatial equilibrium and how those supply and demand curves can be combined with spatial data to apply the framework to the real world. Now we are equipped to describe the many types of questions that can be addressed with such a framework. We classify these questions into three types: those examining the impact of changes to the local geography, those examining the impact of changes to the global geography, and those which extend the framework above to incorporate additional spatial linkages beyond the flow of goods. We make no pretense here of offering a full survey of the literature; instead, our goal is to illustrate the extraordinary range of this work across events, policies, places, and times.

### **Local Geography Shocks**

Consider first the question of how changes to local geography—changes to amenities which shift the supply curve or changes to productivities which shift the demand curve—affect the spatial distribution of economic activity.

Changes in the natural environment due to climate change offer many such examples. Rising sea levels and the resultant flooding both reduce the amount of land available for production and reduce the attractiveness of living in a coastal location, shifting both supply and demand curves in such locations inward, inducing populations to migrate elsewhere. Desmet et al. (2018) study the long-run impact of coastal flooding using a dynamic variation of the framework described here, finding that approximately 1.5 percent of the world population will be displaced by the year 2200 under current projections of the extent of flooding. Changing temperatures and patterns of precipitation also affect the suitability of different locations for producing different types of crops, affecting the productivity of different locations. Costinot, Donaldson, and Smith (2016) examine the long-run impact of estimated future changes in agricultural productivity across the globe to assess its impact on the spatial distribution of economic activity, estimating that climate change will reduce the global value of agricultural output by approximately one-sixth.

Conflict and war can also reduce local productivities and amenities, although it remains an outstanding question for how long after the conflict these effects persist. For example, Davis and Weinstein (2002) examine the rebuilding of Japan after World War II, finding that the postwar distribution of economic activity closely

mirrored the pre-war distribution, suggesting that wartime destruction was not enough to overcome fundamental characteristics of different locations. In contrast, Chiovelli, Michalopoulos, and Papaioannou (2018) find the removal of landmines in the period after Mozambique's civil war had substantial impacts on the spatial distribution of economic activity, especially after accounting for the impacts of the de-mining on market access—that is, on the global geography.

Technological innovations may also increase the productivities in certain locations, shifting the labor demand curve outward. For example, Bustos, Caprettini, and Ponticelli (2016) present evidence that the introduction of genetically modified soybeans in Brazil had heterogeneous effects across areas with different soil and weather characteristics, and also was a labor-saving technology that ended up boosting industry. Caliendo et al. (2018) extend the framework above to incorporate intersectoral linkages along with spatial linkages to examine, for example, how local productivity improvements resulting from California's computer industry boom and the introduction of shale oil production in North Dakota affected the spatial distribution of economic activity. Some interesting topics for future research along these lines include the spatial effects of automation (as in Acemoglu and Restrepo 2020) or new technologies that allow for remote work (as in Dingel and Neiman 2020; Althoff et al. 2022).

Place-based policies enacted by the government can also be viewed as shifts to the local demand or supply curves (depending on the particular nature of the policy). For example, Diamond and McQuade (2019) show that tax credits for low-income housing projects across 129 counties nationwide raised housing prices and reduced crime rates in low-income neighborhoods, but reduced housing prices in high-income neighborhoods. Some recent work seeks to characterize the trade-offs of such policies; for example, how policies that attract high-skill workers to low-wage cities can have broader social benefits and the equity-efficiency trade-offs of focusing place-based policies on locations with a dense concentration of low-income households (for discussion, see Fajgelbaum and Gaubert 2018; Gaubert, Kline, and Yagan 2021).

### **Global Geography Shocks**

Now let us turn our attention to how changes to global geography—changes in the economic distances and the resulting changes in the market access—affect the spatial distribution of economic activity.

Investment in transportation infrastructure which reduces the economic distance between locations is a natural application for evaluating changes to global geography. For example, the US interstate highway system increased US welfare by 1.0 to 1.4 percent of GDP, more than its costs (Allen and Arkolakis 2014); the US railroad system constructed in the second half of the nineteenth century more than doubled the price of land in nearby agricultural counties (Donaldson and Hornbeck 2016); the Los Angeles Metro rail system increased commuting, but with little effect on productivity or amenities, and thus has considerably larger costs than benefits (Severen 2021); the Appalachian Development Highway System started in

1965 did benefit Appalachian counties, but most of the benefits accrued outside the region (Jaworski and Kitchens 2019); and the arrival of the steam railway in mid-nineteenth century London led to a doubling of population and land prices, as well as a geographical separation of workplaces and residences (Heblich, Redding, and Sturm 2020). Recent work has also examined the distributional implications of such infrastructure investments; for example, transportation infrastructure investments in New York City from 1870 to 1940 seem to have caused greater racial sorting and disparities (Lee 2022) and the recently constructed national highway system in China benefits the economy of larger regional cities at the expense of rural regions (Baum-Snow et al. 2020).

While the basic framework above abstracts from the possibility that the economic distances may depend in part on the amount of trade between two locations, Duranton and Turner (2011) demonstrate the empirical relevance of congestion by showing that neither additional roads nor mass transit seem to reduce congestion in US cities. Recent work has made substantial progress incorporating congestion into spatial frameworks like the one described above. For example, Fajgelbaum and Schaal (2020) study optimal transportation networks in the presence of traffic congestion. In applying their framework to European countries, they find that the desirable network depends on whether they focus on flows within countries or flows between countries. In a similar spirit, Allen and Arkolakis (2022) develop a spatial framework that includes congestion and apply it to the US highway network and the Seattle road network. These types of frameworks could also be used to evaluate congestion imposing tolls in specific areas of the cities, such as the London or Singaporean traffic toll system or the congestion price system suggested for downtown Manhattan.

Other recent work has sought to consider congestion in the context of ports, sea routes, and supply chains. In particular, the Allen and Arkolakis (2022) spatial framework for transportation and congestion has been applied to study the effect of several recent events in global shipping on the distribution of economic activity. For example, the 2016 expansion of the Panama Canal expanded trade between pairs of countries using the canal by 9 to 10 percent, although the costs of the expansion were borne by Panama (Heiland et al. 2019); the expansion of container shipping and Chinese-financed development of seaports across Africa and Asia is leading to reallocations away from more expensive ports like Singapore (Ducruet et al. 2020); and entrepôts, defined as shipping hubs that serve an intermediate role between place of origin and destination, play a key role in holding down global shipping costs (Ganapati, Wong, and Ziv 2020).

Another branch of this work looks at intermodal shipping: for example, how the construction of expressways in China early in the twenty-first century boosted exports (Fan, Lu, and Luo 2021) and how to identify the nodes between road, rail, and ports in the US economy that would provide the greatest gains from additional investment (Fuchs and Wong 2022). An exciting new area of work builds on the approach of Brancaccio, Kalouptsi, and Papageorgiou (2020), who develop a model of endogenous route choices of exporters and endogenous transportation

costs to study the global bulk shipping that constitutes 80 percent of world trade and evaluate the effect of large infrastructure projects such as the expansion of the Panama canal. Conwell (2022) combines endogenous route choices and traffic to find that an optimal subsidy on minibus entry in Cape Town, South Africa, may particularly benefit low-skill workers on long routes.

A classic example of changes in global geography arises from changes in international trade policy, like changes in tariffs. For example, Topalova (2010) examines the impact of the 1991 Indian tariff reduction to measure the impact of trade liberalization on poverty and rural districts, in which production sectors more exposed to tariff declines experienced slower decline in poverty and lower consumption growth. The recent escalation of tariff measures by large economies such as the United States and China has generated a renewed interest on the impact of tariff increases on the spatial distribution of economic activity, following the influential work of Fajgelbaum et al. (2020), who find that the recent US-China trade war reduced US real income by \$7.2 billion, with the benefits of tariffs concentrated in politically competitive counties.

A final set of questions can be thought of as how changes to the local geography in some locations affect the economy elsewhere through the global geography. For example, beginning with the influential work of Autor, Dorn, and Hanson (2013), there has been much work on how productivity increases in China have affected workers in the United States and elsewhere through spatial linkages. Autor, Dorn, and Hanson (2013) found that US labor markets that previously included import-competing manufacturing industries experienced job and economic losses from the “rise of China.” Caliendo, Dvorkin, and Parro (2019) use a spatial framework like the one above (expanded to include multiple sectors) to conclude that while there was an overall loss of manufacturing jobs from the rise of China, the US economy as a whole benefited, albeit with considerable variation across sector-state labor markets. The increase in demand elsewhere for goods or services in a location provides another example: Faber and Gaubert (2019) show that increasing international demand for tourism in Mexico causes large and significant local economic gains, which are in part driven by positive spillovers on manufacturing. In contrast, Allen et al. (2021) find that increasing international demand for tourism in Barcelona reduces the welfare of many local residents by increasing prices and crowding out local consumption.

### **Alternative Spatial Linkages**

The framework developed above focuses on spatial linkages between locations that arise through the trade of goods. But of course people interact across space in many ways, including commuting, migration, or even social and business personal networks (for example, Christakis and Fowler 2009). Some recent advances have incorporated other types of interactions into spatial frameworks like the one developed here.

Following the seminal work of Ahlfeldt et al. (2015), which considered how the rise and fall of the Berlin Wall affected the spatial distribution of economic



activity in that city, a number of papers have examined the impact of spatial interactions that arise through commuting flows. For example, Severen (2021), mentioned earlier, separates the commuting effect of the Los Angeles Metro from productivity or amenity effects, while Zárate (2022) find that extensions of subway lines in Mexico City lead to increased commuting and a shift from informal to formal jobs. Monte, Redding, and Rossi-Hansberg (2018) and Allen, Arkolakis, and Li (2015) combine commuting and spatial linkages in a single model: the first study finds that communities which win a competition for location of large plants have greater benefits if they have a more open commuting network; the second considers optimal zoning policy and finds Chicago would benefit from having more residences downtown and more business activity in outlying neighborhoods.

A related literature incorporates spatial linkages arising through altered migration patterns, extending the framework above to a dynamic setting. While the steady state (or balanced growth path) of these models resemble the static framework above, they are also able to yield predictions on the time it takes the economy to adjust to changes in geography. For example, in a global model with realistic geography, Desmet, Nagy, and Rossi-Hansberg (2018) examine different scenarios for migration and how eliminating migration restrictions could triple global welfare. Allen, de Castro Dobbin, and Morten (2018) show that walls built along the US-Mexico border altered migration patterns between Mexican municipalities and US counties. Tombe and Zhu (2019) argue that declining costs of internal migration in China can account for one-third of the aggregate growth in China's labor productivity from 2001 to 2005. Peters (2022) finds that the expulsion of ethnic Germans from eastern Europe after World War II, and their return to West Germany, increased aggregate income per capita by about 12 percent after 25 years. Finally, Kleinman, Liu, and Redding (2021) find that the interaction of migration and capital investment can help to explain why convergence of incomes between US states declined between 1965 to 2015.

Another spatial linkage garnering recent attention is the formation of production linkages across firms. For example, lower costs of searching for and creating linkage between heterogeneous buyers and sellers can drive down marginal costs, as Bernard, Moxnes, and Ulltveit-Moe (2018) and Bernard, Moxnes, and Saito (2019) find in applying their models to improved flow of people in Japan and to Norwegian customs data. Yet another spatial linkage can be measured by taking advantage of new data sets to assess the role of knowledge diffusion. Using nationally representative smartphone data, Couture et al. (2020) examine patterns of travel and communication. While using highly granular smartphone data, Atkin, Chen, and Popov (2022) find substantial returns to what are actually face-to-face interactions in Silicon Valley. Using Facebook data grouped by zip code (and thus anonymized), Chetty et al. (2022a, b) look at personal connections across socioeconomic groups and within cliques to study associations with economic mobility and determinants of connectedness.

Related studies look at the effects of new information technologies, documenting how the spatial spread of information can affect the distribution of

economic activity. For example, Steinwender (2018) finds that the introduction of the trans-Atlantic telegraph in 1866 provided information that affected cotton prices and trade flows, with gains equivalent to 8 percent of export value. Allen (2014) shows that including information frictions can make sense of observed patterns of regional agricultural trade flows prices in the Philippines. Akerman, Leuven, and Mogstad (2022) find that this improved access to information in makes trade patterns more sensitive to distance and economic size using broadband expansion in Norway.

Recent research has incorporated even more types of spatial linkages including electricity transmission (Arkolakis and Walsh 2022), piped water (Coury et al. 2022), and natural gas pipelines (Bachmann et al. 2022). The possibilities of adding additional spatial linkages or combining multiple types (or multiple layers) of linkages seem limitless. Moreover, extending the framework to include such interactions brings more realism and helps to illuminate the many ways in which geography shapes the spatial economy.

## **Conclusion**

This article has sought to serve three purposes. First, it was meant as an introduction to the reader about how geography shapes the spatial distribution of economic activity. In the classic Rosen-Roback framework, the answer depends solely on the “local” geography of each location and the equilibrium spatial distribution can be determined through familiar analysis of supply and demand curves. The major innovation of the new generation of economic geography models is to incorporate the spatial linkages between locations—putting space into the spatial model. The equilibrium can continue to be understood using the same supply and demand curves, but is appropriately augmented to incorporate the impacts of the “global” geography.

The second purpose was to guide the reader through the process of combining these spatial models with spatial data to understand how geography shapes the real world spatial economy. Detailed spatial data are now readily available and researchers can apply these data to the theory using the well-understood process of estimating supply and demand curves. With spatial linkages between locations arise potential pitfalls in estimation, but we offer strategies for traversing such issues. The end result is the ability to recover the underlying local and global geography such that the theory and data exactly correspond, allowing a researcher the ability to assess the impacts of any change in geography on the real world spatial distribution of economic activity.

Finally, we demonstrate the power of this close marriage between theory and data by highlighting the many types of questions that can be addressed. The types of questions and topics that can be examined using the framework here spans an incredibly wide range of topics, such as economic history, environmental, labor, public finance, urban, and international topics, to name a few. This is an exciting

time to be working on spatial issues: we have a new set of tools applicable to many interesting questions, most of which have yet to be tackled.

■ We thank our editors Erik Hurst, Nina Pavcnik, Timothy Taylor, and Heidi Williams as well as Steve Redding, Anson Soderbery, and Sayantan Roy for exceptionally helpful feedback and Saptarshi Majumdar for excellent research assistance. The authors acknowledge support by the National Science Foundation under grants SES-1658838 and SES-1658875.

## References

- Acemoglu, Daron, and Pascual Restrepo.** 2020. “Robots and Jobs: Evidence from US Labor Markets.” *Journal of Political Economy* 128 (6): 2188–2244.
- Adão, Rodrigo, Costas Arkolakis, and Federico Esposito.** 2019. “General Equilibrium Effects in Space: Theory and Measurement.” NBER Working Paper 25544.
- Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf.** 2015. “The Economics of Density: Evidence from the Berlin Wall.” *Econometrica* 83 (6): 2127–89.
- Akerman, Anders, Edwin Leuven, and Maren Mogstad.** 2022. “Information Frictions, Internet, and the Relationship between Distance and Trade.” *American Economic Journal: Applied Economics* 14 (1): 133–63.
- Allen, Treb.** 2014. “Information Frictions in Trade.” *Econometrica* 82 (6): 2041–83.
- Allen, Treb, and Costas Arkolakis.** 2014. “Trade and the Topography of the Spatial Economy.” *Quarterly Journal of Economics* 129 (3): 1085–1140.
- Allen, Treb, and Costas Arkolakis.** 2022. “The Welfare Effects of Transportation Infrastructure Improvements.” *Review of Economic Studies* 89 (6): 2911–57.
- Allen, Treb, and Costas Arkolakis.** 2023. “Replication data for: Economic Activity across Space: A Supply and Demand Approach.” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E186022V1>.
- Allen, Treb, Costas Arkolakis, and Xiangliang Li.** 2015. “Optimal City Structure.” Unpublished.
- Allen, Treb, Costas Arkolakis, and Xiangliang Li.** 2020. “On the Equilibrium Properties of Network Models with Heterogeneous Agents.” NBER Working Paper 27837.
- Allen, Treb, Costas Arkolakis, and Yuta Takahashi.** 2020. “Universal Gravity.” *Journal of Political Economy* 128 (2): 393–433.
- Allen, Treb, Cauê de Castro Dobbin, and Melanie Morten.** 2018. “Border Walls.” NBER Working Paper 25267.
- Allen, Treb, and Dave Donaldson.** 2020. “Persistence and Path Dependence in the Spatial Economy.” NBER Working Paper 28059.
- Allen, Treb, Simon Fuchs, Sharat Ganapati, Alberto Graziano, Rocio Madera, and Judit Montoriol-Garriga.** 2021. “Urban Welfare: Tourism in Barcelona.” Unpublished.
- Althoff, Lukas, Fabian Eckert, Sharat Ganapati, and Conor Walsh.** 2022. “The Geography of Remote Work.” *Regional Science and Urban Economics* 93: 103770.
- Anderson, James E.** 2010. “The Gravity Model.” *Annual Review of Economics* 3: 133–60.
- Anderson, James E., and Eric van Wincoop.** 2003. “Gravity with Gravitas: A Solution to the Border Puzzle.” *American Economic Review* 93 (1): 170–92.
- Arkolakis, Costas, and Conor Walsh.** 2022. “Clean Growth.” Unpublished.
- Armington, Paul S.** 1969. “A Theory of Demand for Products Distinguished by Place of Production.” IMF

- Staff Papers 16 (1): 159–78.
- Atkin, David, M. Keith Chen, and Anton Popov.** 2022. “The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley.” NBER Working Paper 30147.
- Autor, David H., David Dorn, and Gordon H. Hanson.** 2013. “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” *American Economic Review* 103 (6): 2121–68.
- Bachmann, Rüdiger, David Baqaee, Christian Bayer, Moritz Kuhn, Andreas Löschel, Benjamin Moll, Andreas Peichl, Karen Pittel, and Moritz Schularick.** 2022. “What If? The Economic Effects for Germany of a Stop of Energy Imports from Russia.” ECONtribute Policy Brief 028.
- Baldwin, Richard, and Daria Taglioni.** 2006. “Gravity for Dummies and Dummies for Gravity Equations.” NBER Working Paper 12516.
- Baum-Snow, Nathaniel, J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt.** 2020. “Does Investment in National Highways Help or Hurt Hinterland City Growth?” *Journal of Urban Economics* 115: 103124.
- Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito.** 2019. “Production Networks, Geography, and Firm Performance.” *Journal of Political Economy* 127 (2): 639–88.
- Bernard, Andrew B., Andreas Moxnes, and Karen Helene Ulltveit-Moe.** 2018. “Two-Sided Heterogeneity and Trade.” *Review of Economics and Statistics* 100 (3): 424–39.
- Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou.** 2020. “Geography, Transportation, and Endogenous Trade Costs.” *Econometrica* 88 (2): 657–91.
- Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli.** 2016. “Agricultural Productivity and Structural Transformation: Evidence from Brazil.” *American Economic Review* 106 (6): 1320–65.
- Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro.** 2019. “Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock.” *Econometrica* 87 (3): 741–835.
- Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte.** 2018. “The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy.” *Review of Economic Studies* 85 (4): 2042–96.
- Chaney, Thomas.** 2018. “The Gravity Equation in International Trade: An Explanation.” *Journal of Political Economy* 126 (1): 150–77.
- Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebe, Nathaniel Hendren, Robert B. Fluegge, Sara Gong et al.** 2022a. “Social Capital I: Measurement and Associations with Economic Mobility.” *Nature* 608 (7921): 108–21.
- Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebe, Nathaniel Hendren, Robert B. Fluegge, Sara Gong et al.** 2022b. “Social Capital II: Determinants of Economic Connectedness.” *Nature* 608 (7921): 122–34.
- Chiovelli, Giorgio, Stelios Michalopoulos, and Elias Papaioannou.** 2018. “Landmines and Spatial Development.” NBER Working Paper 24758.
- Christakis, Nicholas A., and James H. Fowler.** 2009. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. New York: Little, Brown.
- Conwell, Lucas.** 2022. “Are There Too Many Minibuses in Cape Town? Privatized Provision of Public Transit.” Unpublished.
- Costinot, Arnaud, Dave Donaldson, and Cory Smith.** 2016. “Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 million Fields around the World.” *Journal of Political Economy* 124 (1): 205–48.
- Costinot, Arnaud, and Andrés Rodríguez-Clare.** 2014. “Trade Theory with Numbers: Quantifying the Consequences of Globalization.” In *Handbook of International Economics*, Vol. 4, edited by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, 197–261. Amsterdam: North-Holland.
- Coury, Michael, Toru Kitagawa, Allison Shertzer, and Matthew Turner.** 2022. “The Value of Piped Water and Sewers: Evidence from 19th Century Chicago.” NBER Working Paper 29718.
- Couture, Victor, Jonathan Dingel, Allison Green, and Jessie Handbury.** 2020. “Quantifying Social Interactions Using Smartphone Data.” Unpublished.
- Davis, Donald R., and David E. Weinstein.** 2002. “Bones, Bombs, and Break Points: The Geography of Economic Activity.” *American Economic Review* 92 (5): 1269–89.
- Dekle, Robert, Jonathan Eaton, and Samuel Kortum.** 2008. “Global Rebalancing with Gravity: Measuring the Burden of Adjustment.” *IMF Staff Papers* 55 (3): 511–40.
- Desmet, Klaus, Robert E. Kopp, Scott A. Kulp, Dávid Krisztián Nagy, Michael Oppenheimer, Esteban Rossi-Hansberg, and Benjamin H. Strauss.** 2018. “Evaluating the Economic Cost of Coastal Flooding.” NBER Working Paper 24918.

- Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg.** 2018. "The Geography of Development." *Journal of Political Economy* 126 (3): 903–83.
- Diamond, Rebecca, and Tim McQuade.** 2019. "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development." *Journal of Political Economy* 127 (3): 1063–1117.
- Dijkstra, Edsger Wybe.** 1959. "A Note on Two Problems in Connexion with Graphs." *Numerische Mathematik* 1: 269–71.
- Dingel, Jonathan I., and Brent Neiman.** 2020. "How Many Jobs Can Be Done at Home?" *Journal of Public Economics* 189: 104235.
- Dingel, Jonathan I., and Felix Tintelnot.** 2020. "Spatial Economics for Granular Settings." NBER Working Paper 27287.
- Disdier, Anne-Celia, and Keith Head.** 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *Review of Economics and Statistics* 90 (1): 37–48.
- Donaldson, Dave.** 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108 (4–5): 899–934.
- Donaldson, Dave, and Richard Hornbeck.** 2016. "Railroads and American Economic Growth: A 'Market Access' Approach." *Quarterly Journal of Economics* 131 (2): 799–858.
- Donaldson, Dave, and Adam Storeygard.** 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives* 30 (4): 171–98.
- Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender.** 2020. "All Aboard: The Effects of Port Development." NBER Working Paper 28148.
- Duranton, Gilles, and Matthew A. Turner.** 2011. "The Fundamental Law of Road Congestion: Evidence from US Cities." *American Economic Review* 101 (6): 2616–52.
- Eaton, Jonathan, and Samuel Kortum.** 2002. "Technology, Geography and Trade." *Econometrica* 70 (5): 1741–79.
- Eckert, Fabian, and Michael Peters.** 2022. "Spatial Structural Change." NBER Working Paper 30489.
- Faber, Benjamin, and Cecile Gaubert.** 2019. "Tourism and Economic Development: Evidence from Mexico's Coastline." *American Economic Review* 109 (6): 2245–93.
- Fajgelbaum, Pablo D., and Cecile Gaubert.** 2018. "Optimal Spatial Policies, Geography and Sorting." *Quarterly Journal of Economics* 135 (2): 959–1036.
- Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal.** 2020. "The Return to Protectionism." *Quarterly Journal of Economics* 135 (1): 1–55.
- Fajgelbaum, Pablo D., and Edouard Schaal.** 2020. "Optimal Transport Networks in Spatial Equilibrium." *Econometrica* 88 (4): 1411–52.
- Fan, Jingting, Yi Lu, and Wenlan Luo.** 2021. "Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters." *Review of Economics and Statistics*. [https://doi.org/10.1162/rest\\_a\\_01084](https://doi.org/10.1162/rest_a_01084).
- Fuchs, Simon, and Woan Foong Wong.** 2022. "Multimodal Transport Networks." Federal Reserve Bank of Atlanta Working Paper 2022-13.
- Fujita, Masahisa, Paul Krugman, and Anthony J. Venables.** 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.
- Ganapati, Sharat, Woan Foong Wong, and Oren Ziv.** 2020. "Entrepôt: Hubs, Scale, and Trade Costs." NBER Working Paper 29015.
- Gaubert, Cecile, Patrick M. Kline, and Danny Yagan.** 2021. "Place-Based Redistribution." NBER Working Paper 28337.
- Glaeser, Edward L., and Joshua D. Gottlieb.** 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983–1028.
- Head, Keith, and Thierry Mayer.** 2013. "Gravity Equations: Workhorse, Toolkit, and Cookbook." In *Handbook of International Economics*, Vol. 4, edited by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, 131–95. Amsterdam: North-Holland.
- Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm.** 2020. "The Making of the Modern Metropolis: Evidence from London." *Quarterly Journal of Economics* 135 (4): 2059–2133.
- Heiland, Inga, Andreas Moxnes, Karen Helene Ulltveit-Moe, and Yuan Zi.** 2019. "Trade from Space: Shipping Networks and the Global Implications of Local Shocks." CEPR Discussion Paper 14193.
- Helpman, Elhanan.** 1998. "The Size of Regions." In *Topics in Public Economics: Theoretical and Applied Analysis*, edited by David Pines, Efraim Sadka, Itzhak Zilcha, 33–54. Cambridge, UK: Cambridge University Press.

- Henderson, J. Vernon, Adam Storeygard, and David N. Weil.** 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028.
- Jaworski, Taylor, and Carl T. Kitchens.** 2019. "National Policy for Regional Development: Historical Evidence from Appalachian Highways." *Review of Economics and Statistics* 101 (5): 777–90.
- Kleinman, Benny, Ernest Liu, and Stephen J. Redding.** 2021. "Dynamic Spatial General Equilibrium." NBER Working Paper 29101.
- Kline, Patrick, and Enrico Moretti.** 2014. "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129 (1): 275–331.
- Krugman, Paul.** 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99 (3): 483–99.
- Lee, Sun Kyoung.** 2022. "When Cities Grow: Urban Planning and Segregation in the Prewar US." Unpublished.
- Manson, Steven.** 2020. "IPUMS National Historical Geographic Information System: Version 15.0." IPUMS. <http://doi.org/10.18128/D050.V15.0> (accessed March 1, 2023).
- McCallum, John.** 1995. "National Borders Matter: Canada-U.S. Regional Trade Patterns." *American Economic Review* 85 (3): 615–23.
- Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg.** 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review* 108 (12): 3855–90.
- Nordhaus, William, and Xi Chen.** 2006. "Geographically Based Economic Data (G-Econ)." Yale University. <https://gecon.yale.edu/> (accessed March 1, 2023).
- Peters, Michael.** 2022. "Market Size and Spatial Growth—Evidence from Germany's Post-war Population Expulsions." *Econometrica* 90 (5): 2357–96.
- Redding, Stephen J.** 2016. "Goods Trade, Factor Mobility and Welfare." *Journal of International Economics* 101: 148–67.
- Redding, Stephen J., and Esteban Rossi-Hansberg.** 2017. "Quantitative Spatial Economics." *Annual Review of Economics* 9: 21–58.
- Redding, Stephen, and Anthony J. Venables.** 2004. "Economic Geography and International Inequality." *Journal of International Economics* 62 (1): 53–82.
- Roback, Jennifer.** 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78.
- Rosen, Sherwin.** 1979. "Wage-Based Indexes of Urban Quality of Life." *Current Issues in Urban Economics*: 74–104.
- Sethian, J. A.** 1999. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, Vol. 3. Cambridge, UK: Cambridge University Press.
- Severen, Christopher.** 2021. "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification." *Review of Economics and Statistics*. [https://doi.org/10.1162/rest\\_a\\_01100](https://doi.org/10.1162/rest_a_01100).
- Steinwender, Claudia.** 2018. "Real Effects of Information Frictions: When the States and the Kingdom Became United." *American Economic Review* 108 (3): 657–96.
- Tombe, Trevor, and Xiaodong Zhu.** 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109 (5): 1843–72.
- Topalova, Petia.** 2010. "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India." *American Economic Journal: Applied Economics* 2 (4): 1–41.
- Tsitsiklis, J. N.** 1995. "Efficient Algorithms for Globally Optimal Trajectories." *IEEE Transactions on Automatic Control* 40 (9): 1528–38.
- US Department of Transportation, Bureau of Transportation Statistics; and, US Department of Commerce, US Census Bureau.** (2020-08). 2017 Commodity Flow Survey Datasets 2017 CFS Public Use File (PUF). Available at: [https://data.nber.org/transportation/CommoditiesData/States/states\\_bilateral\\_trade2017.dta](https://data.nber.org/transportation/CommoditiesData/States/states_bilateral_trade2017.dta).
- Zárate, Román D.** 2022. "Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City." World Bank Policy Research Working Paper 9990.