What Explains U.S. House Prices? Regional Income Divergence*

Greg Howard[†] Jack Liebersohn[‡]

January 4, 2019

Abstract

A simple measure of regional income divergence explains much of the variation in U.S. house prices since 1939. We develop an asset-pricing theory to explain why. House prices reflect the discounted value of expected rents, which reflect expected incomes. Higher expected regional income divergence increases the house price premium in rich areas. This raises average prices because house prices in poor areas are largely determined by construction costs. In addition to explaining average prices, our model explains several facts about the housing market, including regional variation in prices, the relationship between rents and prices, and patterns of net inter-state migration.

JEL Codes: R31, G12, E22

Keywords: regional divergence, house prices

*We would like to thank Hansen Shao for his research assistance, as well as everyone who saw Figure 1 before this was a paper and told us it was interesting enough to write down: Vivek Bhattacharya, Ben Roth, and Dejanir Silva. Ye Li, Victor Lyonnais and Nina Karnaukh provided helpful comments. All errors are our own.

[†]University of Illinois, Urbana-Champaign, glhoward@illinois.edu.

[‡]Ohio State Fisher, liebersohn.1@osu.edu

1 Introduction

House prices play a prominent role in explaining many macroeconomic phenomena including business cycles, the decline in labor share, the slowdown in U.S. growth, the rise in wealth inequality, and the stagnation of real wages.[1] Yet the determinants of aggregate house price movements remain mysterious. In this paper, we show that a simple measure of regional divergence does a good job of explaining U.S. house prices over the last eighty years. We also show that this finding is consistent with an intuitive model of the national housing market.

Our model builds upon several well-established facts of housing markets. Housing is an asset, so the price of a house is related to the present value of its future expected rents (Poterba, 1984). A significant fraction of the value of a house is due to its location (Davis and Heathcote, 2007). Higher rents in a specific area reflect a premium to live there, so states with higher income typically also have higher rents (Rosen, 1979; Roback, 1982). In times of high regional income convergence, that premium is expected to erode, whereas when there is regional divergence, the rent premium is expected to grow so house prices are higher. These facts link income divergence to the cross-section of house prices. Because house prices are systematically less volatile in low-income areas, in part due to an elastic housing supply in those regions (Glaeser, Gyourko, and Saiz, 2008), regional divergence increases prices in higher-priced areas a great deal while hardly decreasing prices in lower-priced regions. Therefore, the speed of divergence matters for the average price of housing nationally.

Our measure of regional divergence, κ_t, comes from a regression of the 10-year growth rate of U.S. states on their lagged per capita incomes. When the coefficient on lagged income is positive, rich states are growing faster than poor states. Specifically, we estimate the κ_t's

[1] See for example, Leamer (2007) for how housing is a leading indicator of business cycles, Mian and Sufi (2009, 2011, 2012) and Mian, Rao, and Sufi (2013) for the effect of housing on the Great Recession, Rognlie (2016) for the effects of housing on the decline in the labor share, Mahtani and Miller (2017) for the effect of housing on the stagnation of real wages, Aladangady, Albouy, and Zabek (2017) for the effect of housing on wealth inequality, Hsieh and Moretti (2015) and Herkenhoff, Ohanian, and Prescott (2018) for the effect of housing on the slowdown of U.S. growth.

using the following regression:

$$\Delta_{10} \log(\text{Per Capita Personal Income}_{st}) = \sum_t \kappa_t \log(\text{Per Capita Personal Income}_{s,t-10}) + \eta_t + \epsilon_{st}$$

where s denotes a U.S. state, t is a year, and η_t is a year fixed effect. Log Per Capita Personal Income is measured by the Bureau of Economic Analysis, and obtained through FRED.

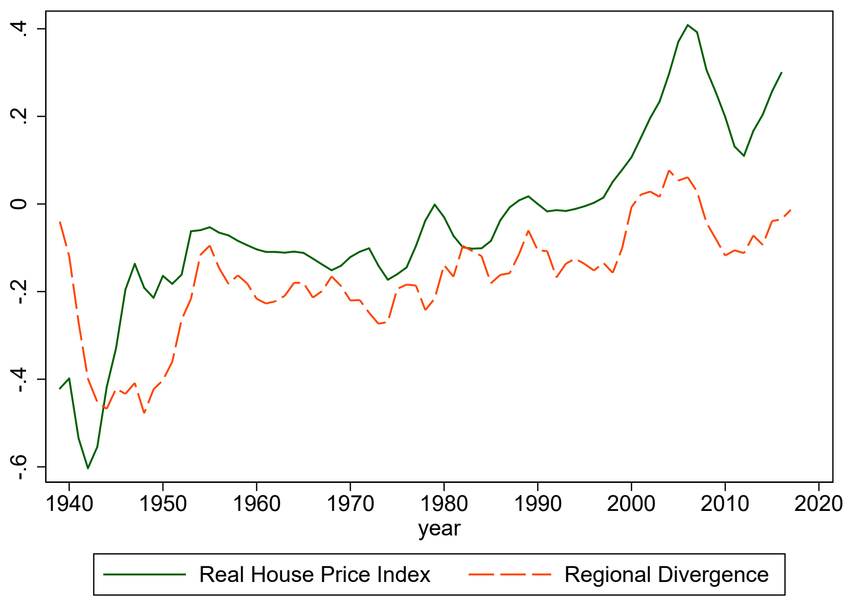

Figure 1: Regional Divergence and Real House Prices. Nominal house prices come from the Jorda, Shularick, and Taylor (2017) Macro History database and are deflated by CPI. Regional divergence is measured in each year by a cross-state regression of 10-year growth rates in per capita personal income (from the BEA) on 10-year lagged per capita personal income. When it is higher, richer states are growing faster than poorer states.

The convergence measure is plotted, along with real U.S. house prices, in Figure 1.[2] Consistent with the literature, divergence has risen over time and is now close to zero (Giannone, 2017; Ganong and Shoag, 2017). What is striking about Figure 1 is that regional divergence parallels real house prices since before World War II.[3]

In the rest of this paper, we show why this relationship is so strong. Our theory relies

[2]House prices are taken from the Jorda, Shularick, and Taylor (2017) Macro History Database, and are deflated by U.S. CPI in that same database. See Knoll, Schularick, and Steger (2017).

[3]We provide a more detailed description of this correlation in Appendix A.

on two main arguments. The first is that differences in house prices across states are driven by the differences in the present discounted value of rents across states, which in turn reflect differences in the present discounted value of income. This assumption implies that cross-sectionally, house prices should be increasing in wages, and that that relationship should be stronger when there is more regional divergence. The data support these implications.

The second argument is that the increase in the dispersion of house prices increases house prices overall. In regions that are growing, house prices will be bounded below by the cost of construction, an assumption that is supported empirically. If construction anchors prices in poorer regions, then any increase in dispersion will also raise the average. This argument also implies that national rents should be related to the current dispersion of regional income, a prediction we confirm in the data, and which we find interesting in and of itself.

If income divergence explains house prices, what explains income divergence? Following Barro and Sala-I-Martin (1991), we believe that the most important factor is differences in regions' exposure to technological and industry-specific shocks. We provide evidence for this using a shift-share measure to predict actual income growth using a combination of historical regional industry shares and national industry growth. This measure predicts divergence well. Our view that technological shocks matter for local incomes is also supported by Giannone (2017).

Ganong and Shoag (2017) also point out that both regional divergence and house prices have been rising.[4] The main difference to this paper is that they propose a relationship in which the causality runs the other way: high house prices due to supply regulations lead to less regional convergence. In Section 5, we discuss ways to disentangle the two stories. A key difference between the two models concerns the relationship between house prices and migration. The model in Ganong and Shoag (2017) implies that populations should flow away from places with rising house prices, whereas our model implies the opposite. The relationship between price fluctuations and migration appear to support our model more

[4]They note that both variables have been trending upward, but we believe we are the first to document the close relationship in the time series between these two series.

strongly. However, while we think our story is the primary reason that house prices and divergence are related, we do find support for one of the main points of Ganong and Shoag (2017), that housing supply elasticities have declined over time.

Most closely related to our work is Van Nieuwerburgh and Weill (2010), which shows that income changes can help explain why house price dispersion increased from 1975 to 2007. We build on what they do, showing that regional divergence also affects house price dispersion through the expectation of future income divergence, and we also tie house price dispersion to house prices overall.

A recent literature has focused on explaining the 2000's housing boom, focused mainly on the expansion of credit and speculation. For example, Kaplan, Mitman, and Violante (2017) shows the importance of expectations to explain house price movements in a calibrated DSGE model, and Mian and Sufi (2018) show that changes in credit supply affect expectations which affect house prices at the micro-level. Our paper provides highlights a critical reason for changes in expectations, based on recent economic growth. We do not think that our findings in any way suggest that credit does not play a role in explaining the cross-sectional house price movements, and that credit could have a large effect on the aggregate if it changes expectations along the right dimensions. However, our measure of regional divergence does explain a large chunk of the recent boom, consistent with the fact that income explains much of the cross-sectional variance and changes in house prices (Van Nieuwerburgh and Weill, 2010).

Several papers have studied how a national housing demand shock heterogeneously affects different regions (Mian and Sufi, 2009, 2011; Mian, Rao, and Sufi, 2013; Guren, McKay, Nakamura, and Steinsson, 2018). Because there is a positive association between local house prices and local consumption, several papers have found that increases in regional or national house prices cause higher economic activity in inelastic areas.[5] Some implications of this are closely related to our findings, but we can again distinguish our channel by focusing

[5]Inelastic areas are highly correlated with richer areas.

on migration. National housing demand shocks, coming from changing tastes or changing credit availability, for example, predict that populations should increase more in the poorer, more-elastic areas of the country.[6] But in the data, populations increase in rich areas during times of high regional divergence, as our channel implies. Of course, a positive effect of house prices on consumption would amplify our channel, because richer states would grow even faster when their house prices increase.

For our channel to predict the empirical result shown in Figure 1, there must be a close relationship between past divergence and expected future income divergence. Estimates using a survey of household income expectations support such a relationship. In particular, a divergence measure that uses expected income changes rather than past changes is very similar to Figure 1. This is expected because of the high persistence of regional income growth shocks, as measured by Glaeser and Nathanson (2017) and Head, Lloyd-Ellis, and Sun (2014). The former paper also provides evidence that housing price expectations are affected by recent experiences, which is analogous to our assumption about recent income experiences and income expectations.[7]

In addition to research on income convergence, our paper is related to the large literature on house prices and expectations. Several papers, such as Case and Shiller (1988), have linked realized house price growth to expectations of future house prices. We contribute to the literature on house price expectations by showing a theoretical and empirical relationship between house prices and expectations of future *incomes*, rather than prices per se.

[6]There are two reasons for this. One is that if housing demand increases, areas with cheap housing become more attractive. Second, if there is a general increase in housing demand, cities with elastic housing supply are better able to expand their housing stock. While it is theoretically possible that the change in local economic conditions is so strong that it overwhelms the direct effect of a change in housing demand, this indirect channel would have to be implausibly large compared to existing empirical estimates.

[7]Rozsypal and Schlafmann (2017) and Nagel (2012) show further evidence that expectations are influenced by recent income experience. Kwan and Cotsomitis (2004) demonstrates that it does carry over into real behavior.

2 Theory

The mechanism through which regional divergence causes high house prices is illustrated in the following toy model. Consider two regions, $i \in \{1, 2\}$, where Region 1 has higher income. Consumers are risk-neutral and we assume there are no bubbles, so house prices are given by the present discounted value of rents:

$$p_{t,i} = \sum_{s=t}^{\infty} \frac{1}{(1+R)^{s-t}} r_{s,i} \tag{1}$$

In each period, perfectly mobile consumers with Cobb-Douglass utility over a tradable good and housing choose where to live, which determines their wages, leading to the following indifference condition:[8]

$$\log w_{t,1} - \alpha \log r_{t,1} = \log w_{t,2} - \alpha \log r_{t,2} \tag{2}$$

Wages converge exponentially in logs by parameter κ, which corresponds to the regional divergence line in Figure 1, divided by 10.

$$\log w_{t,1} - \log w_{t,2} = (1+\kappa)(\log w_{t-1,1} - \log w_{t-1,2}) \tag{3}$$

Finally, in Region 2, house prices are given by the cost of construction, which we normalize.

$$p_{2,t} \equiv 1 \tag{4}$$

Proposition 1: Average house prices are increasing in κ.

Proof: By (1)-(4), house prices in Region 1 are given by

$$p_{1,t} = \frac{1+R}{R} \sum_{s=0}^{\infty} \frac{1}{(1+R)^{s-t}} \left(\frac{w_{1,t}}{w_{2,t}}\right)^{(1+\kappa)^{s-t}/\alpha} \tag{5}$$

[8]This indifference follows in the footsteps of Rosen (1979) and Roback (1982).

That expression is increasing in κ. Because house prices in Region 1 are increasing in κ, and house prices in Region 2 are constant, overall house prices are also increasing in κ. □

If $\frac{w_1}{w_2} \approx 1$, then we can approximate the change in log house prices:

$$\log \frac{p_1}{p_2} \approx \frac{1}{\alpha} \frac{R}{R-\kappa} \log \frac{w_1}{w_2}$$

So as κ increases, we would expect house prices to increase as well, and if it gets close to R, the effect would be stronger. In our model, there is only a small role for R to play in explaining house price movements, and the effect of the interest rate on house prices changes signs depending on κ.

Proposition 2: Fix the wage growth of city 1. Suppose city 1 has positive housing supply elasticity,[9] and that the housing market clears. Then population growth in city 1 is increasing in regional divergence.

Proof: The growth rate of housing supply of city 1 is increasing in p_1, so therefore is increasing in κ by Proposition 1. At the same time, the growth rate of r_1 is increasing in κ, so the growth in per capita housing is slowing. Because growth rate of total housing is larger and the growth rate of per capita housing is smaller, the growth rate in population must be higher. □

Proposition 2 is important because it will help us distinguish our story from reverse causality, which hinge on the idea that regional divergence is high because population movements to city 1 are low.

3 Regional Divergence

In the next two sections, we show empirical support for our theory. First, in this section, we focus on the causes of regional divergence and its effects on expectations. In the next

[9]We define positive elasticity to mean that the growth rate of housing quantity is increasing in the price of housing.

section, we focus on how regional divergence affects the housing market.

This section shows two important facts about regional divergence. First, we show that a shift-share measure of divergence created using historical industry exposure is very similar to the main divergence measure. This finding reduces concerns that our findings are due to contemporaneous location-specific factors, such as changes in housing regulation, and justifies treating it as an exogenous variable in our model. Second, we construct a measure of *expected* divergence from survey data on consumer expectations and show that actual and expected divergence are similar.

Regional divergence is an inherently aggregate phenomenon, making it difficult to identify its origins. We think the major factors behind regional divergence are determined by a combination of technology and factors exogenous to the United States. For example, Giannone (2017) demonstrates that skill-biased technological change can explain much of the decline in regional convergence since 1980. Barro and Sala-I-Martin (1991) argue that many of the fluctuations in regional convergence can be explained by nation-wide industy changes, such as oil in the 1980s. Acemoglu and Restrepo (2017) or Autor, Dorn, and Hanson (2013) document large geographic differences in the technological effects of robots and the effects of Chinese import competition.

To study the role of industry composition, we use a shift-share methodology to predict each state's GDP growth using only lagged industry shares and national rates of growth per employee by industry.[10] This measure thus purges from the divergence measure any idiosyncratic, location-specific factors, such as local changes in local housing supply regulation or local construction costs. Panel regressions of 10-year state income growth on predicted

[10] We measure national industry GDP and employment at the national and state levels using historical data from the BEA. For each state and each year, we multiply 10-year national industry growth rates by 10-year lagged industry shares of state GDP and sum this over all industries to predict state GDP growth per employee. National industry GDP data is available from 1947, so 1957 is the first year we can measure 10-year industry growth rates. Industry-specific regional accounts data is not available before 1963, however, so we use 1963 industry shares for 1957-1963. Because SIC industry data is not available after 1997, we use the SIC-NAICS concordance file created for Autor, Dorn, and Hanson (2013) to concord industry GDP and employment. This file, which is available on David Dorn'as web site, is concordance from 6-digit NAICS codes to 4-digit SIC codes, which we aggregate to the level of NIPA industries using NAICS national employment by industry in 1998 .

growth have an R^2 of 0.77, which increases to 0.91 with year fixed effects, suggesting that income growth is mostly explained by industry exposure.

We also use predicted income growth instead of actual income growth to measure regional divergence. We do this using the same equation as before but with predicted income plugged in for realized income. The actual and predicted divergence series, shown together in Figure 2, are similar to each other.[11] Interestingly, the early-2000s housing boom and bust is predicted a few years in advance by the shift share measure.[12] Overall, we believe the evidence suggests that industry-specific shocks actually explain regional growth quite well.

The similarity of the shift-share measure of divergence and the measure of actual regional divergence has two implications for understanding what causes divergence. First, it provides support for the view that plausibly-exogenous industry changes are an important factor. Second, it means that our results would also be hard to explain from contemporaneous, region-specific factors alone, like changes in regulation.

What ultimately matters for the mechanism described in Section 2 is not realized divergence, which is easily measurable, but expectations of future regional divergence, which are hard to measure. Ideally, we would have expectations data for the growth rates of multiple states from the same person. If this data exists, we are not aware of it. Instead, we compare expectations of future personal income growth across people in different states. We find that this measure of expected divergence is closely related to actual divergence, providing support for our mechanism.

The data on expected incomes comes from the Michigan Survey of Consumers. This nationally-representative consumer survey provides, to the best of our knowledge, the longest-

[11] Dropping sectors related to housing, i.e. construction and real estate finance, does not change the results visually. Neither does using a "leave-one-out" construction. However, our state-industry level data begins in 1963, so we cannot construct a "leave-one-out" measure before 1973, and therefore show the shift-share based on national shifts, including the state's own contribution. Graphs of these alternative measures are in the Appendix.

[12] Another difference between the shift-share measure and the actual measure is that the shift-share predicts more divergence at the beginning of the sample than actually occurred. This is probably an artifact of the data construction: The pre-1973 divergence estimate uses pre-1963 shift shares, which use future industry composition, as described in Footnote 10.

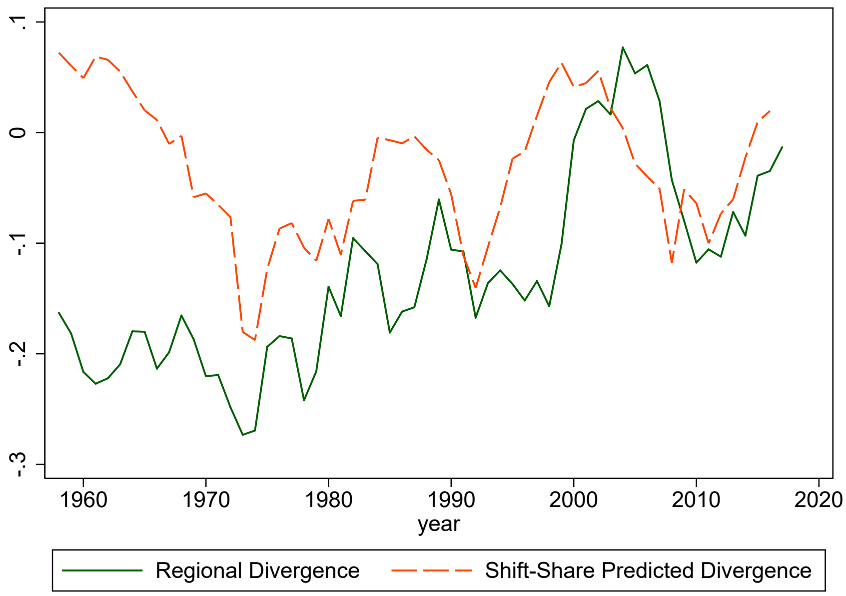

Figure 2: Regional GDP Divergence and Shift-Share Prediction of GDP Divergence. Regional GDP divergence is measured in each year by a cross-state regression of 10-year growth rates in per capita personal income on 10-year lagged per capita personal income. Predicted divergence in GDP per employee is measured using a cross-state regression of 10-year growth rates in predicted GDP per employee on lagged 10-year per capita income, where predicted GDP per employee is created using a shift-share methology.

running measure of household income expectations and consumer sentiment in the United States. It includes several questions about recent and expected household incomes, business conditions, employment, etc., and has done so since 1978. We use a custom, respondent-level sample of households' income expectations for the next year. The respondent-level data is aggregated to a panel of state-by-year income expectations, which we take a running average of to smooth.[13] We then create a measure of *expected* income divergence that is analogous to our measure of realized income divergence but uses expected rather than actual income. Specifically, we regress the expectations of future personal income growth on lagged ten-year state income and take the coefficients from each year.

[13]The specific question we use asks "By about what percent do you expect your (family) income to (increase/decrease) during the next 12 months?" This question is known as INEX in the survey. We drop DK and missing values and censor the remaining observations at the 90th percentiles, and then take state-by-year averages weighted by sample weights. Our estimates are robust to censoring at other values. After taking the state-by-year average of income expectations, we calculate the centered, 5-year running average of the series. The results are very similar with running averages using fewer or more years.

Our divergence expectations measure is shown in Figure 3. Expected divergence is very similar to realized divergence for the years that both series are available. One reason for this is that households may correctly understand that local income shocks are persistent, and therefore use recent income growth to project future income growth. Evidence from both psychology and economics suggests that households form expectations in an extrapolative way in many domains (Bordalo, Gennaioli, and Shleifer, 2018). While extrapolation may not always be rational, it is a very reasonable way for households' to form expectations of future income in particular, because local income shocks are persistent (Glaeser and Nathanson, 2017; Head, Lloyd-Ellis, and Sun, 2014). For this reason, we do not find it surprising that expected and realized divergence are so similar. Nonetheless, the similarity of these lines is important because it gets to the heart of our mechanism.[14]

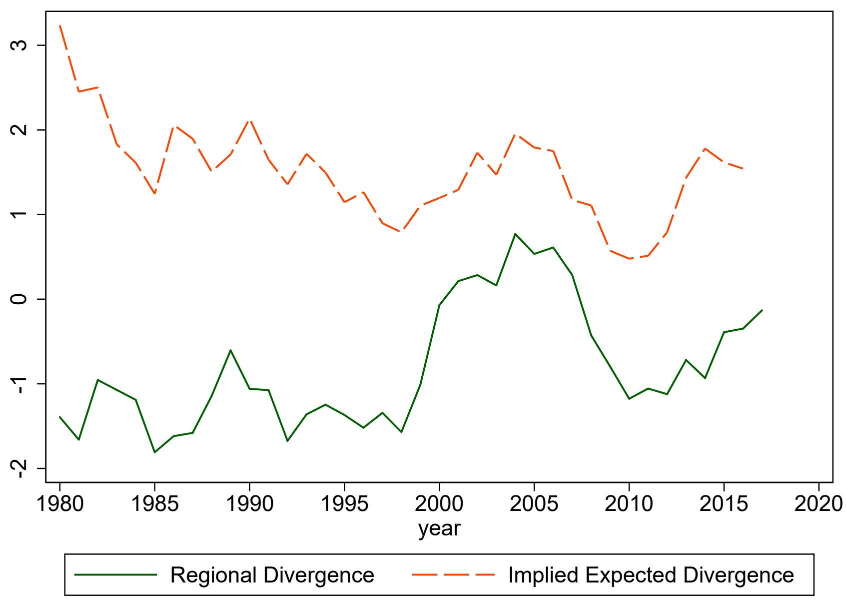

Figure 3: Regional Divergence and Expectations of Future Regional Divergence. Divergence expectations are created by regressing one-year income expectations on 10-year lagged incomes. Income expectations are from the Michigan Survey of Consumers.

[14]An obvious feature of Figure 3 is that expectations are higher than regional divergence. We do not take a stand on why these are generally higher, but it is helpful later in explaining that the regression of house prices on income leads to a higher coefficient in most years than the regression of rents on income. See Figure 4. The ratio of these should be $\frac{R}{R-\kappa}$, which is bigger than one only if the expectations of future κ are positive.

4 House Price Dispersion

The second part of our argument is that expectations of future regional divergence lead to more disperse and higher house prices. We break this claim into three predictions, check these three predictions in the data, and find that all three hold. We do not know of previous research which has made or verified the first or third predictions. The second prediction is not surprising given other work done on housing supply elasticities (e.g. Saiz, 2010), but we are not aware of other papers that have focused on income.

Prediction 1. The cross-sectional relation between house prices and income should be positive, and should be more positive when κ is high. We would expect a less-volatile positive relationship with rents.[15]

Prediction 2. House prices in poorer areas should be less volatile than those in richer areas, and when the difference between high income and low income house prices are higher, average house prices are higher.

Prediction 3. The cross-sectional dispersion of income should be a good predictor of national rents. This is because, like house prices, rents in poorer more-elastic areas are not volatile, so when dispersion increases, the average rent does as well.

Prediction 1 is shown in Figure 4. The orange line is created by regressing log house prices on 10-year-lagged log income, year-by-year, and plotting the coefficient.[16] From 1940-2000, we use Census data on median house values, and from 1975-2016, we use the housing series created by Davis and Heathcote (2007). The thick gray line is regional divergence, measured as in Figure 1. The regression coefficient, as expected, has been trending up over time, similar to regional divergence, and boomed in the early 1990s and the 2000s, as did regional convergence and house prices. Income is thus a good predictor of house prices. The R^2 of a regression of prices on income within a given year is typically about 0.5, and every

[15]If rents were truly only a flow value, they would not be expected to move at all. But since a rental contract typically does bind a renter to a location for some amount of time, and other moving costs exist, it may fluctuate a little bit.

[16]We use ten-year lagged income to lessen correlations due to reverse causality, but the results are similar using current log income.

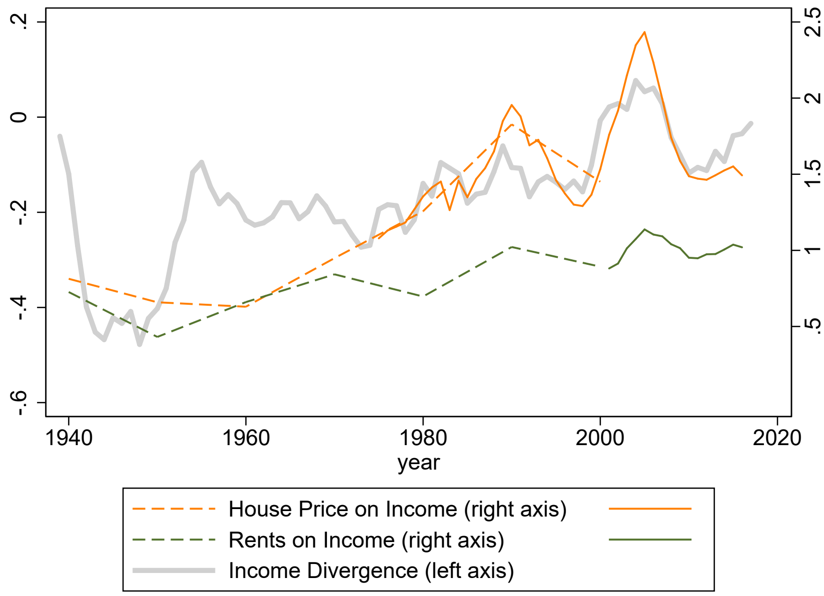

Figure 4: Regional Divergence and House Price Dispersion. House Prices come from the Census (dashed line) or from Davis and Heathcote (2007) (solid line). Gross rents come from the Census (dashed line) or from the American Community Survey (solid line).

year is statistically significant with $p < 0.001$. While rents have also fluctuated in similar ways, their fluctuations are much smaller.

For Prediction 2, we look at movements in house prices since 1975 in different counties. Using BEA statistics, we split counties into four bins (roughly equal in population) by their per capita income.[17] Figure 5 reveals that house prices in the poorest quartile have increased much less and are much less volatile that prices in the middle quartiles, and especially compared to the richest quartile. In fact, real house prices for this quartile in 2016 are 2.5 percent higher than they were in 1975, and the most they changed was 22 percent in 2007.

The assumption that house prices in poorer areas are pinned down by construction costs is supported by the fact that these areas have a more elastic housing supply. Howard and Liebersohn (2018) construct a measure of the elasticity of house prices for every commuting

[17]By using counties, we can capture a lot more disaggregated variation. The same picture holds true for states, but there are larger differences with counties. The FHFA does not keep data on every county, especially going all the way back to 1975, so there is certainly some bias because larger and richer counties are more likely to show up in the sample. The line for the poorest quantile would likely be even flatter if this data were available.

zone in the United States, which we use. Consistent with what we would expect based on Figure 5, we estimate a correlation of -0.5 between average housing supply elasticity and log average state income, indicating that housing supply is more elastic in lower-income states.

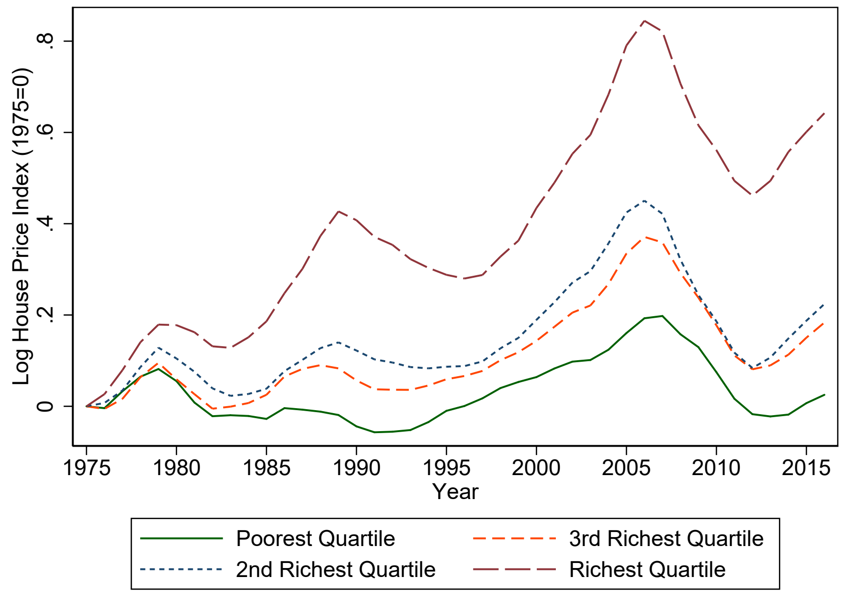

Figure 5: House price movements by quartile of per capita income in 1975. Per capita income measured by the BEA. House price indices from the FHFA, deflated by CPI.

Prediction 3 is shown in Figure 6. To measure dispersion, we simply use the standard deviation of log income across states, within a year.[18] This is shown in orange. In green, we show the CPI measure of rent of primary residence, normalized by overall CPI. These lines follow one another since the 1930's. We find this result to be particularly interesting in and of itself. It also strongly supports our central claim because it shows income dispersion has important implications for the levels of rents, and therefore expectations of future dispersion should have important implications for current house prices.

[18]Our model does not specify a specific measure of regional income dispersion. Various other measures have very similar trends.

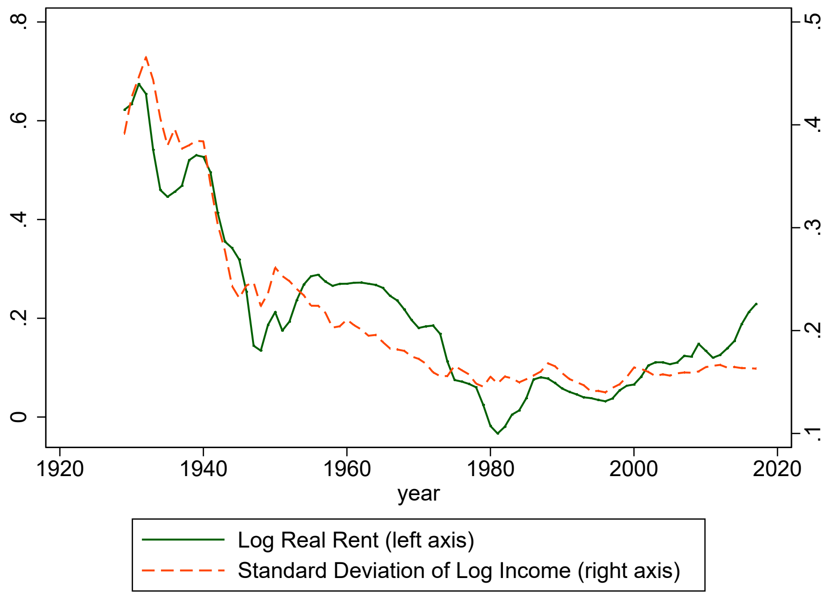

Figure 6: Income Dispersion and Rental Prices. Per capita personal income by state comes from the Census, and dispersion is calculated as the within-year standard deviation. Rents come from the BLS, and are normalized by CPI.

5 Disentangling Reverse Causality

Ganong and Shoag (2017) and Herkenhoff, Ohanian, and Prescott (2018) present models in which housing supply regulations cause high house prices, which slows regional convergence. Both papers note the long-term trends in the two series and interpret the causality in reverse of the theory presented here. Some additional facts are helpful to help disentangle our story from theirs. Furthermore, the facts that we discuss show that the theories behind Mian and Sufi (2009), and Guren, McKay, Nakamura, and Steinsson (2018) cannot completely explain our results.

Fact 1. Population moves towards richer states when regional convergence is low. For the models in Ganong and Shoag (2017) and Herkenhoff, Ohanian, and Prescott (2018), building on Blanchard and Katz (1992), the story is that high house prices prevent adjustment, i.e. people cannot move to high income areas. In contrast, in our model, times of higher regional divergence are exactly when people most want to move to richer areas. So if population grows more in richer states during times of regional divergence, it would support our story, whereas if population grows less in richer states during times of regional divergence, it would support their story.

As it turns out, in Figure 7, population movements are towards richer states exactly at times when house prices and regional divergence are highest. The series in green consists of the coefficients of the regression:

$$\Delta_{10} \log \text{Population Change}_{st} = \sum_t \beta_t \log \text{Per Capita Personal Income}_{s,t-10} + \eta_t + \epsilon_{st}$$

When this series is high, people are moving to high income states. While it is true that population movement toward high income states has declined over the entire time, the fluctuations support our story. People move to high-income areas when there is regional divergence, especially in the late 1980s, and a smaller bump during the 2000s, and again during the most recent rise.

A related story is that higher house prices have caused additional sorting, leading to regional divergence because more productive workers live in more expensive areas. To see if this might cause our results, we look at the population movements of only people over 25 without a college degree. We run the same regression but use an estimate of the population over 25 with less than four years of college.[19] Interestingly, there is no bump during the 2000s, but a large increase in the last few years of the data, during the latest regional divergence and increase in house prices. So while there has been more sorting in recent years, we do not think that it corresponds well to some of the more high-frequency movements in either regional convergence or house prices.

We also provide some evidence that is consistent with Ganong and Shoag (2017). Figure 7 shows that overall migration has trended down since about 1980, even as both house prices and overall divergence have increased. These patterns are consistent with their finding that

[19] We use data from the CPS to estimate the share of people that fit both those categories, and multiply that by the population estimates.

restrictions on construction have both raised house prices and decreased net migration.

Fact 2. Regional divergence is a leading indicator of house prices. This makes sense if regional divergence helps people form expectations of future rents. It would be hard to reconcile with reverse causality. This fact is most apparent in the recent boom, where the regional divergence line rise and peaks before aggregate house prices. To make our claim about timing more formal, we run a vector auto-regression and do a Granger causality test. For this test, we use a different version of regional convergence that is noisier but less backward looking, regressing one-year changes in log per capita personal income on a one-year lag of log per capita personal income. Because that measure and real house prices are non-stationary, we take first differences and run a VAR, using six lags as selected by the Akaike Information Criterion (Akaike, 1974).[20] We reject the hypothesis that regional divergence does not Granger-cause house prices with $p = .002$. We cannot reject the hypothesis that house prices do not Granger-cause regional divergence.

Fact 3. Under the reverse-causality story, we would expect the rent-income gradient to be as steep as house prices. This is because these models make the same predictions for rents as for house prices. In reality, the rent-income gradient is not as steep as the house price gradient, as we showed in Figure 4. Our theory correctly predicts a bigger movement in house prices than rents.

6 Conclusion

The main result of our paper is that aggregate U.S. house prices and regional divergence move together in the time series. We explain this with is an asset pricing framework in which housing in high-income areas is more valuable if the rest of the country is not expected to catch up. We present evidence for this theory and show that alternative theories cannot explain all of the available facts.

In addition to providing a new explanation for aggregate house price movements, this

[20] The Granger causality test gives a p-value less than .02 for all lags between 2 and 20 years.

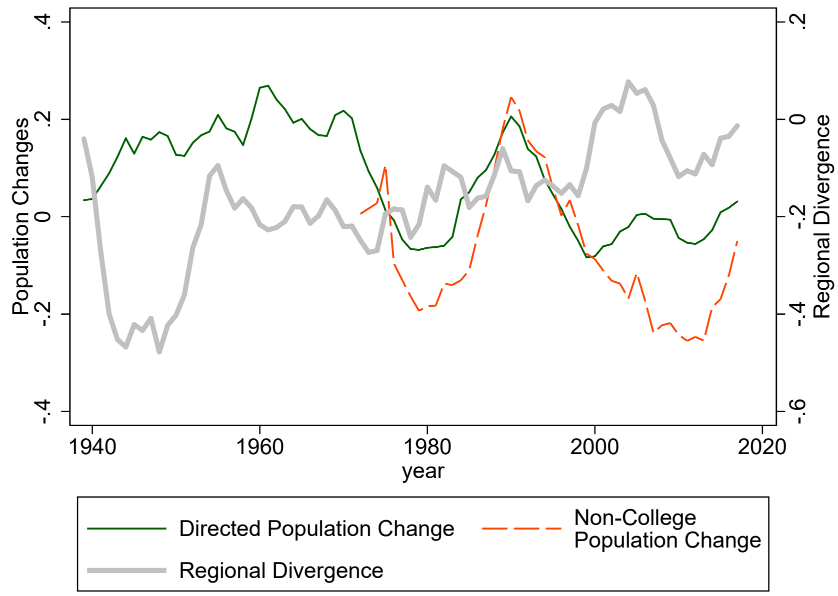

Figure 7: Regional Convergence and Population Movements Toward Higher Income States. Income movement series comes from a regression of population flows on lagged per-capita incomes.

paper provides a new understanding of macroeconomic phenomena that are directly impacted by housing. As an example, it cements the importance of house prices for inequality. In the toy model, if the averge income is always 1, then the welfare of consumers is decreasing in κ, meaning that as regional divergence increases, the welfare of non-homeowners everywhere declines. All of the declines are a transfer to homeowners in Region 1.

Another area our paper sheds light on is the role housing plays in business cycles. Several recent papers have argued that housing wealth affects aggregate consumption and hence income. By contrast, in our model, house prices are related to expectations of future income, which is an alternative channel mechanism whereby house prices and income may be linked. Interactions between these channels may be a fruitful area for future research.

References

Acemoglu, D., and P. Restrepo (2017): "Robots and jobs: Evidence from US labor markets," .

Akaike, H. (1974): "A new look at the statistical model identification," *IEEE transactions on automatic control*, 19(6), 716–723.

Aladangady, A., D. Albouy, and M. Zabek (2017): "Housing Inequality," .

Autor, D., D. Dorn, and G. H. Hanson (2013): "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 103(6), 2121–68.

Barro, R. J., and X. Sala-I-Martin (1991): "Convergence Across States and Regions," *Brookings Papers on Economic Activity*, 1991(1), 107–182.

Blanchard, O., and L. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018): "Diagnostic expectations and credit cycles," *The Journal of Finance*, 73(1), 199–227.

Case, K. E., and R. J. Shiller (1988): "The behavior of home buyers in boom and post-boom markets," *New England Economic Review*, (Nov), 29–46.

Davis, M. A., and J. Heathcote (2007): "The price and quantity of residential land in the United States," *Journal of Monetary Economics*, 54(8), 2595–2620.

Ganong, P., and D. Shoag (2017): "Why has regional income convergence in the US declined?," *Journal of Urban Economics*, 102, 76–90.

Giannone, E. (2017): "Skilled-biased technical change and regional convergence," Discussion paper, University of Chicago Working Paper.

Glaeser, E. L., J. Gyourko, and A. Saiz (2008): "Housing supply and housing bubbles," *Journal of Urban Economics*, 64(2), 198 – 217.

Glaeser, E. L., and C. G. Nathanson (2017): "An extrapolative model of house price dynamics," *Journal of Financial Economics*, 126(1), 147–170.

Guren, A. M., A. McKay, E. Nakamura, and J. Steinsson (2018): "Housing Wealth Effects: The Long View," Discussion paper, National Bureau of Economic Research.

Head, A., H. Lloyd-Ellis, and H. Sun (2014): "Search, liquidity, and the dynamics of house prices and construction," *American Economic Review*, 104(4), 1172–1210.

Herkenhoff, K. F., L. E. Ohanian, and E. C. Prescott (2018): "Tarnishing the golden and empire states: Land-use restrictions and the US economic slowdown," *Journal of Monetary Economics*, 93, 89–109.

Howard, G., and J. Liebersohn (2018): "The Geography Channel of House Price Appreciation," .

Hsieh, C.-T., and E. Moretti (2015): "Housing Constraints and Spatial Misallocation," Working Paper 21154, National Bureau of Economic Research.

Jorda, O., M. Shularick, and A. M. Taylor (2017): "Macrofinancial History and New Business Cycle Facts," *NBER Macroeconomics Annual 2016*, 31.

Kaplan, G., K. Mitman, and G. Violante (2017): "Consumption and House Prices in the Great Recession: Model Meets Evidence," .

Knoll, K., M. Schularick, and T. Steger (2017): "No price like home: Global house prices, 1870-2012," *American Economic Review*, 107(2), 331–53.

Kwan, A. C., and J. A. Cotsomitis (2004): "Can consumer attitudes forecast household spending in the United States? Further evidence from the Michigan Survey of Consumers," *Southern Economic Journal*, pp. 136–144.

Leamer, E. E. (2007): "Housing is the business cycle," in *Proceedings-Economic Policy Symposium-Jackson Hole*, pp. 149–233. Federal Reserve Bank of Kansas City.

Mahtani, S., and C. Miller (2017): "Don't Blame the Robots: How Housing Prices and Market Power Explain Wage Stagnation," *Foreign Affairs*.

Mian, A., K. Rao, and A. Sufi (2013): "Household Balance Sheets, Consumption, and the Economic Slump*.," *Quarterly Journal of Economics*, 128(4).

Mian, A., and A. Sufi (2009): "The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis," *The Quarterly Journal of Economics*, 124(4), 1449–1496.

——— (2011): "House prices, home equity–based borrowing, and the us household leverage crisis," *The American Economic Review*, 101(5), 2132–2156.

——— (2012): "What explains high unemployment? The aggregate demand channel," Discussion paper, National Bureau of Economic Research.

——— (2018): "Credit supply and housing speculation," Working paper, National Bureau of Economic Research.

Nagel, S. (2012): "Macroeconomic experiences and expectations: A perspective on the great recession," .

Poterba, J. M. (1984): "Tax subsidies to owner-occupied housing: an asset-market approach," *The Quarterly Journal of Economics*, 99(4), 729–752.

Roback, J. (1982): "Wages, rents, and the quality of life," *Journal of Political Economy*, 90(6), 1257–1278.

Rognlie, M. (2016): "Deciphering the fall and rise in the net capital share: accumulation or scarcity?," *Brookings Papers on Economic Activity*, 2015(1), 1–69.

Rosen, S. (1979): "Wage-based indexes of urban quality of life," *Current Issues in Urban Economics*, pp. 74–104.

Rozsypal, F., and K. Schlafmann (2017): "Overpersistence bias in individual income expectations and its aggregate implications," .

Saiz, A. (2010): "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 125(3), 1253–1296.

Van Nieuwerburgh, S., and P.-O. Weill (2010): "Why has house price dispersion gone up?," *The Review of Economic Studies*, 77(4), 1567–1606.

A Measuring Regional Divergence

In this appendix, we show the robustness of our measure of regional divergence, and explore the main causes. We begin by studying the timing of the relationship between house prices and regional divergence. Our findings suggest that the correlation between divergence and prices has remained strong throughout the sample period, but that at some times, one series has led or lagged the other. The visual evidence in Figure 1 suggests that during the 2000s, house prices responded more to regional divergence than during the 1940s and 1950s, but that they did so with a greater lag. Possible reasons for this may include changes in how households form expectations and changes in the persistence of income or housing market shocks.

Are our results robust to varying the time period? We investigate the timing and magnitude by measuring the correlation between divergence and house prices using a variety of lags and in several time periods, shown in Table 1. The correlations in Panel A show the relationship between contemporaneous house prices and 10-year divergence, and we separately estimate this correlation for the entire sample period as well as three sub-periods of equal length. The correlation increased from 0.38 during the 1939-1964 period to 0.81 from 1991-2018, and the correlations in all three sample periods statistically significant at conventional levels.[21]

Next, we ask what lag structure most closely matches the observed relationship between house prices and divergence during each period. To answer this, we estimate pairwise correlations between convergence and house prices at a variety of leads and lags and, in each time period, selected the convergence measure with the highest correlation. These results are shown in Panel B. As implied by the visual evidence in Figure 1, house price changes precede changes to regional divergence during the earliest years that data are available, but move nearly contemporaneously in more recent years.

Because the divergence measure is backwards-looking, the estimates in Table 1 do not

[21] We find a similar pattern in linear regression estimates.

imply that house price changes precede changes in regional divergence. Indeed, formal tests described in Section 5 imply that they do not. What the estimates do show is that the relative timing of the two series has changed even as their relationship has remained statistically and economically significant.

Table 1: Correlation Between National House Prices and Regional Divergence

	(1) 1939-2018	(2) 1939-1964	(3) 1965-1990	(4) 1991-2018
	A. Correlations With 10-Year Divergence			
Pairwise Correlation	0.76	0.38	0.51	0.81
Significance Level	0.00	0.06	0.01	0.00
Divergence Lag Years	10	10	10	10
	B. Correlation With Maximum-Correlation Divergence			
Pairwise Correlation	0.75	0.70	0.55	0.89
Significance Level	0.00	0.00	0.00	0.00
Lag on Prices	2	6	3	-1
Years	78	26	26	26

B Robustness of Shift-Share Growth

In this appendix, we show two other measures of shift-share growth similar to Figure 2, that address potential concerns about exogeneity. In the first, shown in Figure 8, we exclude real estate finance and construction sectors from the orange line, while the green line is exactly the same. Given these lines are hardly distinguishable from one another, we conclude that trends in the real estate sector are unlikely to be driving the co-movement between regional divergence and house prices.

In Figure 9, we would like to show the same shift-share growth, but using only the growth rates in other states, instead of the national rates. Unfortunately, we do not have a long panel of state-by-year-by-industry employment data we would need to do this, since we calculate

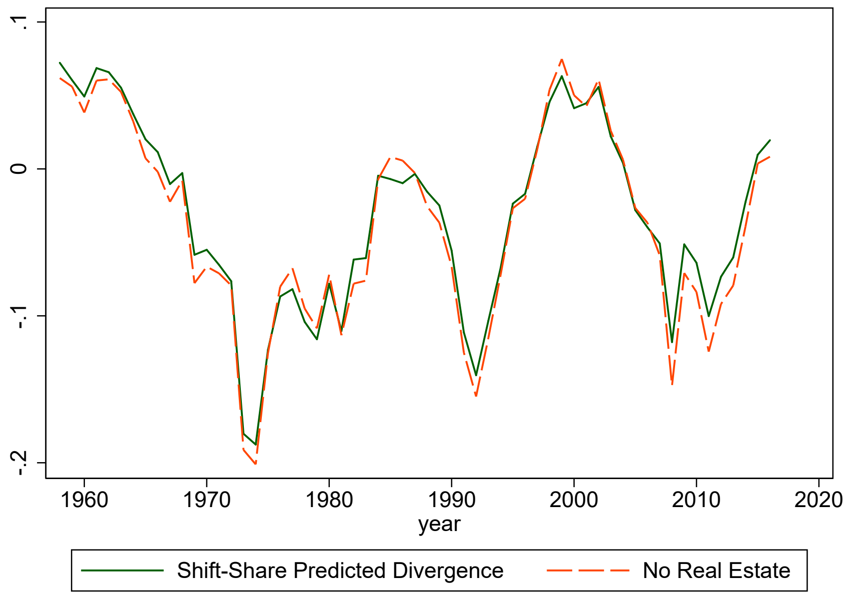

Figure 8: Shift Share Divergence Prediction, Primary Measure and Without Real Estate. The shift share measure without real estate is created in the same way as the primary measure, but excludes the real estate finance and construction sectors when calculating state industry GDP shares.

industry growth rates using GDP per employee. What we can do is construct this measure using growth rates of GDP. While it will not be the same as the predicted divergence we present in the main text, it will give us a sense if using a leave-one-out methodology is likely to change our results. The green line is the predicted divergence using GDP growth instead of GDP-per-employee growth. Many features are similar, although it does a less good job of tracking actual regional income divergence. In orange, we construct the measure using the leave-one-out methodology. These lines are quite close, so we conclude that using a leave-one-out methodology is unlikely to change our impression that much of regional divergence is because of industrial change.

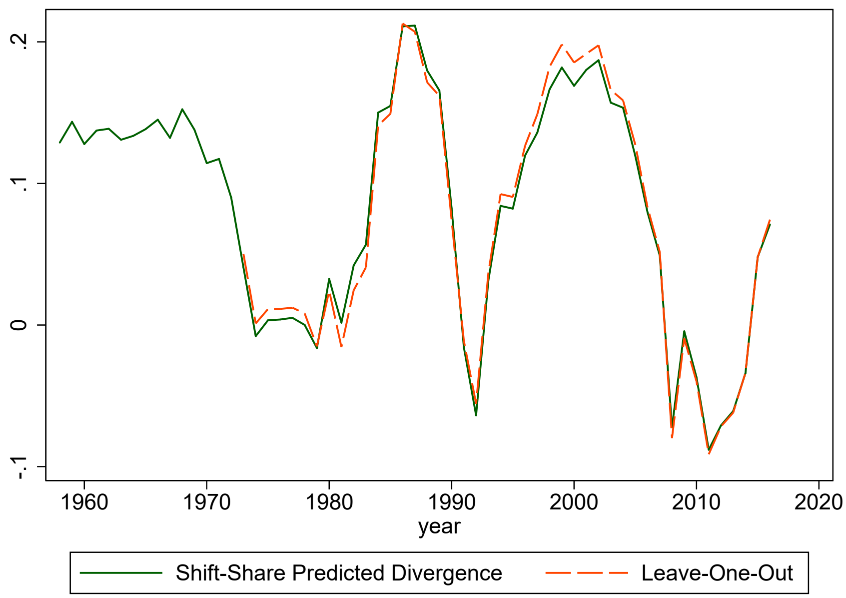

Figure 9: Shift Share Divergence Prediction and Leave-One-Out Measure. The leave-one-out shift share measure without real estate is created in the same way as the primary measure, but excludes each state's own industries when calculating national GDP growth by industry. The leave-one-out measure is only available from 1973 onwards, because this is when state-by-industry GDP data becomes available (the primary measure uses 1963 industry shares when predicting pre-1963 growth).