NBER WORKING PAPER SERIES

THE ANATOMY OF CONCENTRATION: NEW EVIDENCE FROM A UNIFIED FRAMEWORK

Kenneth R. Ahern
Lei Kong
Xinyan Yan

Working Paper 32057
http://www.nber.org/papers/w32057

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2024

For helpful comments, we thank Reena Aggarwal, David Hirshleifer, Mariassunta Giannetti, Jerry Hoberg, John Matsusaka and seminar participants at the University of Iowa, the University of Southern California, and the AFFECT Workshop at the AFA. The authors have no financial relationships to disclose. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Anatomy of Concentration: New Evidence From a Unified Framework
Kenneth R. Ahern, Lei Kong, and Xinyan Yan
NBER Working Paper No. 32057
January 2024
JEL No. C46,D40,L11

ABSTRACT

Concentration is a single summary statistic driven by two opposing forces: the number of firms in a market and the evenness of their market shares. This paper introduces a generalized measure of concentration that allows researchers to vary the relative importance of each force. Using the generalized measure, we show that the widely-cited evidence of increasing industrial employment concentration is driven by the Herfindahl Index's over-weighting of evenness and under-weighting of firm counts. We propose an alternative, equally-weighted measure that has an equivalent economic meaning as the Herfindahl Index, but possesses superior statistical attributes in typical firm size distributions. Using this balanced measure, we find that employment concentration decreased from 1990 to 2020. Finally, decomposing aggregate diversity into meaningful geographic and industry subdivisions reveals that concentration within regional markets has fallen, while concentration between markets has risen.

Kenneth R. Ahern
Marshall School of Business
University of Southern California
3670 Trousdale Parkway, HOH 718
Los Angeles, CA 90089
and NBER
kenneth.ahern@marshall.usc.edu

Lei Kong
The University of Alabama
Alston 233
Tuscaloosa, AL 35487
lkong2@ua.edu

Xinyan Yan
University of South Florida
xyan@usf.edu

A wealth of recent research shows a dramatic surge in industrial concentration since 1990.[1] In response, academics are raising concerns about increasing markups and declining competitiveness, while regulators are calling for stronger enforcement of antitrust regulations (Zingales, 2017; De Loecker, Eeckhout, and Unger, 2020; Federal Trade Commission, 2022). In light of this concerning trend in industrial concentration, it is imperative that we understand the theoretical and methodological assumptions that underlie the evidence.

In this paper, we re-evaluate the evidence on concentration and find that the prior results are sensitive to methodological assumptions. Using alternative, but equally valid assumptions, we find that from 1990 to 2020, employment concentration *decreased*. These findings cast doubt on the wide-spread belief that markets have become less competitive.

Why do our results differ from the prior evidence? The difference lies in the nature of concentration. Market concentration is defined by two distinct properties: the number of firms (richness) and the uniformity of their market shares (evenness). Each property is crucial for the definition of concentration. Without considering richness, a market with two equally sized firms would be as concentrated as a market with 100 equally-sized firms. Without considering evenness, a market with two equally sized-firms would be as concentrated as a market in which one firm had 99% of the market share and the other had 1%. Standard measures of concentration, such as the Hirschman-Herfindahl Index (HHI), make implicit assumptions to compress these two properties into a single summary statistic of concentration. While it is clear that a meaningful measure of concentration must account for both richness and evenness, the relative importance given to each property is a choice of the econometrician.

To make the methodological assumptions explicit, we introduce a generalized measure of concentration that parametrizes the weight placed on evenness and richness. Using this measure, we show that HHI is a special case of the generalized measure in which the importance

[1]See Gutiérrez and Philippon (2017), Grullon, Larkin, and Michaely (2019), Autor, Dorn, Katz, Patterson, and Van Reenen (2020), Autor, Patterson, and Reenen (2023), and Rossi-Hansberg, Sarte, and Trachter (2020).

of evenness is overweighted and the importance of richness is underweighted. These implicit weightings of HHI are not trivial. When we assign equal weights to evenness and richness in our generalized measure, we find that the stylized facts of concentration documented in recent research do not hold. Thus, the widely-reported increase in concentration is conditioned on implicit assumptions made by researchers.

Our generalized measure of concentration is based on a parsimonious set of axioms derived by both economists and ecologists (Hannah and Kay, 1977; Encaoua and Jacquemin, 1980; Daly, Baetens, and De Baets, 2018). Just as economic concentration reflects the lack of variety of economic entities within the economy, ecological diversity reflects the variety of species or traits within a given ecosystem. Following the mathematical ecology literature, our measure of 'true diversity' is indexed by order q and measured in units of 'effective firms.' Though q can take any value, we focus on three special cases of true diversity. When $q = 0$, true diversity puts all weight on the count of firms (*Count Diversity*). When $q = 1$, true diversity is equally sensitive to firm counts and market shares (*Balanced Diversity*). When $q = 2$, true diversity is the inverse of HHI and is more sensitive to the evenness of market shares than firm counts (*Dominance Diversity*). As q goes to infinity, all weight is placed on the firm with the largest market share.

We consider both economic and statistical criteria to determine the relative merits of balanced diversity versus dominance diversity (HHI). First, though HHI is the status quo among researchers and regulators, there is little economic theory that ties HHI to important economic outcomes (Demsetz, 1973, 1974; Syverson, 2019). In particular, though HHI is positively related to profitability under some theoretical assumptions (Stigler, 1964), it is also negatively related to profitability under other equally valid assumptions (Tirole, 1988, p. 223). Second, HHI is not unique. The theoretical results in Encaoua and Jacquemin (1980) show that for any weighting on richness and evenness across a variety of models of competition, the generalized measure of diversity can be related to the Lerner index. Similarly, in the canonical Cournot model with symmetric firms, evenness is canceled out, which means that any weighting on evenness and richness will generate the same relationship between concentration and profitability. Thus, though HHI is ubiquitous, it is not a superior measure based on economic theory.

Based on statistical criteria, HHI is also not superior to other weightings. In particular, following prior work on the firm size distribution (de Wit, 2005), we study the statistical traits of our generalized concentration measure under the assumption that market shares follow fat-tailed distributions, such as the Pareto or lognormal distribution. First, we show that HHI is not mathematically defined when the Pareto parameter value α is less than two. This is a serious limitation because nearly all empirical estimates of α are close to one. This means that the true value of HHI is not a valid mathematical construct for most distributions of market shares, especially when concentration is high. In contrast, balanced diversity is defined for nearly all estimates of α. Furthermore, we show that in lognormal distributions, balanced diversity has better statistical properties than HHI. Finally, we show that dominance diversity (HHI) exhibits much larger small-sample biases in both Pareto and lognormal distributions than does balanced diversity. Thus, we argue that our new results using balanced diversity provide a more accurate representation of concentration trends than existing research based on HHI.

Beyond the flexibility of weighting schemes, our framework has two additional advantages over traditional concentration measures. First, it measures diversity in intuitive units of 'effective firms,' defined as the number of firms with equal shares necessary to generate an equivalent measure of diversity (Adelman, 1969; Hill, 1973). For example, using weights based on HHI, a market with six firms with shares of (0.49, 0.28, 0.10, 0.054, 0.04, 0.016) has three effective firms because it has the same diversity as a market with three firms with equal market shares. This provides a more meaningful quantification of diversity than standard measures. For example, the economic meaning of an industry with an HHI of 250 is not readily apparent. In contrast, a market with 40 equally-sized effective firms is more easily understood.

The second advantage of our framework is that is allows diversity to be systematically decomposed from an aggregate into within-industry and between-industry components, with meaningful units. For example, the true diversity of national industries can be decomposed into true diversities of industry-counties. In this setting, within-diversity reflects the number of effective firms in the average industry-county and between diversity reflects the number of 'effective industry-counties.' If three markets have completely different firms from each other, then the between-industry diversity is three effective markets. In contrast, if the same set of firms have the same market shares in each of the three markets, then the between-industry diversity is one effective market. Between these two extremes, between-industry diversity reflects the degree of overlap between two markets, weighing firm counts and market shares based on parameter q. This decomposition allows us to quantify the difference between local and national diversity, which standard measures, such as HHI, cannot.

To apply our new approach to the data, we use establishment-level observations of employment from the National Establishment Time Series (NETS) database from 1990 to 2020, used by Rossi-Hansberg, Sarte, and Trachter (2020). To address concerns of over-sampling of very small firms in the NETS data (Barnatchez, Crane, and Decker, 2017), we restrict the data to include only firms with at least 20 employees. Using this data filter, the time-series of firm counts in NETS is highly correlated with the time-series of firm counts reported in Census data. To further address concerns about imputed data in NETS, we show that all of our results hold if we only use non-imputed data in NETS.

Within the average 4-digit national industry, dominance diversity (the inverse of HHI) of employment was 46.8 effective firms in 1990, relative to count diversity of 941.4 firms. From 1990 to 2013, dominance diversity fell by 40%, consistent with prior evidence of an increase in HHI at the national-industry level. Extending the sample period through 2020, we find that dominance diversity remained stable at around 30 effective firms from 2015 to 2020. In contrast, the balanced diversity of the average industry was 649.9 firms in 1990, increasing to 693.2 effective firms by 2020, for an increase of 6.7%. Thus, in contrast to conventional wisdom, when firm counts and evenness are equally weighted, concentration actually decreased slightly over the last 30 years.

We next decompose aggregate diversity into between-industry diversity. In 1990, there were 21.6 effective industries based on dominance diversity and 130.7 effective industries based on balanced diversity, relative to a count diversity of 875 industries. Using dominance diversity, the number of effective industries increased to 36.3 by 2020, while the number of effective industries decreased to 113.2 using balanced diversity. These results reflect that the declining prevalence of multi-segment conglomerates is concentrated among larger firms.

Next, decomposing industry diversity by geography, we find that the diversity within the average county and industry-county increased for both dominance and balanced diversity, while the diversity across counties and industry-counties decreased. These results indicate that counties and industry-counties have become significantly more homogenous over time, even as the diversity within the average industry-county has increased. These results provide quantifiable evidence that is consistent with the differing trends in HHI at the local and national levels presented in Rossi-Hansberg, Sarte, and Trachter (2020).

This paper has two main contributions. First, this paper advances a new, more flexible method to quantify economic concentration. Since the development of HHI by Hirschman (1945) and Herfindahl (1950), there have been few lasting advances in the measurement of concentration. Adelman (1969) first showed that HHI can be transformed into a numbers equivalent statistic, though few papers adopted his approach. Finkelstein and Friedberg (1967) advocated for an entropy-based measure of concentration, similar to the true diversity when $q = 1$. However, in a published comment on the paper, Stigler (1967) argued that that entropy does not have a theoretical connection to competition, whereas HHI does. However, subsequent research shows that the connection between HHI and competition is tenuous (Demsetz, 1968; 1973; 1974). Our framework extends the axiomatic approach of Hannah and Kay (1977) and Encaoua and Jacquemin (1980) by borrowing from the vast literature on diversity in ecology. For reviews of recent literature on methodologies in ecological diversity, see Daly, Baetens, and De Baets (2018) and Roswell, Dushoff, and Winfree (2021).

The second contribution of this paper is to use our new method to show that the widely-cited evidence of increases in market concentration relies on a set of ad hoc assumptions. In particular, we show that HHI is just one measure of concentration based on arbitrary assumptions, but other measures are equally valid and have superior statistical properties for typical firm size distributions. In particular, using balanced diversity, we show that market concentration of employment has declined, not increased, as widely reported. These results imply that concentration is more nuanced than suggested by prior evidence and that regulatory policies based on HHI could have unintended consequences.

I. Firm Counts, Market Shares, and Diversity Order q

In economics, the concept of market concentration is used to refer to the degree to which market share is concentrated among a few dominant firms or individuals. The inverse of concentration is diversity, or the variety and dispersal of market shares. While concentration is the standard framing used in economics research, we frame our mathematical approach in terms of diversity rather than concentration, which aligns our work more closely with recent advances in ecology. As we describe below, adopting this diversity perspective allows for a more intuitive unit of measurement that follows the mathematical formulation of diversity more closely.

I.A. Axiomatic Approach to Defining Diversity

There are four key axioms that define a useful measure of diversity independently proposed by economists and ecologists (Hannah and Kay, 1977; Encaoua and Jacquemin, 1980; Daly, Baetens, and De Baets, 2018).

1. *Evenness.* Diversity is maximized when market shares are equal across all firms in a market.
2. *Principle of transfers.* A transfer of market share from a firm with greater market share to one with lesser market share increases diversity.
3. *Entry.* The entrance of a new firm increases diversity.
4. *Replication principle.* If there are m sub-groups of firms such that each sub-group has equal diversity and no firms spans two subgroups, then the diversity of the super-group of the pooled m sub-groups is m times the diversity of a single sub-group.

Axioms 1, 2, and 3 are equivalent to the axioms specified by Hirschman (1945) and Herfindahl (1950) to define HHI. In particular, both Hirschman and Herfindahl argued that a measure of concentration should increase with the relative dispersion of market shares and it should decrease with the number of firms in the market. Hirschman noted that the axiom of entry was required to distinguish between measures of concentration and inequality. As noted above, without the axiom of entry, a market comprised of two equally sized firms would be as concentrated as a market with 100 equally sized firms. Thus, the number of unique firms is a critical component of diversity.

Axiom 4 has been proposed by ecologists, but it has not been included among the axioms of standard measures of economic concentration. However, it has important consequences for both the intuitive value of a diversity measure as well as its ability to be decomposed into sub-units. In particular, this axiom requires that a diversity measure scales linearly. For example, if the dollar values of sales for three different firms in Market A is $[10, 5, 5]$ and the dollar values for six different firms in Market B is $[10, 5, 5, 10, 5, 5]$, this axiom requires the intuitive result that the diversity of Market B should be double the diversity of Market A. We show below that this axiom means that diversity should be measured in intuitive units, not arbitrary indices, and also that a diversity measure should be able to be decomposed into smaller sub-groups while maintaining a consistent and intuitive unit of measurement. As an example, this axiom implies that the value of diversity of an industry that is measured at the national level should be intuitively related to the aggregation of the industry's diversity values that are measured at the local level.

I.B. Classical Diversity Measures

We next assess classical diversity indices and their alignment with the standard axioms. Nearly all classical indices of diversity are transformations of the following measure (Chao, Chiu, and Jost, 2014):

$$ {}^{q}\lambda = \sum_{i=1}^{N} p_i^q, \tag{1} $$

where parameter q determines the influence of common versus rare firms on the measure, such that the larger is q, the greater is ${}^{q}\lambda$'s sensitivity to the uniformity of market shares.

The first classical measure is richness, which is generated when $q = 0$. Richness is simply the number of unique firms in a market, without considering their market shares. The second classical measure is Shannon entropy derived as $\lim_{q\to 1} {}^{q}H = (1 -{}^{q} \lambda) / (q - 1)$, which equals $-\sum_{i=1}^{N} p_i \ln(p_i)$. Though this measure was originally developed in information theory (Shannon, 1948), it was first proposed in an economics setting by Finkelstein and Friedberg (1967) and it is widely used by ecologists as a measure of diversity. Finally, when $q = 2$, the third classical index of diversity, known as Simpson's index (Simpson, 1949) in ecology, is identical to HHI. Ecologists transform Simpson's index from a measure of concentration to a measure of diversity, as $1-{}^{2}\lambda$, which is known as the Gini-Simpson index. Though q is typically set to 0, 1, or 2, ${}^{q}\lambda$ is not limited to integer values for q.

This formulation reveals that HHI is a special case of a more general family of diversity measures. Moreover, the choice of $q = 2$ to define HHI is arbitrary. For example, defining HHI as ${}^{1.5}\lambda$ or ${}^{3}\lambda$ would also have satisfied HHI's axioms. Both Hirschman and Herfindahl recognize that HHI is not a unique solution to their axioms, but they chose ${}^{2}\lambda$ as their measure for ease of calculation and for its relationship to statistical measures of dispersion. As we show below, this seemingly innocuous assumption has important implications for the interpretation of HHI and its usefulness in empirical settings characterized by fat-tailed distributions of firm sizes.

Though the classical diversity measures are widely used, recent research argues that they have two significant drawbacks (Ellison, 2010; Roswell, Dushoff, and Winfree, 2021). First, they do not satisfy all of the axioms of diversity. Richness does not satisfy the principle of transfers because it ignores market shares. Shannon entropy and the Gini-Simpson index do not satisfy the replication principle because they do not scale linearly.

Second, the units and interpretation of classical measures of diversity vary across measures. Richness is measured in the number of firms, while the Gini-Simpson index is the probability that two random sales in the same market belong to different firms. Similarly, HHI is the probability that two random sales belong to the same firm. Further, Shannon entropy is measured in units of information and reflects the uncertainty of a firm of a randomly chosen sale in a market. When q takes values other than 0, 1, or 2, classical diversity measures are even harder to interpret. Thus, these measures are not easily compared across different orders of q and they do not have intuitive interpretations as 'diversity.'

I.C. True Diversity

To address the limitations of classical diversity measures, a seminal paper in mathematical ecology by Jost (2006) advocates for the adoption of diversity indices variously known as 'numbers equivalents' (Adelman, 1969), Hill numbers after Hill (1973), or what Jost calls 'true diversities.' Using true diversities is now the consensus among ecologists (Ellison, 2010; Chao and Ricotta, 2019) and in this paper, we advocate for their use in economics, just as

Adelman did 50 years ago. The following equations transform ${}^q\lambda$ into true diversities:

$$(2) \qquad {}^qD = ({}^q\lambda)^{\frac{1}{1-q}} = \left(\sum_{i=1}^{N} p_i^q\right)^{\frac{1}{1-q}} \quad \text{if } q \neq 1$$

and

$${}^qD = \exp\left(-\sum_{i=1}^{N} p_i \ln p_i\right) \text{ if } q = 1.$$

The transformations of the three classical diversity indices are easily calculated. If $q = 0$, then ${}^0D$ is the count of the firms in a market, or richness. If $q = 1$, then ${}^1D$ is the exponential of Shannon entropy. If $q = 2$, then ${}^2D$ is the inverse of HHI.

True diversity (${}^qD$) exhibits at least two advantages over classical diversity measures (Daly, Baetens, and De Baets, 2018). First, true diversity satisfies all four axioms, including the replication principle, for all $q > 0$. Second, true diversity possesses a consistent unit and interpretation across all q values. These units are 'effective firms' and represent the number of firms with equal market shares that would yield the same ${}^qD$ value as the observed market.

To illustrate the notion of effective firms, Figure I presents four different distributions of market shares. The first chart on the left represents a market with three firms with equal market shares. Moving to the right, the charts represent markets with increasing numbers of firms, though with more unequal market shares. However, the value of ${}^1D$ for each of these four markets equals three, which means that all four markets have a true diversity of three effective firms. In other words, when $q = 1$, the diversity in each market is equivalent to the diversity of a market with three firms with equal market shares of 1/3, as in the first chart.

This example highlights that true diversity measures have precise interpretations that can be easily compared across different distributions of market shares, in contrast to classical diversity indices, such as HHI. For instance, the concentration of a market with three firms with equal market shares is easy to comprehend. In contrast, the concentration of a market

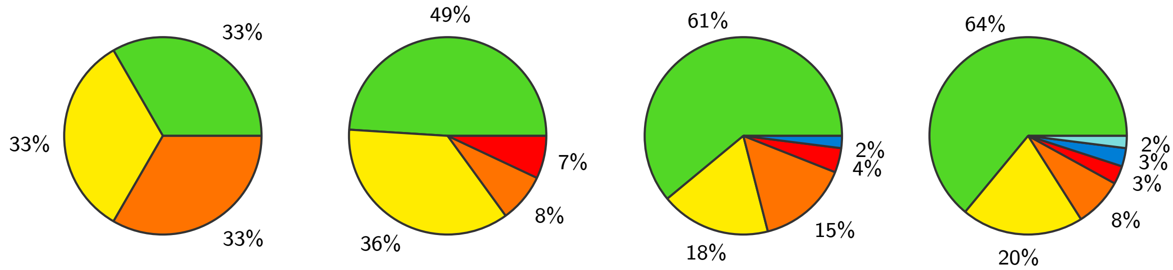


FIGURE I
EFFECTIVE NUMBER OF FIRMS
This figure illustrates the concept of effective firms. Each chart has the same true diversity of 3 at a diversity of order $q = 1$. The diversity of each market is equivalent to a market with three equally-sized firms.

with six firms with market shares of 64%, 20%, 8%, 3%, 3%, and 2% is not easily grasped. Using true diversities reveals that both markets are equally diverse when $q = 1$. Though we have shown equivalence with $q = 1$ in this case, the same general concept applies for any q.

I.D. Firm Counts versus Market Dominance and Diversity Order q

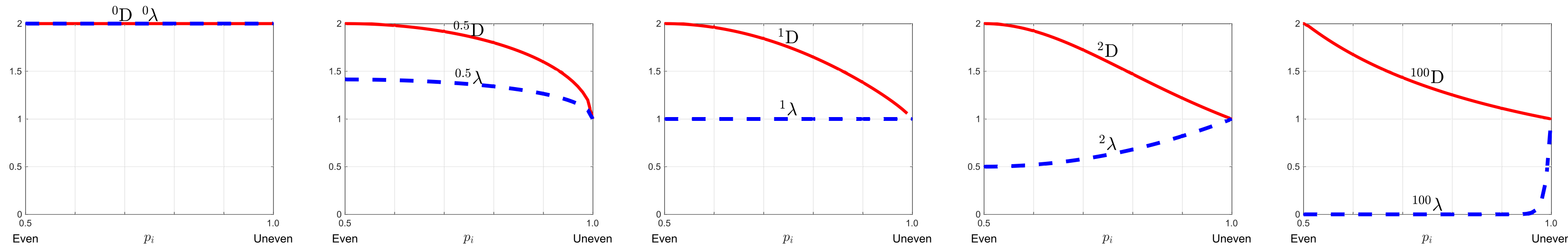
Figure I highlights the central trade-off of diversity: richness versus evenness. The entrance of a new firm increases diversity, but greater unevenness in firms' market shares decreases diversity. In our example, the increase in diversity caused by additional firms is exactly offset by the decrease in diversity caused by greater unevenness in market shares. The sensitivity of diversity to the count of firms and the evenness of market shares is determined by the order q. For values of q less than unity, true diversity is more influenced by the count of firms than the evenness of their market shares. For values of q greater than unity, true diversity is more influenced by the evenness of market shares than by the number of firms. At $q = 1$, counts and evenness have equal influence.

Figure II illustrates the role of q on diversity. The top five panels represent markets with two firms. In each panel, the horizontal axis represents variation in evenness from a perfectly even 50-50 share to a perfectly uneven share in which one firm has a market share approaching 100% and the other approaches 0%. The vertical axis represents the value of true diversity, ${}^qD$ and the value of ${}^q\lambda$ as the evenness of market shares changes. From Panel A to Panel E, the value of q increases from 0 to 100.

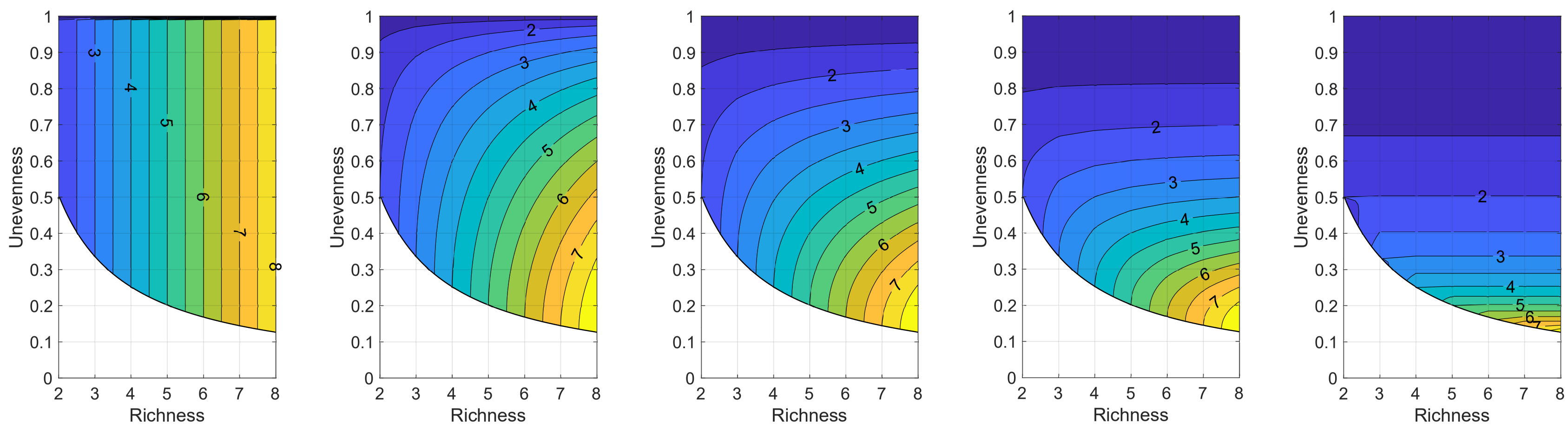
The top five panels show that for any value of q, when market shares are perfectly even, true diversity is two firms. Likewise, for any value of q, when the market shares are perfectly uneven, the true diversity is one firm. However, the transition from a diversity of two to one is controlled by the order q. When $q = 0$, evenness has no effect on diversity because market shares are ignored. As q increases, the effect of the unevenness of market shares on diversity intensifies. This is driven by the convexity of ${}^q\lambda$. When $q < 1$, ${}^q\lambda$ is concave, which gives more influence to small firms with low market shares. When $q > 1$, ${}^q\lambda$ is convex, which gives more influence to large firms with high market shares. The larger is q, the more convex is ${}^q\lambda$. The point at which evenness has a constant influence on diversity is when $q = 1$, as shown in Panel C. This is the unique value of q that weighs the frequency of each observation in proportion to its market share.

The bottom five panels of Figure II further illustrate the impact of q on the trade-off between the count of firms and the evenness of their market shares. Evenness is defined in these figures as the market share of the largest firm, p_1, assuming the remaining $n - 1$ firms have equal market shares of $\frac{1-p_1}{n-1}$ (Jost, 2010). The minimum level of unevenness is represented by the parabola $1/n$ and the maximum is one. The true diversity of each market is indicated by color-coded iso-bars representing equal diversity across variation in counts and evenness.

In Panel F, $q = 0$ and diversity is richness (${}^0D$). At this extreme, diversity is completely independent of the evenness of market shares. At the other end of the spectrum, in Panel J, when $q = 100$, ${}^qD$ is highly sensitive to evenness and insensitive to the number of firms. Panels G, H, and I, represent intermediate steps in which $q = 0.5$, 1.0, and 2.0. As q increases, the sensitivity of diversity to evenness increases as q increases. At $q = 1$, consistent with Panel C, the substitution is balanced equally between counts and unevenness. Given this



(A) $q = 0.0$ (B) $q = 0.5$ (C) $q = 1.0$ (D) $q = 2.0$ (E) $q = 100$



(F) $q = 0.0$ (G) $q = 0.5$ (H) $q = 1.0$ (I) $q = 2.0$ (J) $q = 100$

FIGURE II
RICHNESS AND EVENNESS

Panels A through E present diversity component ${}^q\lambda = p_1^q + (1 - p_1)^q$ and true diversity ${}^qD = ({}^q\lambda)^{\frac{1}{1-q}}$ of a market with two firms for q from zero (Panel A) to 100 (Panel E), where p_1 is firm 1's market share. The horizontal axis represents the evenness of the market shares. It is perfectly even when both firms have market shares of 50% and perfectly uneven when one firms has a market share that is 100% and the other has a market share of zero. Panels F through J present markets with n firms (richness), in which the market share of one firm, p_{max}, moves from a minimum of $\frac{1}{n}$ to the maximum of 1, while all other firms have shares $\frac{1-p_i}{n-1}$ (unevenness). Iso-lines representing constant true diversity ${}^qD$ appear as a contour plot for $q = 0$ in Panel F to $q = 100$ in Panel J.

discussion, for ease of exposition in the rest of the paper, we refer to ${}^0D$ as *Count Diversity*, ${}^1D$ as *Balanced Diversity*, and ${}^2D$ as *Dominance Diversity*.[2]

II. Diversity of Employment: Preliminary Evidence

To illustrate the importance of the underlying assumptions of concentration measures, we estimate true diversity of the employment shares of employers. Our first set of tests show strikingly different results for balanced and dominance diversity at the national industry level. Therefore, after these tests, we discuss the relative merits of using balanced diversity versus dominance diversity. Then, we consider decompositions of national-level industry diversity into within- and between-industry diversity based on local geographic markets.

II.A. Data Sources

To estimate diversity within an industry requires firm-level data on market shares of both public and private firms. In addition, as we show later, to estimate diversity between industries and counties requires detailed geographic and segment-level data to account for the overlap between industries created by multi-segment and multi-regional firms. To address these requirements, we use data from the National Establishment Time Series (NETS), as used to study concentration in Rossi-Hansberg, Sarte, and Trachter (2020).[3] The NETS data cover over 74 million establishments from 1990 to 2020, both public and private, where an establishment is a business or plant at a single physical location. For each establishment, NETS provides the location, industry code, ultimate owner, employment level, and sales. To

[2]In Section A2 of the Internet Appendix, we provide a mathematical proof that $q = 1$ is the unique case that perfectly balances evenness and richness. In particular, when market shares are uneven, transferring a small amount of market share from one firm to another is of the same order as a change in the number of firms only when $q = 1$.

[3]Other papers that use NETS include Bernstein, McQuade, and Townsend (2021), Faccio and Hsu (2017), Farre-Mensa, Hegde, and Ljungqvist (2020), Crouzet and Mehrotra (2020), and Borisov, Ellul, and Sevilir (2021). For a detailed description of the NETS database, see Kolko, Neumark, and Lefebvre-Hoang (2007) and Barnatchez, Crane, and Decker (2017).

our knowledge, NETS is the most comprehensive establishment-level dataset available other than confidential Census micro-data.

To make NETS data more closely match official data sources, we apply filters following suggestions from prior research. First, we exclude SIC codes related to government entities and federal reserve banks. Second, we filter by firm size. In particular, Barnatchez, Crane, and Decker (2017) find that though NETS is consistent with official data sources in terms of distributions across industries and geographic regions, NETS tends to mismeasure the incidence of very small firms. To address this issue, Barnatchez et al. show that excluding establishments with less than ten employees in the NETS data from 1990 to 2014 substantially improves the correlation between the number of establishments in NETS and official sources from the Census Bureau. We confirm the findings of Barnatchez et al. for 1990 to 2014, but we observe a large disparity in firm counts between official sources and NETS for the period from 2015 to 2020. Instead, we find that the correlation between NETS and Census records across the entire sample period is highest when we exclude firms with less than twenty employees, rather than ten.

Panel A of Figure III compares the time-series of firm counts from 1990 to 2020 from the Census Bureau's Statistics of U.S. Business (SUSB) to firms counts from NETS in which both series are restricted to firms with at least 20 employees. The figure shows a similar time-series pattern for both series with similar magnitudes, though the SUSB data displays more extreme volatility in the second half of the sample period. The results in Panel A suggest that NETS lags the SUSB time series. Therefore, Panel B presents the three-year leading NETS data which appears to match the official SUSB data more closely.

To quantify the correlation between NETS and SUSB, we estimate the following equations using yearly observations from 1990 to 2020:

$$\ln(SUSB_t) = \underset{(1.841)}{2.916} + \underset{(0.135)}{0.760} \ln(NETS_t) + \varepsilon_t \tag{3}$$

$$\ln(SUSB_t) = \underset{(1.959)}{-0.215} + \underset{(0.143)}{0.988} \ln\left(NETS_{t+3}\right) + \varepsilon_t \tag{4}$$

where $SUSB_t$ and $NETS_t$ are the number of firms with at least twenty employees in the SUSB and NETS databases in year t. Standard errors of the coefficient estimates are reported in parentheses. These results show that the time-series of firm counts in NETS is statistically and economically significantly correlated with the time-series of firm counts in SUSB. The results also suggest that the official data lead the NETS data by three years. Based on these results, in all of our empirical tests we use NETS data restricted to firms with at least twenty employees and note that the patterns we observe may better represent the period three years prior to the date in NETS.

A related criticism of NETS data is that many of its observations are imputed, especially for sales data.[4] To address this concern, we re-run all of our analyses using only observations

[4] It is important to note that all micro-data, including administrative data collected by government agencies, include imputed data, especially for small firms. For example, Chow, Fort, Goetz, Goldschlag, Lawrence,

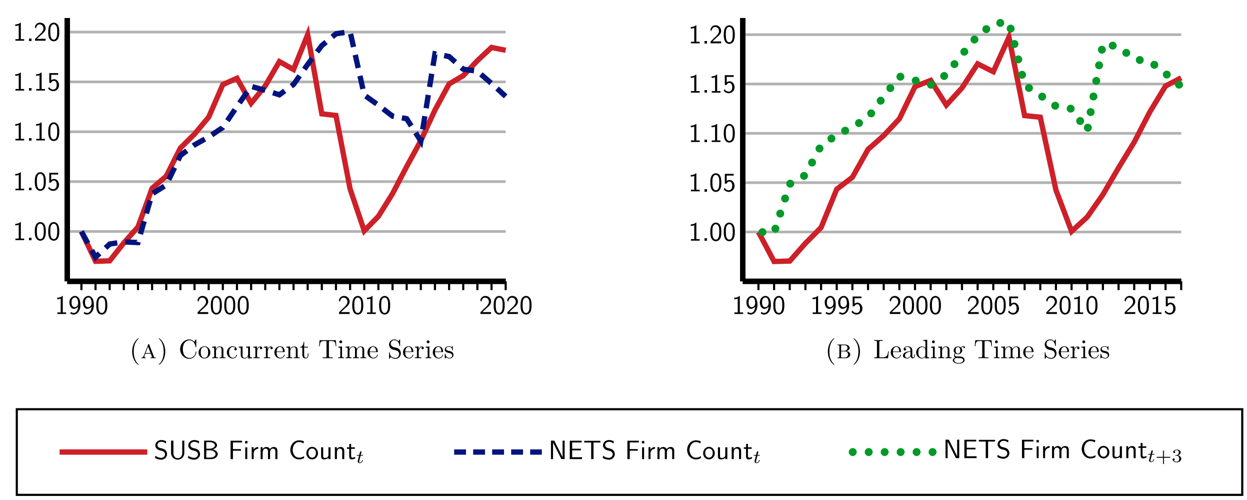


Figure III

Census Bureau Data vs. Filtered NETS Data

This figure presents the time-series of firm counts normalized by 1990 values in the Statistics of U.S. Business (SUSB) compared to the National Establishment Time Series (NETS) from 1990 to 2020. Panel A presents concurrent time-series. Panel B presents three-year leading NETS data. In both series, firm counts include only firms with at least 20 employees.

in NETS that are flagged as not imputed. We find that the results are nearly identical for employment shares, but the results for market shares of sales deviate in the non-imputed data in the later part of our sample. Therefore, throughout the paper we use only employment data from NETS and when we refer to firms, we mean firms with at least twenty employees. All of the results using non-imputed employment data are reported in Section A6 of the Internet Appendix.

By limiting attention to employee data of firms with at least 20 employees in NETS, we help mitigate concerns about data limitations. However, the contribution of this paper is methodological, not empirical. Limitations in the data do not reflect limitations in our methodological approach. The empirical tests we offer are meant to be illustrative, not definitive, and we encourage researchers to apply our methodological approach to alternative datasets to compare results.

II.B. Count, Balanced, and Dominance Industry Diversity

Table I provides industry-weighted averages of count, balanced, and dominance diversity of employment, for years 1990, 2000, 2010, and 2020.[5] To provide a more tangible picture of the data, Table II presents the industries at key percentiles of diversity. For example, the industry with the largest number of firms is eating places with over 83,000 firms, while the least diverse industry at $q = 1$ and $q = 2$ is household laundry equipment.

Perlman, Stinson, and White (2021) explain that the Census's confidential Longitudinal Business Database (LBD) relies on imputed data during intercensal years to 'smooth out' the bunching of establishment births and deaths that are observed in years in which the more complete Economic Census is conducted. In particular, total employment at the firm level is allocated across establishments either reported or imputed in intercensal years. Imputation is not uncommon because most small and medium size single-establishment firms with fewer than 500 employees are not included in the annual Report of Organization. According to CBP data, in 2019, only 0.3% of firms in the U.S. have more than 500 employees.

[5]Diversity measures, including HHI, rely on sample-based population estimates prone to measurement error, including the omission of rare species. The smaller is q, the bigger is the effect of undercounting rare species. Chao and Jost (2015) provide a bootstrap procedure to calculate standard error bounds and address undercounting by estimating the distribution of uncounted species using the number of species with one or two observations in the sample. In our setting, where sampling units are number of employees in national samples, the bootstrapped confidence intervals are extremely small. Moreover, the sampling procedure in ecology is different than in economics, so the correction for undercounting does not apply.

Table I
Employment Diversity

	Year			
	1990	2000	2010	2020
Count Diversity ($q = 0$)	941.4	1,111.2	1,086.1	1,106.2
Balanced Diversity ($q = 1$)	649.9	646.3	708.7	693.2
Dominance Diversity ($q = 2$)	46.8	42.1	34.4	28.0

Notes: This table presents snapshots of weighted average industry diversity of employment at the SIC 4-digit nationwide industry level for diversity orders $q = 0$, 1, and 2. Diversity is in units of effective firms. Data are from NETS and include firms with at least 20 employees.

Table I shows that in 1990, the average industry had 941.4 firms, increasing over the sample period, to 1,106.2 by 2020. This represents an average annual growth rate of firms of 0.539%. This growth rate is very close to the growth rate in SUSB data of 0.558%, which validates that our sample closely approximates official Census data.

Consistent with prior evidence of increasing HHI, dominance diversity ($q = 2$) fell substantially from 1990 to 2020. The number of effective firms in the average industry when $q = 2$ was 46.8 in 1990. By 2020, dominance diversity had fallen by 40% to 28.0 effective firms. However, when we consider balanced diversity, the results contradict the prior evidence of increasing concentration. Using balanced diversity, the average industry had 649.9 effective firms in 1990, rising slightly to 693.2 effective firms by 2020. Figure IV presents the time-series changes in each diversity measure from 1990 to 2020. The figure shows that in contrast to the dramatic decline in dominance diversity, balanced diversity increased by nearly 20% from 1990 to 2012, followed by a dip, then a recovery to end around 6.7% higher in 2020 than in 1990.

Table II
Diversity Percentiles by Industries

Rank/Percentile	Industry	Effective Firms
Panel A: Count Diversity ($q = 0$)		
Minimum	House Slippers	6.0
5th	Creamery Butter	28.0
25th	Drapery, Curtain, and Upholstery Stores	110.0
50th	Sewing, Needlework, and Piece Goods	284.0
75th	Florists	803.0
95th	Repair Shops, Not Elsewhere Classified	3951.0
Maximum	Eating Places	83002.0
Panel B: Balanced Diversity ($q = 1$)		
Minimum	Household Laundry Equipment	2.1
5th	Malt Manufacturing	9.2
25th	Aeronautical, and Nautical Systems and Instruments	31.4
50th	Mobile Home Dealers	79.5
75th	Newspapers - Publishing, or Publishing and Printing	243.7
95th	Engineering Services	1487.8
Maximum	Eating Places	13454.4
Panel C: Dominance Diversity ($q = 2$)		
Minimum	Household Laundry Equipment	1.5
5th	Household Vacuum Cleaners	5.3
25th	Softwood Veneer and Plywood	13.4
50th	Men's and Boys' Suits, Coats, and Overcoats	33.6
75th	Airports, Flying Fields, and Airport Terminal Services	77.5
95th	Local Trucking With Storage	415.7
Maximum	Elementary and Secondary Schools	1629.6

Notes: This table presents the nationwide SIC 4-digit industries for a range of diversity percentiles for diversity orders $q = 0$ (Count Diversity), 1 (Balanced Diversity), and 2 (Dominance Diversity). Data are from NETS and include firms with at least 20 employees.

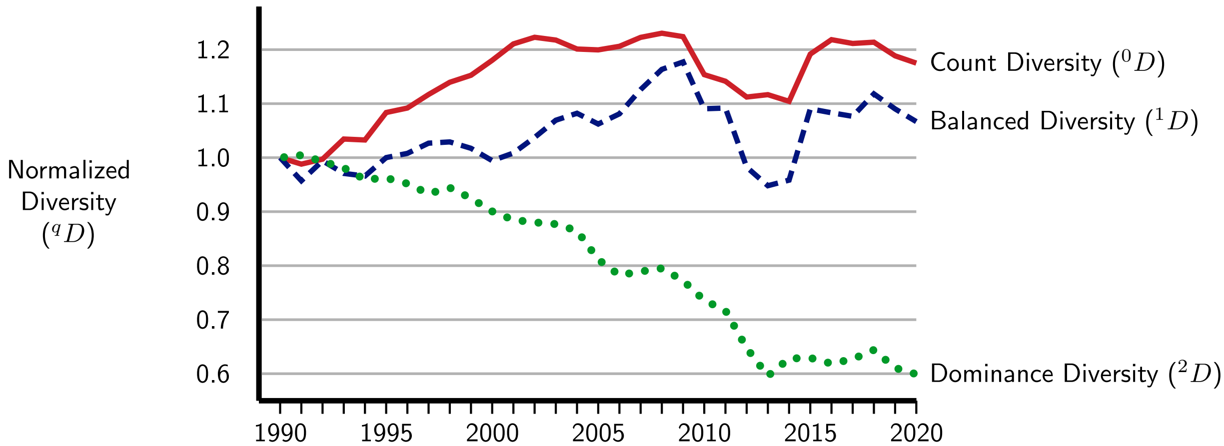


FIGURE IV
WEIGHTED AVERAGE INDUSTRY DIVERSITY

This figure presents the time-series of employment diversity at the SIC 4-digit nationwide industry level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Data are from NETS and include firms with at least 20 employees.

The striking difference between balanced and dominance diversity is driven by the relative weighting of firm counts and unevenness of market shares. Because dominance diversity underweights firm counts, it is less affected by the increasing number of firms in the average industry, as observed in both NETS and the Census data. These results illustrate that the choice of q, and hence, the weight given to richness versus unevenness, can have large implications for our understanding of diversity.

For a more disaggregated view, Figure V presents the time series of industry-weighted diversity at the sector level. Count diversity rose in nearly all sectors and the net gain of balanced diversity is larger than dominance diversity in nearly all sectors. Some sectors realized small changes in diversity, including mining, manufacturing, and wholesale, while other sectors realized large increases in balanced diversity, including agriculture and construction, and others realized large decreases, including retail and finance.

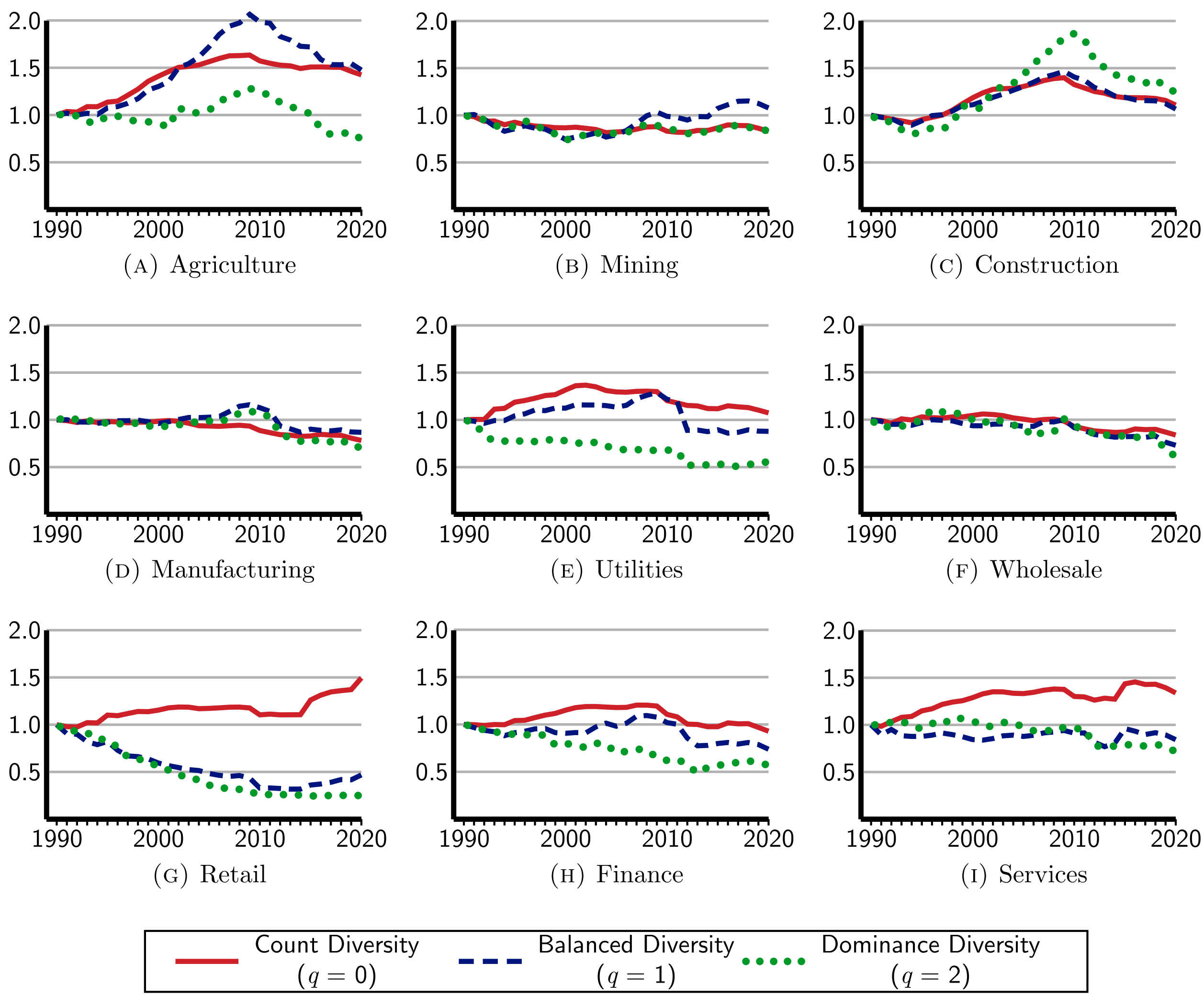


Figure V
Weighted Average Industry Diversity by Sector

This figure presents the time-series of employment diversity at the sector level from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Data are from NETS and include firms with at least 20 employees.

III. Which Diversity Measure Is Best?

The foregoing empirical evidence reveals that the choice of q has important consequences for the meaning of diversity. In particular, HHI's assumption that $q = 2$ is non-trivial. Using balanced weights with $q = 1$ provides a starkly different picture of the diversity of employment in the average industry. Thus, it is important to ask which q is most appropriate. To answer this question, we consider both economic and statistical criteria for the choice of q.

III.A. Economic Rationale for q

Even though the vast majority of economics research relies on HHI to measure concentration, industrial organization economists have long recognized that HHI is not well grounded in economic theory (Bresnahan, 1989). In particular, though a line of research starting with Stigler (1964) connects HHI with profitability in Cournot oligopoly models,[6] subsequent research showed that these models rely on unrealistic assumptions about firm entry and their results do not hold empirically (Demsetz, 1968, 1973, 1974).

Second, just as HHI is not a unique solution to the axioms of Herfindahl and Hirschman, its theoretical properties are also not unique. In particular, Encaoua and Jacquemin (1980) show that any measure of concentration that satisfies the Herfindahl and Hirschman axioms, including both HHI and balanced diversity, can be theoretically related to the Lerner index in a variety of models of competition, including non-cooperative oligopoly with differentiated and homogenous products and in both static and dynamic price leadership models. This means that for any value of $q > 0$, the properties of ${}^{q}D$ are equivalent to the properties of HHI in a wide range of theoretical settings.

To illustrate why HHI is not unique, consider the canonical symmetric Cournot model, where each firm is assumed to have identical costs and market shares. In this model, it is

[6]See Dansby and Willig (1979) and Donsimoni, Geroski, and Jacquemin (1984) for subsequent papers.

well known that total industry profit is increasing with HHI. However, as our discussion of concentration has shown, when market shares are equal, true diversity equals the count of firms in the market for all q. Intuitively, the symmetric Cournot model eliminates variation in market shares and focuses solely on firm counts. Thus, any true diversity measure would generate the same relationship between concentration and profitability as does HHI in a symmetric Cournot model because all true diversities are equivalent to the count of firms in a perfectly even market. Thus, HHI does not represent a unique measure in symmetric Cournot models.

Alternatively, in asymmetric Cournot models, cost functions are not identical across firms, and market shares are not perfectly even. Thus, the choice of q will influence true diversity. However, as we have shown above, HHI can be held constant, while changing the number of firms and their market shares. For HHI to be a reliable predictor of profitability, profitability should be constant if HHI is constant, even if the number of firms vary. However, HHI fails this criterion. In Internet Appendix Section A1, we show that total profit varies as firm counts change, even holding HHI constant. This means that HHI does not uniquely identify profit in asymmetric Cournot models.

In summary, while HHI is ubiquitous, it is not an optimal measure of concentration based on existing economic theory. Our analysis based on Cournot models and the extensive analysis in Encaoua and Jacquemin (1980) show that HHI does not have unique properties for measuring economic outcomes. More generally, the economic relationship between any measure of concentration and profitability is tenuous. Tirole (p. 231, 1988) sums up the vagueness of concentration in existing economic theory, writing "... [concentration indices] have no systematic relationship with economic variables of interest for assessing changes in cost, demand, or policy."

III.B. Statistical Rationale for q

Given that the link between concentration and economic theory is tenuous, we next evaluate the statistical properties of diversity measures. Though concentration may not directly relate to economic fundamentals such as competitive intensity or profit, it still represents a key descriptive statistic of market conditions. Without any economic interpretation, concentration measures serve as summary statistics for statistical distributions.

To assess the relative statistical merits of different diversity measures, we focus on two statistical distributions commonly predicted in stochastic growth models: lognormal and Pareto. In a seminal work, Gibrat (1931) shows that with a fixed number of firms that receive random growth shocks drawn from a common distribution, the firm size distribution approaches a lognormal distribution. Subsequent work allowed for the entry of small firms which led to a Pareto, or power law, distribution, or its discrete counterpart, a Yule distribution (Simon, 1955, 1960). Further refinements, such as restricting firm sizes to be larger than a minimum threshold, allowing for a selection mechanism for survival, or assuming growth rates that are decreasing in firm size also generate Pareto-like distributions (Levy and Solomon, 1996; Luttmer, 2007; de Wit, 2005). Thus, the majority of economic theory predicts that market shares will follow a fat-tailed distribution, whether Pareto or lognormal.[7]

Consistent with theory, the vast majority of empirical evidence is consistent with Pareto and lognormally-distributed firm sizes. Using U.S. Census data, Axtell (2001) finds that firm sizes follow a Pareto distribution with a shape parameter close to one, consistent with Zipf's law. More recently, Kondo, Lewis, and Stella (2022) find evidence in support of lognormally-distributed firm sizes using U.S. Census microdata. Similarly, in French and U.K. data, evidence supports both Pareto (di Giovanni, Levchenko, and Rancière, 2011; Garicano, Lelarge, and Reenen, 2016; Montebruno, Bennett, van Lieshout, and Smith, 2019)

[7]See de Wit (2005) for an extensive summary of the stochastic growth literature.

and lognormally-distributed firm sizes (Clarke, 1979; Bee, Riccaboni, and Schiavo, 2017). Similar findings are reported for firms in Portugal (Cabral and Mata, 2003), Belgium (Artige and Bignandi, 2023), India (Amirapu and Gechter, 2020), and across a range of other countries (Giovanni and Levchenko, 2013; Gaffeo, Gallegati, and Palestrini, 2003).

Given the strength of the evidence for lognormal and Pareto distributions, we provide analytic expressions of ${}^qD$ assuming firm sizes follow these two distributions.[8] First, we assume that firm sizes follow the Pareto distribution $F(x) = 1 - (k/x)^\alpha$ for $x > k$, where k is the location parameter and α is the shape parameter. Using this distribution, we can rewrite Equation 2 as follows,

$$\text{(5)} \qquad {}^qD_{\text{Pareto}} = N \cdot \frac{\alpha(\alpha-q)^{\frac{1}{q-1}}}{(\alpha-1)^{\frac{q}{q-1}}} \quad \text{if } q \neq 1 \text{ and } \alpha > q,$$

$$ {}^1D_{\text{Pareto}} = N \cdot \left(\frac{\alpha}{\alpha-1}\right) e^{-\frac{1}{\alpha-1}} \quad \text{if } \alpha > q = 1.$$

This formulation shows that in Pareto-distributed data, ${}^qD$ reflects the product of the number of firms in the market and a shrinkage factor that ranges from zero to one based on the shape parameter and the diversity order. As α approaches infinity, the distribution approaches perfect evenness such that ${}^qD = N$ and the shrinkage factor approaches one for all q. As α approaches q, the distribution becomes less even, the shrinkage factor approaches zero, and ${}^qD$ decreases. Likewise, as q goes to zero, the shrinkage factor goes to one.

Equation 5 reveals a critical boundary condition of ${}^qD$ in Pareto-distributed data: ${}^qD_{Pareto}$ is only defined when $\alpha > q$. This is because higher values of q rely on higher moments and higher moments of the Pareto distribution only exist for higher values of α. For example, HHI is based on squared values of market shares, similar to the variance. However, the variance of the Pareto distribution is infinite for $\alpha \in (1, 2]$. Thus, if firm sizes are distributed following a Pareto distribution with $\alpha < 2$, the variance of firm sizes is infinite and diversity measures based on the second moment, including HHI and ${}^2D$, are not mathematically valid

[8]This analysis is an extension of Hart (1975). See the Internet Appendix for details.

measures. This is a serious limitation for HHI and ${}^2D$ because nearly all empirical estimates of α in firm size distributions are below two.

In contrast, when $q = 1$, the expression for ${}^1D_{Pareto}$ only requires that $\alpha > 1$, which is consistent with the large majority of estimates of α in the empirical literature. While any true diversity with $q < 1$ will provide reliable statistics for Pareto distributions, because most firm size distributions have $\alpha > 1$, there is no advantage from using lower values of q. Thus, statistically, ${}^1D$ is an ideal measure of diversity in the most commonly observed Pareto distributions. In contrast, HHI is mathematically undefined.

Second, if firm sizes X follow a lognormal distribution where σ is the standard deviation of $\log(X)$, the true diversity of order q is

$$(6) \qquad {}^qD_{\text{Lognormal}} = Ne^{-\frac{\sigma^2}{2}q}.$$

As in Pareto-distributed data, this formulation shows that in lognormally-distributed data, ${}^qD_{\text{Lognormal}}$ reflects a shrinkage of the number of firms in the market. If there is no variance in the firm size distribution ($\sigma = 0$), then the shrinkage factor $e^{-\frac{\sigma^2}{2}q}$ equals one, market shares are perfectly even, and ${}^qD_{\text{Lognormal}} = N$ for all q. As σ increases, market shares become less even and true diversity shrinks. Likewise, as q goes to zero, the shrinkage factor goes to one, and ${}^qD_{\text{Lognormal}}$ is equivalent to richness.

Diversity order $q = 1$ represents a unique value in lognormally-distributed data. In Section A2 of the Internet Appendix, we show that the elasticity of richness (N) with respect to σ is $q\sigma^2$. In other words, holding diversity fixed, a one percent increase in the standard deviation of $\log(X)$ is exactly offset by a $q\sigma^2$ percent increase in firm counts. Thus, elasticity is scaled by q. When $q < 1$, the elasticity of richness and evenness is smaller; when $q > 1$, the elasticity is larger; and when $q = 1$, diversity is equally balanced between richness and evenness.

Second, ${}^1D$ is a natural statistic for lognormal distributions because it represents a weighted geometric mean. In particular, Equation 2 can be rewritten as the reciprocal of a weighted generalized mean of market shares:

$$(7) \qquad {}^qD = \frac{1}{\left(\sum_{i=1}^{N} p_i \cdot p_i^{q-1}\right)^{\frac{1}{q-1}}}.$$

Based on the definition of the generalized mean, when $q = 0$, the generalized mean is the arithmetic mean; when $q = 1$, it is the geometric mean; and when $q = 2$, it is the harmonic mean.[9] It is well-known that the most natural measure of central tendency for lognormally distributed data is the geometric mean (Aitchison and Brown, 1958). In particular, the geometric mean of lognormally distributed data is e^{μ}, compared to the arithmetic mean of $e^{\mu+\sigma^2/2}$ and the harmonic mean of $e^{\mu-\sigma^2/2}$. Thus, $q = 1$ represents a stable measure of location that is independent of variance, in contrast to $q < 1$ and $q > 1$.

III.C. Small-Sample Bias of ${}^qD$

The analytic formulations of ${}^qD$ for lognormal and Pareto distributions allow us to evaluate the small-sample properties of ${}^qD$ relative to their population values. For example, though ${}^2D$ is not analytically defined for $\alpha < 2$, it can always be calculated in sample data. Thus, the sample-based estimate of HHI is biased downward when $\alpha < 2$ because the population value of HHI is infinite. The severity of the bias depends on the sample size and the parameters of the assumed distribution, and importantly, on the value of q used to estimate diversity.

In Section A4 of the Internet Appendix, we show that ${}^1D$ has substantially less small-sample bias than ${}^2D$ for both lognormal and Pareto distributions. First, when $q = 1$, we show that the estimates of σ and α are nearly identical when we assume that our sample data represents the entire population as when we assume that our sample data does not represent

[9] The generalized mean inequality states that the generalized mean of $(p_1, p_2, \ldots, p_n)$ when $q = a$ is less than or equal to the mean when $q = b$ if $a < b$, with equality if $p_1 = p_2 = \cdots = p_n$. This explains why count diversity is larger than balanced diversity which is larger than dominance diversity. It also explains why true diversity is the same for any q when market shares are equal.

the entire population. In contrast, when $q = 2$, the bias is considerably larger. In particular, when $q = 1$, the deviation between data-derived estimates and analytically-derived estimates of σ and α are less than 1%, on average, compared to more than 50% when $q = 2$. This implies that balanced diversity calculated using sample data better reflects population-level diversity than does dominance diversity.

Second, we calculate the mean value of ${}^1D$ and ${}^2D$ across 10,000 random samples of a given sample size N and shape parameter α or σ. We then calculate the ratio of the empirical estimate of ${}^1D$ and ${}^2D$ to their analytical values. The upward small-sample bias of ${}^2D$ is substantially larger than the bias for ${}^1D$ in both Pareto and lognormal distributions for all sample sizes and parameter values. For example, in a sample of 100 firm shares drawn from a lognormal distribution with $\sigma = 1$, the empirical estimate of ${}^1D$ is 1.8% larger than the analytical value compared to 13.3% larger for ${}^2D$. For smaller samples drawn from a Pareto distribution the bias is magnified. The estimate of ${}^1D$ is 3.7% larger than its analytical value in a sample of 25 firms, while the empirical estimate of ${}^2D$ is 39.6% larger.

III.D. Summary: Balanced Diversity is Superior to Dominance Diversity

In sum, though HHI is widely used in practice, we have shown that it does not have superior properties from either an economic or statistical view. In terms of economic rationales, in the relatively rare settings in which HHI can be linked to economic outcomes, such as static Cournot models, Encaoua and Jacquemin (1980) show that the same relation holds for any q. From a statistical view, the properties of HHI limit its usefulness when firm sizes follow Pareto and lognormal distributions, as commonly observed. In contrast, though ${}^1D$ does not have special economic properties, we have shown that it is has superior statistical properties in markets characterized by either Pareto or lognormal distributions. In addition, we show that HHI exhibits substantial small-sample bias in lognormal and Pareto distributions. In contrast, balanced diversity has little small-sample bias. Thus, for an average market, ${}^1D$ is likely to be a better measure of diversity than ${}^2D$ or HHI.

III.E. Diversity Profiles

While the prior discussion indicates that ${}^1D$ is a better measure of diversity for most settings, there is no reason to restrict attention to one level of q. A central theme of this paper is that diversity indices combine a multidimensional concept into a single number. This compression necessarily hides important information about the underlying structure of the market. Rather than choosing one q to measure diversity, it is more informative to analyze diversity for multiple values of q (Patil and Tallie, 1979; Chiu, Jost, and Chao, 2014).

To analyze diversity for multiple values of q, we use a diversity profile, which plots ${}^qD$ as a function of q. In particular, a diversity profile helps to visualize a market's degree of evenness and firm counts. The slope of the profile indicates the evenness of the market. A perfectly even market has constant diversity for any level of q, and a profile slope of zero. The more uneven is a market, the steeper is its diversity profile, as diversity falls faster as more weight is placed on the unevenness of market shares. If the diversity profile for one market is higher than another market for all values of q, we can conclude that it is more diverse unconditionally. Instead, if the diversity profile of one market intersects the diversity profile of another market, we can only conclude that the first is conditionally more diverse within some interval of q.

Figure VI presents the diversity profile for employment for $q = 0$ to $q = 2$ at the national industry level for years 1990, 2000, 2010, and 2020. The profile reveals that when using diversity measures that emphasize the number of firms ($q < 1$), diversity increased slightly from 1990 to 2020. However, using diversity measures that emphasize evenness ($q > 1$), such as HHI, diversity fell from 1990 to 2020. In contrast, balanced diversity of sales ($q = 1$) has remained roughly constant over the last 30 years. A second trait of the diversity profile is that the profiles for 2010 and 2020 are both steeper than 1990 and 2000 and also intersect their profiles. The steeper profile in 2010 and 2020 indicates that market shares have become

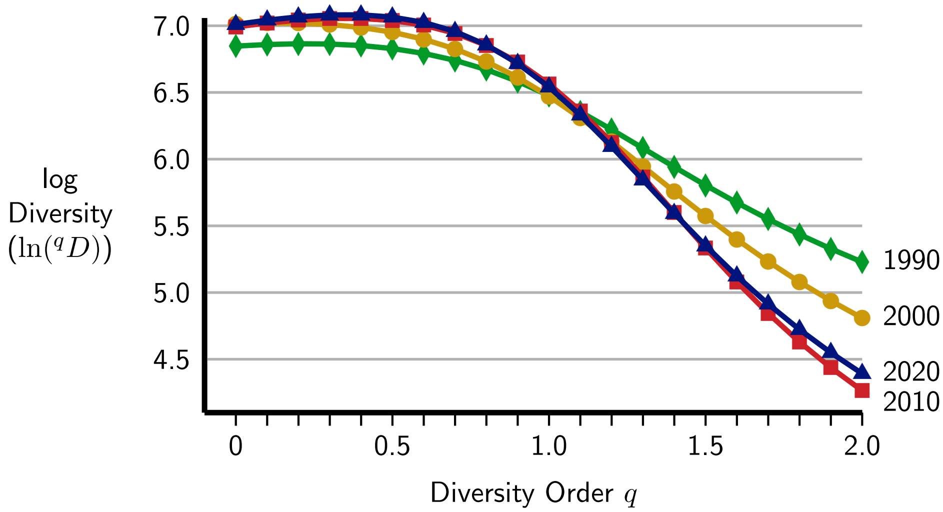


Figure VI
Employment Diversity Profiles

This figure presents the profile of diversity for industry-level employment for years 1990, 2000, 2010, and 2020. Diversity reflects the number of effective firms within an industry. The diversity order q controls how much weight is given to unevenness versus richness. Data are from NETS and include firms with at least 20 employees.

more uneven over time. However, the intersection indicates that we cannot conclude that diversity decreased unconditionally from 1990 to 2020.

The diversity profile results provide a more nuanced view of diversity than is possible by using just one indicator of diversity. In particular, the prior literature's sole focus on HHI overemphasizes the evenness of market shares and underemphasizes the importance of firm counts. Thus, the common wisdom about the time series of diversity is a conditional result, not an unconditional fact.

IV. Partitioning Diversity: Alpha, Beta, and Gamma

So far, we have only considered the diversity within a single market. However, it is also important to understand how diversity varies at different levels of aggregation. Markets can be defined at national, state, and local levels, or by industry, product, and brand levels and the disparity between aggregate and local HHI is an important finding in recent research on concentration (Rossi-Hansberg, Sarte, and Trachter, 2020).

Recent research studies local concentration using HHI. However, HHI does not allow for a systematic decomposition of aggregate HHI into local HHI. In contrast, true diversities allow for aggregate diversity to be systematically decomposed, using common units, into a component based on the diversity within a subdivision and a component based on the diversity across subdivisions.

Following the terminology of ecology, gamma diversity (${}^qD_\gamma$) is the diversity of the aggregate market; alpha diversity (${}^qD_\alpha$) is the diversity of a particular market; and beta diversity (${}^qD_\beta$) is the diversity between markets (Whittaker, 1972). For example, if gamma diversity is the diversity of the entire economy, alpha diversity could be defined as the diversity within industries and beta diversity as the diversity between industries. Likewise, alpha diversity could represent the diversity within particular geographic regions and beta diversity could represent the diversity between the regions.

Following Whittaker (1972), Jost (2007) defines the relationship between alpha, beta, and gamma diversity as,

$$ {}^qD_\gamma = {}^qD_{\overline{\alpha}} \times {}^qD_\beta, \tag{8} $$

where ${}^qD_{\overline{\alpha}}$ is a weighted average of the true alpha diversity of each sub-market, defined as

$$ {}^qD_{\overline{\alpha}} = \left[\frac{\sum_{j=1}^{M} w_j^q \times {}^q\lambda_j}{\sum_{j=1}^{M} w_j^q} \right]^{\frac{1}{1-q}} \qquad \text{for } q \neq 1, \tag{9} $$

and

$$ {}^{1}D_{\overline{\alpha}} = \exp\left[-\sum_{j=1}^{M}\sum_{i=1}^{N} w_j p_{ij} \ln p_{ij}\right] \qquad \text{for } q = 1, \tag{10} $$

for $j = 1, 2, \ldots, M$ sub-markets, where ${}^{q}\lambda_j$ is the diversity index for sub-market j, and w_j is the weight on the jth sub-market, such that $\sum w_j = 1$. Notice that ${}^{q}D_{\overline{\alpha}}$ is a generalized mean of the individual sub-markets' diversities, where weights are raised to the order q to preserve the relationship ${}^{q}D_{\gamma} = {}^{q}D_{\alpha} \times {}^{q}D_{\beta}$. In practice, beta diversity is derived from Equation 8 by first calculating alpha diversity using data on sub-markets and gamma diversity using pooled data.

As in the diversity of a single market, using true diversity versions of alpha, beta, and gamma diversity, provides intuitive and meaningful measures of diversity. In particular, beta diversity represents the number of distinct sub-markets, without any overlaps with any other sub-market, needed to account for the total diversity, given the diversity within an average sub-market, represented by alpha diversity. In particular, if the firms that operate in one market are completely different than those in another market, there is greater diversity across markets. If the same firms operate in both markets, there is less diversity across markets, holding the diversity within the markets fixed.

As a numerical illustration, consider an economy with three industries and six firms, as depicted in Panel A of Figure VII. Industry 1 includes 5 firms and produces an output of \$95 (Firm A: 40, B: 28, D: 3, E: 10, and F: 14); Industry 2 includes 2 firms and produces an output of \$60 (D: 37 and F: 23); and Industry 3 includes 4 firms and produces \$85 (B: 12, C: 40, E: 30, and F: 3).

The aggregate output is \$240, equally contributed by the six firms, which means that gamma diversity is 6; the aggregate diversity is equivalent to six firms with equal market shares. At $q = 1$, the gamma diversity of 6 is decomposed into an alpha diversity of 3 and a beta diversity of 2. This means that the diversity in this economy is equivalent to an economy

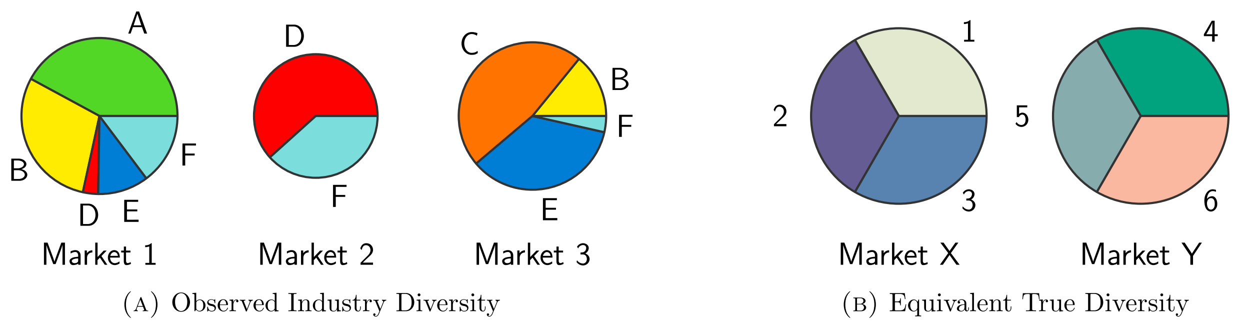


Figure VII
Decomposition Example

Panel A presents the market shares of firms A–F in three different markets, 1–3. Panel B presents the decomposition of the three markets into effective firms and effective markets, in which the total diversity in Panel A is equal to the total diversity in Panel B when $q = 1$ (balanced diversity).

with two industries (${}^1D_\beta = 2$), each with three unique firms (${}^1D_{\overline{\alpha}} = 3$) that command equal market shares only in one industry, as represented in Panel B of Figure VII. Though there are three actual industries, because some firms operate in multiple industries, the effective number of industries is smaller. The greater is the overlap in firms across industries, the fewer are the number of unique industries.

This example illustrates that using true diversities allows for a meaningful decomposition of aggregate diversity into within-industry concentration in the average industry and across-industry concentration. In contrast, though we can calculate HHIs of 0.30, 0.53, and 0.37 in the example industries and a HHI of 0.17 in the aggregate, these numbers are not particular intuitive on their own or in relation to aggregate diversity. Moreover, they do not account for overlapping firms across industries. Thus, HHI values do not reveal how each firm and industry contributes to total diversity.

A crucial axiom for decomposing diversity is that gamma diversity is always greater than or equal to alpha diversity (Jost, 2007). If this axiom is violated, it would imply that one sub-market has a larger diversity than the aggregate market. It would also imply that beta diversity is less than one, which means that there is less than one unique market. Both implications are unappealing and non-intuitive. Jost shows that when the weights of sub-markets are unequal, the only orders of q that satisfy this axiom are $q = 0$ (count diversity) and $q = 1$ (balanced diversity). Because count diversity ignores market shares, its usefulness is limited. This leaves balanced diversity as the only true diversity that allows for a consistent decomposition of diversity and provides an intuitive measure. This means that widely used diversity measures based on $q = 2$, including dominance diversity and HHI, do not provide valid decompositions of aggregate diversity into local diversity. This provides an additional advantage of balanced diversity over dominance diversity. In our empirical decompositions, we will decompose gamma diversity using $q = 0$, 1, and 2, with the caveat that when $q = 2$, the decomposition axiom may be violated.

V. The Decomposition of Diversity: Empirical Evidence

Rossi-Hansberg, Sarte, and Trachter (2020) show that while HHI has increased for industries defined at the national level, concentration has decreased for industries defined by smaller geographic regions. To illustrate our proposed methods, we evaluate this finding by decomposing national diversity into smaller regional components.

V.A. Gamma Diversity

First, we define gamma diversity using the widest definition available: the entire country. Thus, a firm's market share is defined by its total employment across all industries and geographic regions as a fraction of the total employment of all firms in the data. This is not intended to define a competitive market, but rather to quantify diversity at the largest scale possible. We then decompose gamma diversity into sub-markets defined by industries, counties, and industry-counties.

Column 1 of Table III presents gamma diversity estimates in the NETS data for 1990, 2000, 2010, and 2020, for count diversity, balanced diversity, and dominance diversity. Panel

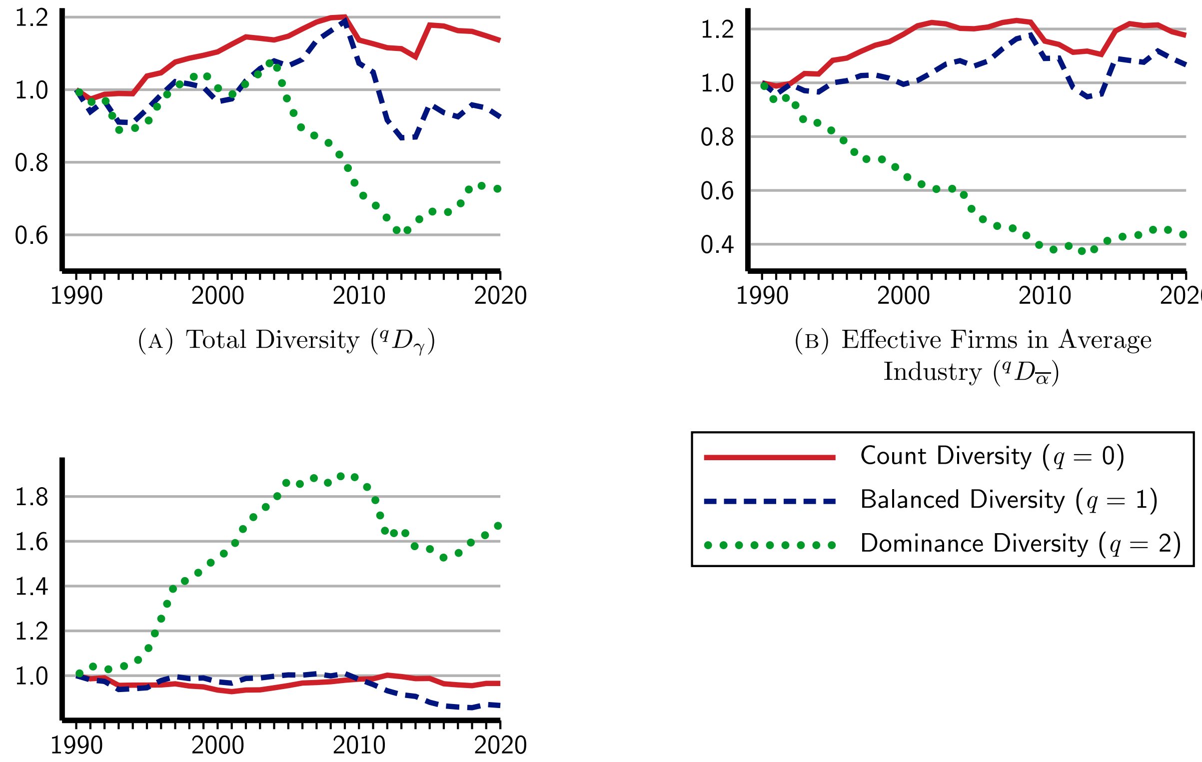
A of Figure VIII presents the time series of gamma diversity normalized by 1990 estimates. In 1990, NETS reports about 761,000 firms with at least 20 employees. This number grew to 864,000 by 2020, as illustrated in Figure VIII.

Table III
Industry, County, and Industry-County Diversity Decomposition

		Industry		County		Industry-County	
Year	Gamma	Alpha	Beta	Alpha	Beta	Alpha	Beta
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Count Diversity ($q = 0$)							
1990	760,968.0	941.4	808.3	349.9	2,174.7	3.9	197,419.8
2000	840,343.0	1,111.2	756.3	475.7	1,766.5	4.2	201,895.2
2010	865,102.0	1,087.2	795.7	464.0	1,864.3	4.0	216,096.8
2020	864,089.0	1,107.4	780.3	495.7	1,743.3	4.3	198,755.8
Panel B: Balanced Diversity ($q = 1$)							
1990	84,911.6	649.9	130.7	497.9	170.5	6.4	13,223.9
2000	82,100.8	646.3	127.0	558.3	147.1	6.8	12,122.4
2010	91,098.8	708.7	128.5	626.8	145.3	7.2	12,599.1
2020	78,508.9	693.2	113.2	610.9	128.5	7.9	9,975.2
Panel C: Dominance Diversity ($q = 2$)							
1990	4,033.0	186.5	21.6	230.8	17.5	5.5	727.0
2000	4,052.4	122.6	33.1	228.8	17.7	5.7	713.8
2010	2,888.6	71.1	40.6	283.4	10.2	6.4	452.5
2020	2,926.3	80.7	36.3	245.6	11.9	6.3	464.6

Notes: This table presents snapshots of alpha, beta, and gamma diversity of employment diversity at the SIC 4-digit nationwide industry level, county level, and industry-county level for diversity orders $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within an industry, county or industry-county. Beta diversity reflects the number of effective industries, counties, or industry-counties. Gamma diversity reflects the total diversity as the product of alpha and beta diversity. Data are from NETS and include firms with at least 20 employees.

Panel B of Table III reports that the balanced diversity across the entire economy in 1990 was equivalent to an economy with roughly 85,000 firms with equal shares of employment.



(A) Total Diversity ($^{q}D_{\gamma}$)

(B) Effective Firms in Average Industry ($^{q}D_{\overline{\alpha}}$)

(C) Effective Industries ($^{q}D_{\beta}$)

Figure VIII
Diversity Decomposition: Within and Between Industries

This figure presents the time-series of alpha, beta, and gamma diversity of employment at the SIC 4-digit nationwide industry level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within an industry (within diversity). Beta diversity reflects the number of effective industries (between diversity). Gamma diversity reflects the total diversity as the product of alpha and beta diversity. Data are from NETS and include firms with at least 20 employees.

Balanced gamma diversity declined by about 7% to 78,500 effective firms in 2020. In contrast, dominance diversity at the national level fell by about 27% from 1990 to 2020. These estimates represent the total diversity across the entire economy. Next, we decompose these aggregate estimates into diversity within smaller subdivisions and diversity across smaller subdivisions.

V.B. Industry-Level Diversity

Columns 2 and 3 of Table III decompose the national level of diversity into a within-industry and between-industry component, using 4-digit SIC code definitions. Panel B of Figure VIII presents the time series of average within-industry diversity. Balanced diversity is 649.9 effective firms in 1990. In comparison, the dominance diversity of the average industry in 1990 is 186.5 effective firms. Consistent with prior evidence on HHI, dominance diversity in the average industry fell significantly from 1990 to a low in 2013, and then partially reversed. In contrast, balanced diversity was slightly higher in 2020 than in 1990.

Next, column 3 of Table III and Panel C of Figure VIII report beta diversity at the industry level. The nearly constant solid line represents count diversity and indicates that the number of industries is close to identical across years. Using balanced diversity, the beta diversity between industries remained constant from 1990 to 2010 at around 128 effective industries, then fell to 113 effective industries in 2020. These units imply that the observed diversity in 1990 across industries is equivalent to an economy with 113 equally-sized industries, in which no firms operated in more than one industry. The decline in beta diversity reflects an increase in the prevalence of firms operating in multiple industries. In contrast, the beta dominance diversity increased substantially from 1990 to 2010, indicating that conglomerate firms with large employment shares became less common, and industries were less likely to be spanned by large firms.

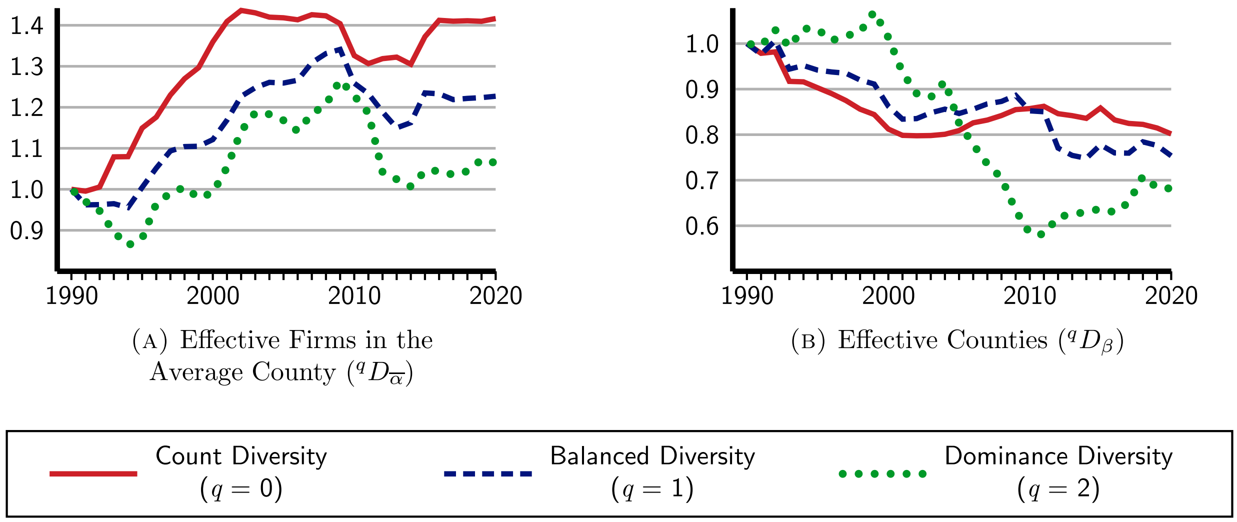
These results are important for at least two reasons. First, true diversities provide an intuitive meaning to concentration. In particular, using dominance diversity, we can characterize the average industry in 1990 as one with 186.5 symmetric firms. In 2010, there were 71.1 symmetric firms. If we use balanced diversity, there were 649.9 symmetric firms in the average industry in 1990, rising to 708.7 symmetric firms in 2010. This provides a more meaningful and intuitive interpretation of concentration than does HHI. For instance, though the change in dominance diversity represents a substantial decline of 62%, it is not obvious how an industry with 187 equally-size firms is substantially different economically than one with 71 equally-sized firms.

Second, these results provide a quantitative estimate of the importance of within- versus between-industry diversity for national diversity. Based on balanced diversity, the results show that the decline in aggregate gamma diversity began around 2010 and was driven in part by a decline in diversity between industries, even as within-industry diversity remained relatively stable from 2012 to 2020. This presents an explanation for changes in diversity that is not obvious from observing industry averages alone.

V.C. County-Level Diversity

Figure IX and columns 4 and 5 of Table III present the decomposition of aggregate diversity by county. Based on balanced diversity, there were about 500 effective firms in the average county in 1990. The between-counties diversity was only 170.5 for employment, relative to the roughly 2,500 counties in the sample. This shows that the average firm-share of employment is highly similar across counties. From 1990 to 2020, balanced diversity within the average county increased by 23% in the average county, while between-county balanced diversity decreased by about 25%. This reveals that the decrease in aggregate gamma diversity at the national level was driven by less diversity across counties, even while within-county diversity increased. These results indicate that as firms enter new geographic areas they increase the diversity of employment at the local level and also make local employment markets more homogenous across counties. The same rough patterns appear for dominance diversity which indicates that the homogenization across counties is driven by firms of all sizes.

To illustrate the importance of the choice of q, Figure X presents the cross-sectional variation in diversity across counties in California in 2020. These figures represent the abnormal level of diversity, relative to population for $q = 0, 1, 2$. In particular, abnormal diversity is the residual from the regression $\ln(1 + {}^q D) = \alpha + \beta \ln(1 + \text{population}) + \varepsilon$, which is estimated



(A) Effective Firms in the Average County (${}^qD_{\overline{\alpha}}$)

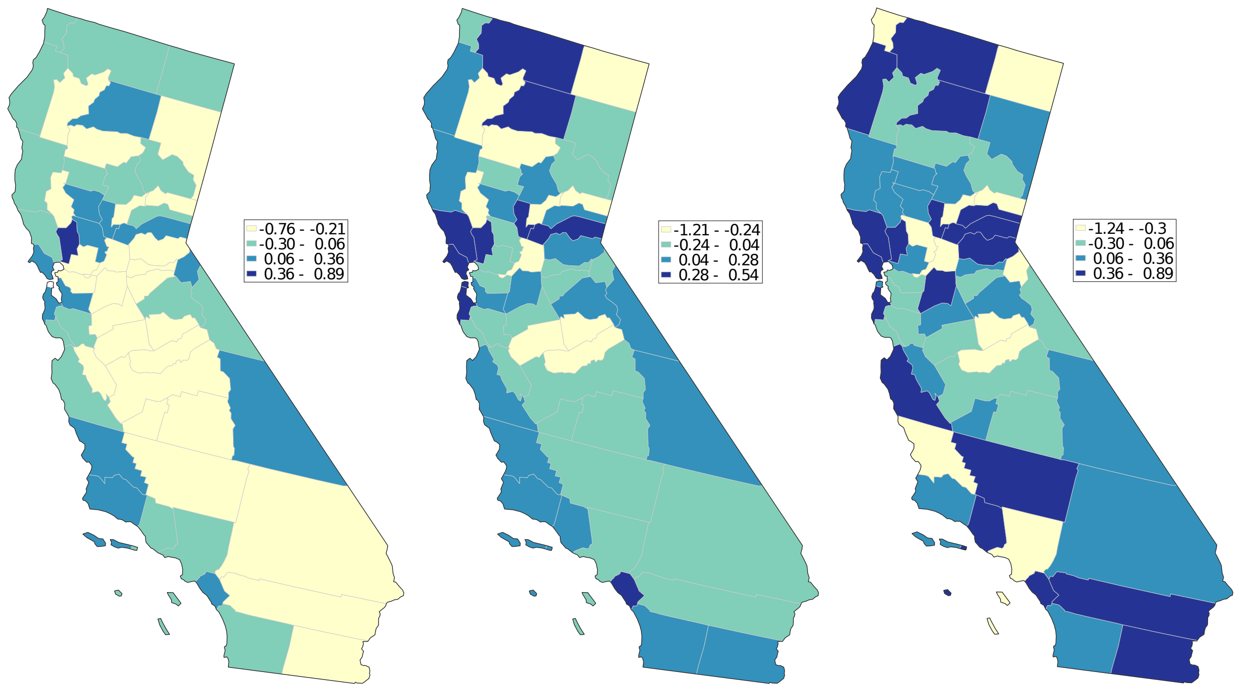
(B) Effective Counties (${}^qD_{\beta}$)

Figure IX

Diversity Decomposition: Within and Between Counties

This figure presents the time-series of alpha and beta diversity of employment at the county level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within a county or industry-county. Beta diversity reflects the number of effective counties or industry-counties. Data are from NETS and include firms with at least 20 employees.

over all counties in the U.S. Using count diversity, California counties tend to have less than normal numbers of firms given the size of their populations. However, using dominance diversity, California counties tends to have more effective firms than predicted based on their populations. This implies that when more weight is given to a firm's employment share, counties in California tend to be more diverse than average. This suggests that the employment shares in California are relatively more even than in other places. This pattern also varies across California counties. For instance, while the counties connecting Sacramento to the Lake Tahoe area become more diverse when more weight is given to employment shares, Los Angeles county becomes less diverse. This indicates that in Los Angeles, employers with large shares of employment decrease the number of effective firms, while the counties near Sacramento are characterized by relatively even market shares. Panel B represents the



(A) Count Diversity (B) Balanced Diversity (C) Dominance Diversity

Figure X

Abnormal Diversity per Capita in California in 2020

This figure presents abnormal diversity for $q = 0$ (Panel A), $q = 1$ (Panel B), and $q = 2$ at the county level in California in 2020. Abnormal diversity is the county-level residual in a regression of $\log(1 +^q D)$ on a constant and log(1+population). Darker values represent higher abnormal diversity per capita. Categories of abnormal diversity are based on quartile values across all counties in the US. Internet Appendix Figure III presents maps of the lower 48 states. Data are from NETS and include firms with at least 20 employees.

balanced version of these two extremes. Maps of the entire U.S. are reported in Internet Appendix Figure III.

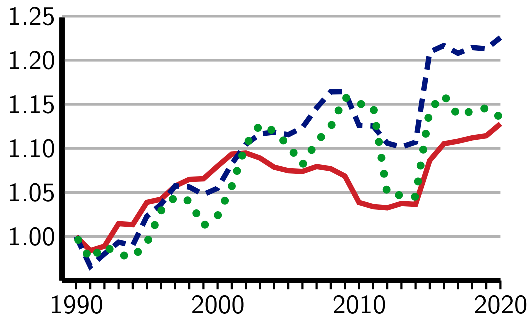
V.D. Industry-County-Level Diversity

Figure XI and columns 6 and 7 of Table III decompose aggregate diversity into industry-counties. The within-market balanced diversity in the average industry-county increased significantly from 6.4 firms in 1990 to a peak in 2009 and then decreased slightly by 2020 to 7.9 firms, which still represents a 23% increase from 1990 levels. In contrast, beta diversity of employment between industry-counties decreased significantly from 13,224 effective industry-counties down to 9,975. A similar pattern appears for dominance diversity. These results indicate that industry-counties have become significantly more homogenous over time, even as the diversity within the average industry-county has increased.

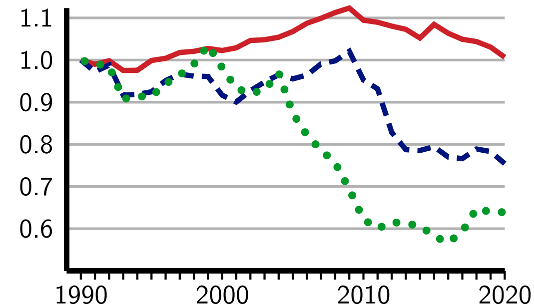
These results confirm the findings in Rossi-Hansberg, Sarte, and Trachter (2020), who show that HHI decreased within local levels, while it increased at the national level. We extend their results because the decomposition method we use allows us to quantify these trends using intuitive, but precise estimates. Using effective firms and industry-counties allows us to quantify the effect originally documented in Rossi-Hansberg et al. in a meaningful way. Moreover, to identify the effects of large firms on this trend, Rossi-Hansberg et al. compare HHI measures after dropping large firms. In our setting, by using a higher order of q, we put more weight on the unevenness of market shares. We find that the within industry-county dominance diversity of sales remained roughly constant over time, but the between-industry-county diversity fell considerably.

V.E. Diversity Profiles at the County and Industry-County Level

Figure XII provides diversity profiles at the county and industry-county levels for years 1990, 2000, 2010, and 2020. Because the profiles are steep, we take the log values of diversity to help understand the patterns by year. The profiles reveal that alpha diversity in 2020 is unconditionally higher than diversity in 1990 at both the county and industry-county levels. This means that for any choice of weighting of richness and evenness, diversity in 2020 was substantially higher than in 1990. However, in 2020, the profile at the county-level is significantly steeper, indicating greater unevenness in employment shares within the average county.



(A) Effective Firms in the Average Industry-County ($^{q}D_{\overline{\alpha}}$)



(B) Effective Industry-Counties ($^{q}D_{\beta}$)

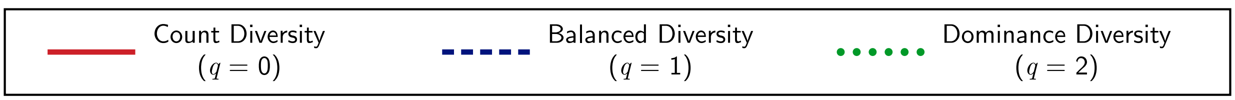
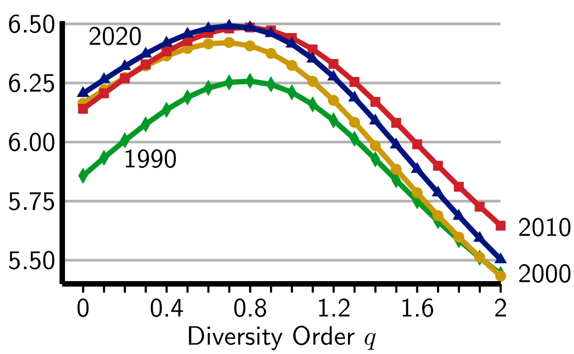


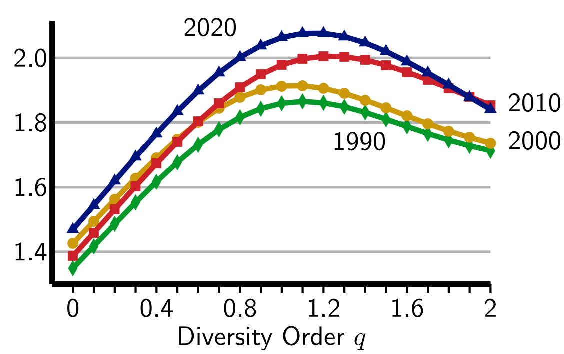
FIGURE XI

DIVERSITY DECOMPOSITION: WITHIN AND BETWEEN INDUSTRY-COUNTIES

This figure presents the time-series of alpha and beta diversity of employment at the industry-county level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within a county or industry-county. Beta diversity reflects the number of effective counties or industry-counties. Data are from NETS and include firms with at least 20 employees.



(A) County Level



(B) Industry-County Level

FIGURE XII

DIVERSITY PROFILES AT COUNTY AND INDUSTRY-COUNTY LEVELS

This figure presents the profile of log alpha diversity at the county level (Panel A) and industry-county level (Panel B) for years 1990, 2000, 2010, and 2020. Alpha diversity reflects the number of effective firms within a county or industry-county. The diversity order q controls how much weight is given to unevenness versus richness. Data are from NETS and include firms with at least 20 employees.

VI. Additional Tests

VI.A. Non-Imputed Data

As mentioned previously, Section A6 of the Internet Appendix provides results of our analysis using only non-imputed data in NETS. The absolute values of effective firms are mechanically reduced because our sample is smaller. However, the time series patterns and diversity profiles provide comparable analysis as using the full dataset. The qualitative patterns are nearly identical in the non-imputed data as in the imputed-data. In particular, at the national level, dominance diversity of the average industry declined, consistent with prior research, while balanced diversity increased. Likewise, the decomposition of aggregate diversity into county and industry-county diversity follows nearly identical patterns as in the full dataset. The robustness of these results mitigates concerns that imputation of data in NETS drives our results.

VI.B. Estimates of Sales Diversity: 1963–1992

One of the advantages to the analytic formulations of diversity in Equations 5 and 6 is that they allow diversity to be estimated with relatively little information. Under an assumption of lognormal or Pareto-distributed firm sizes, if the total number of firms N is known and if we have parameter estimates of σ and α, we can calculate ${}^qD$. Moreover, given estimates of HHI and N, we could back out the distribution parameters, then calculate diversity for any value of q under each firm-size distribution assumption. This analysis depends on the validity of the distributional assumptions, but if these assumptions are reasonably close to reality, then the analytic formulas provide a reasonable estimate of diversity.

Unfortunately, it is rare to observe all necessary parameters in public data. For example, though the U.S. Economic Census reports the number of firms per industry, it only reports HHI for the top 50 firms in the industry. To estimate the parameters, we require HHI for the entire population of firms, not just the top 50. To address this limitation, we estimate

σ and α using the 4-, 8-, 20-, and 50-firm concentration ratios reported in the Economic Censuses from 1963 to 1992 using the machine-readable tabulations provided by Keil (2017). In particular, for a given value of N in the Census reports, we find the parameter value that produces an estimated concentration ratio from a Monte Carlo simulation that is nearest the concentration ratio reported by the Census. The Census only reports enough information to estimate sales diversity at the 4-digit SIC code-level for the manufacturing sector.

Internet Appendix Figure II reports the results of the simulation. We find that diversity for each diversity order q declined from 1963 to the mid 1970s, but then increased through 1992. Dominance diversity was roughly equivalent in 1992 as it was in 1963. The increase in diversity during the 1970s and 1980s may reflect the general trend towards deregulation. The figure also reveals that the three diversity measures comove closely over time. This indicates that the driving force is the number of firms in the sector, not the evenness of their market shares. Indeed, the estimate of σ is highly persistent over the sample period.

This analysis illustrates the benefit of the analytical formulation of diversity under an assumed distribution of firm sizes. If a distributional assumption is reasonable, to calculate the full set of diversity measures only requires an estimate of the shape parameter and the number of firms in the market.

VII. Conclusion

Though concentration is ubiquitous in economic research, there is little discussion of its underlying assumptions or how best to measure it. This paper adopts advances from ecology to introduce a generalized measure of economic diversity that makes the underlying assumptions explicit. This measure has three key advantages. First, it allows the researcher to vary the weight given to the two components of diversity: the number of firms and their market shares. By using a range of weightings, this measure provides a more complete understanding of diversity. Second, the measure is in units of 'effective firms,' defined as the number of equally-sized firms needed to match the observed diversity. Using these units allows for an intuitive understanding of diversity and also satisfies the replication principle which implies that a doubling of the number of effective firms is equivalent to a doubling of diversity. Third, the measure can be systematically decomposed into components that reflect the diversity within an industry and the diversity between industries. The most widely-used measure of concentration in economics, HHI, lacks these advantages.

We apply this method to the diversity of employment. We show that stylized facts of large increases in concentration are conditional on the use of diversity measures, such as HHI, that over-weight market shares relative to firm counts. As an alternative to HHI, we present *balanced diversity* which gives equal weight to firm counts and market shares. We show that balanced diversity has an equivalent economic meaning as HHI, but displays superior statistical properties and less small-sample bias than HHI. In contrast to the stylized facts, we find that employment diversity increased, rather than decreased, when measured with balanced diversity. Further empirical decompositions show that well-known stylized facts are sensitive to methodology choices.

We believe the diversity methods we have presented here provide a useful and flexible framework for future economic research. Though we have used market shares of firms to describe the method, it is not limited to this setting. For instance, in unreported tests, we use our decomposition approach to quantify the the rise in 'niche' household consumption as documented in Neiman and Vavra (2023). In particular, we find that the average household purchased 311 effective products in 2004, but only 221 in 2019, a decline of about 20%. At the same time, the number of effective households increased by 64%, indicating a rise in 'niche' households. Another use of this method could be to calibrate the diversity parameter q to maximize an objective function, such as a correlation between diversity and markups. Alternatively, this method could help motivate new theoretical research that incorporates a broader definition of diversity.

In addition, the broader ecology literature provides many future avenues for economics research. For example, ecologists have developed diversity measures that account for shared evolutionary history and biological traits. A natural analogy for economic diversity is to account for the similarity of firms' production functions and customer bases. Another avenue for future research is to apply new methods from ecology to address sampling bias. Typical ecological surveys are likely to under-represent rare species, which will affect estimates of diversity. Economists could use ecological sampling methods to estimate the size of the informal economy that is absent in administrative survey data.

References

Adelman, M. A., "Comment on the "H" Concentration Measure as a Numbers-Equivalent," *Review of Economics and Statistics,* 51 (1969), 88–101.

Aitchison, J., and J. A. C. Brown, *The Lognormal Distribution With Special Reference To Its Uses in Economics* (Cambridge, England: Cambridge University Press, 1958).

Amirapu, Amrit, and Michael Gechter, "Labor Regulations and the Cost of Corruption: Evidence from the Indian Firm Size Distribution," *Review of Econoimcs and Statistics,* 102 (2020), 34–38.

Artige, Lionel, and Sousso Bignandi, "The firm size distribution: Evidence from Belgium," *Applied Economics,* 55 (2023), 907–923.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," *Quarterly Journal of Economics,* 135 (2020), 645–709.

Autor, David, Christina Patterson, and John Van Reenen, "Local and National Concentration Trends in Jobs and Sales: The Role of Structural Transformation," *NBER Working Paper,* (2023).

Axtell, Robert L., "Zipf Distribution of U.S. Firm Sizes," *Science,* 293 (2001), 1818–1820.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker, "An Assessment of the National Establishment Time Series (NETS) Database," *Finance and Economics Discussion Series, Federal Reserve System,* (2017).

Bee, Marco, Massimo Riccaboni, and Stefano Schiavo, "Where Gibrat meets Zipf: Scale and scope of French firms," *Physica A: Statistical Mechanics and its Applications,* 481 (2017), 265–275.

Bernstein, Shai, Timothy McQuade, and Richard R. Townsend, "Do Household Wealth Shocks Affect Productivity? Evidence from Innovative Workers During the Great Recession," *Journal of Finance,* 76 (2021), 57–111.

Borisov, Alexander, Andrew Ellul, and Merih Sevilir, "Access to public capital markets and employment growth," *Journal of Financial Economics,* 141 (2021), 896–918.

Bresnahan, Timothy F., "17," in *Empirical Studies of Industries with Market Power,*Richard Schmalensee and Robert Willig, eds.

Cabral, Luís M. B., and José Mata, "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review,* 93 (2003), 1075–1090.

Chao, Anne, Chun-Huo Chiu, and Lou Jost, "Unifying Species Diversity,Phylogenetic Diversity,Functional Diversity, and Related Similarity andDifferentiation MeasuresThrough Hill Numbers," *Annual Review of Ecology, Evolution, and Systematics,* 45 (2014), 297–324.

Chao, Anne, and Lou Jost, "Estimating diversity and entropy profiles via discovery rates of new species," *Methods in Ecology and Evolution,* 6 (2015), 873–882.

Chao, Anne, and Carlo Ricotta, "Quantifying evenness and linking it to diversity, beta diversity, and similarity," *Ecology,* 100 (2019), 1–15.

Chiu, Chun-Huo, Lou Jost, and Anne Chao, "Phylogenetic beta diversity, similarity, and differentiation measuresbased on Hill numbers," *Ecological Monographs,* 84 (2014), 21–44.

Chow, Melissa, Teresa C. Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White, "Redesigning the Longitudinal Business Database," *US Census Bureau Center for Economic Studies,* (2021).

Clarke, Roger, "On the lognormality of firm and plant size distributions: Some U.K. evidence," *Applied Economics,* 11 (1979), 415–434.

Crouzet, Nicolas, and Neil R. Mehrotra, "Small and Large Firms over the Business Cycle," *American Economic Review,* 110 (2020), 3549–3601.

Daly, Aisling J., Jam M. Baetens, and Bernard De Baets, "Ecological Diversity: Measuring the Unmeasurable," *Mathematics,* 6 (2018), 119.

Dansby, Robert E., and Robert D. Willig, "Industry Performance Gradient Indexes," *American Economic Review,* 69 (1979), 249–260.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics,* 135 (2020), 561–644.

de Wit, Gerrit, "Firm size distributions: An overview of steady-state distributions resulting from firm dynamics models," *International Journal of Industial Organization,* 23 (2005), 423–450.

Demsetz, Harold, "Why Regulate Utilities?," *Journal of Law and Economics,* 11 (1968), 55–65.

——— , "Industry Structure, Market Rivalry, and Public Policy," *Journal of Law and Economics,* 16 (1973), 1–9.

——— , "Two Systems of Belief About Monopoly," in *Industrial Concentration: The New Learning*, H. Goldschmid H. M. Mann and J. F. Weston, eds. (Little, Brown Company, 1974).

di Giovanni, Julian, Andrei A. Levchenko, and Romain Rancière, "Power laws in firm size and openness to trade: Measurement and implications," *Journal of International Economics,* 85 (2011), 42–52.

Donsimoni, Marie-Paule, Paul Geroski, and Alexis Jacquemin, "Concentration Indices and Market Power: Two Views," *Journal of Industrial Economics,* 32 (1984), 419–434.

Ellison, Aaron, "Partitioning Diversity," *Ecology,* 91 (2010), 1962–1963.

Encaoua, David, and Alexis Jacquemin, "Degree of Monopoly, Indices of Concentration and Threat of Entry," *International Economic Review,* 21 (1980), 87–105.

Faccio, Mara, and Hung-Chia Hsu, "Politically Connected Private Equity and Employment," *Journal of Finance,* 72 (2017), 539–574.

Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist, "What is a Patent Worth? Evidence from the U.S. Patent "Lottery"," *Journal of Finance,* 75 (2020), 639–682.

Federal Trade Commission, "Policy Statement Regarding the Scope of Unfair Methods of Competition Under Section 5 of the Federal Trade Commission Act," *Federal Trade Commission Policy Statement,* (2022).

Finkelstein, Michael O., and Richard M. Friedberg, "The application of an entropy theory of concentration to the Clayton Act.," *The Yale Law Journal,* 76 (1967), 677–717.

Gaffeo, Edoardo, Mauro Gallegati, and Antonio Palestrini, "On the size distribution of firms: Additional evidence from the G7 countries," *Physica A,* 324 (2003), 117–123.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review,* 106 (2016), 3439–3479.

Gibrat, Robert, *Les inégalit'es économiques applications: aux inegaliteés des richesses, á la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc. d'une loi nouvelle, la loi de l'effet proportionnel* (Paris: Recueil Sirey, 1931).

Giovanni, Julian Di, and Andrei A Levchenko, "Firm entry, trade, and welfare in Zipf's world," *Journal of International Economics,* 89 (2013), 283–296.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, "Are U.S. Industries Becoming More Concentrated?," *Review of Finance,* 23 (2019), 697–743.

Gutiérrez, Germán, and Thomas Philippon, "Declining Competition and Investment in the U.S.," *NBER Working Paper,* (2017).

Hannah, Leslie, and J. A. Kay, *Concentration in Modern Industry: Theory, Measurement, and the U.K. Experience* (London and Basinstroke: The MacMillan Press Ltd., 1977).

Hart, P.E., "Moment Distributions in Economics: An Exposition," *Journal of the Royal Statistical Society. Series A (General),* 138 (1975), 423–434.

Herfindahl, Orris Clemens, 1950, "Concentration in the Steel Industry," Ph.D. thesis Columbia University.

Hill, Mark O., "Diversity and evenness: a unifying notation and its consequences," *Ecology,* 54 (1973), 427–432.

Hirschman, Albert O., *National Power and the Structure of Foreign Trade* (Berkeley and Los Angeles: University of California Press, 1945).

Jost, Lou, "Entropy and Diversity," *Oikos,* 113 (2006), 363–375.

——— , "Partitioning Diversity Into Independent Alpha and Beta Components," *Ecology,* 88 (2007), 2427–2439.

——— , "The Relation Between Evenness and Diversity," *Diversity,* 2 (2010), 207–232.

Keil, Jan, "The trouble with approximating industry concentration from Compustat," *Journal of Corporate Finance,* 45 (2017), 467–479.

Kolko, Jed David, David Neumark, and Ingrid Lefebvre-Hoang, *Business location decisions and employment dynamics in California* (Public Policy Instit. of CA, 2007).

Kondo, Illenin O., Logan T. Lewis, and Andrea Stella, "Heavy Tailed, but not Zipf: Firm and Establishment Size in the U.S.," *Working Paper,* (2022).

Levy, Moshe, and Sorin Solomon, "Power laws are logarithmic Boltzmann laws," *International Journal of Modern Physics C,* 7 (1996), 595.

Luttmer, Erzo G. J., “Selection, Growth, and the Size Distribution of Firms,” *Quarterly Journal of Economics,* 122 (2007), 1103–1144.

Montebruno, Piero, Robert J. Bennett, Carry van Lieshout, and Harry Smith, “A tale of two tails: Do Power Law and Lognormal models fit firm-size distributions in the mid-Victorian era?,” *Physica A,* 523 (2019), 858–875.

Neiman, Brent, and Joseph Vavra, “The Rise of Niche Consumption,” *American Economic Journal: Macroeconomics,,* (2023).

Patil, G. P., and C. Tallie, “ in *An overview of diversity,*J. F. Grassle G. P. Patil W. Smith and C. Tallie, eds.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicolas Trachter, “Diverging Trends in National and Local Concentration,” *NBER Macoreconomics Annual,* 35 (2020), 115–150.

Roswell, Michael, Jonathan Dushoff, and Rachael Winfree, “A Conceptual Guide to Measuing Species Diversity,” *Oikos,* 130 (2021), 321–338.

Shannon, Claude Elwood, “A mathematical theory of communication,” *The Bell System Technical Journal,* 27 (1948), 379–423.

Simon, Herbert A., “On a class of skew distribution functions,” *Biometrika,* 52 (1955), 425–440.

——— , “Some further notes on a class of skew distribution functions,” *Information and Control,* 3 (1960), 80–88.

Simpson, E.H., “Measurement of Diversity,” *Nature,* 163 (1949), 688–688.

Stigler, George J., “A Theory of Oligopoly,” *Journal of Political Economy,* 72 (1964), 44–61.

——— , “Comment: The application of an entropy theory of concentration to the Clayton Act.,” *The Yale Law Journal,* 76 (1967), 718–720.

Syverson, Chad, “Macroeconomics and Market Power: Context, Implications, and Open Questions,” *Journal of Economic Perspectives,* 33 (2019), 23–43.

Tirole, Jean, *The Theory of Industrial Organization* (Cambridge, MA: The MIT Press, 1988).

Whittaker, Robert H., “Evolution and measurement of species diversity,” *Taxon,* 21 (1972), 213–251.

Zingales, Luigi, “Towards a Political Theory of the Firm,” *Journal of Economic Perspectives,* 31 (2017), 113–130.

Internet Appendix

"The Anatomy of Concentration: New Evidence from a Unified Framework"

Kenneth R. Ahern, Lei Kong, and Xinyan Yan

A1. Cournot Competition and Generalized Diversity

In this section we show how diversity is related to profitability in Cournot models of competition. First, we provide numerical examples to show that HHI is not a sufficient statistic for profitability under symmetric or asymmetric cost assumptions. Then we provide analytical solutions that demonstrate the same result.

The non-cooperative Cournot model of competition generates market shares based on the cost functions of the firms. If cost functions are identical, then all firms have equal market shares, so there is perfect evenness, and all firms have equal profits. In this case, only the number of firms (richness) matters for profits. More firms reduce total profits in aggregate. If cost functions are not identical, then firms with lower costs will have higher market shares and there will be unevenness, which will affect diversity.

For example, assume a linear aggregate demand curve $P(q) = A - b\sum^n q_i$, where $q = \sum^n q_i$ and q_i is the output of firm i. Firm i profit is $\pi_i = P(q)q_i - c_i q_i$, where c_i is the marginal cost of firm i. The equilibrium output per firm q_i, total output Q and P is as follows,

$$q_i = \frac{1}{b}\left(\frac{A + \sum^n c_i}{n+1} - c_i\right)$$

$$Q = \frac{1}{b}\left(\frac{nA - \sum^n c_i}{n+1}\right)$$

$$P = \frac{A + \sum^n c_i}{n+1}$$

I.A. Numerical Examples

A.1. Example 1

Let $A = 10$ and $b = 1$ in a market with two firms who have zero marginal costs, each firm has an equal market share and produces 3.3 units of output at a price of 3.3. The HHI of this market is 0.5 and the aggregate profit is 22.2.

If we add another firm with zero marginal cost, each firm produces 2.5 units at a price of 2.5. The HHI is 0.33 and the aggregate profit is 18.75. Both markets are perfectly even, but the second market has more diversity (less concentration) because it has more richness.

In the third market, assume that firm 1 has zero marginal costs, but two other firms each have marginal costs of 3. In this case, the low-cost firm produces 4 units and the other two firms produce 1 unit at a price of 4. Therefore, the low-cost firm has a market share of 0.67 and the other two firms have market shares of 0.167. Now, the HHI of this market is 0.5, equivalent to the first market with two symmetric firms.

Compared to the first market, HHI decreased with the addition of another firm, but increased by the unevenness in their cost functions. In this case, these two forces balanced such that the HHI remained at 0.5. However, the aggregate profit in the third market is 18, compared to 22.2 in the first market, even though the HHI is the same. Thus, HHI does not predict profitability.

A.2. Example 2

Consider a more extreme example using the same aggregate demand function, but different numbers of firms and cost functions. In the first market there are five firms. The low-cost firm has a marginal cost of 2 and the four high cost firms have marginal costs of 3.2. At a market price of 4.13, the low-cost firm will sell 2.13 units and earn profit of 4.55, while the high-cost firms will each sell 0.93 units and earn profits of 0.871. The low-cost firm has a market share of 0.36 and the high-cost firms each have market shares of 0.16. This generates an HHI of 0.233 in a market in which aggregate profit is 8.03 and average profit per firm is 1.606.

In the second market, there are 31 firms. This serves to decrease concentration. However, assume that the low-cost firm has zero marginal costs, while the other 30 firms have marginal costs of 3.1. At a market price of 3.215, the low-cost firm will sell 3.22 units and earn profits of 10.34. The remaining firms each sell 0.119 units and earn profits of 0.014. The greater disparity in costs leads the low-cost firm to command 0.47 of the market, while each of the high cost firms has a market share of 0.0175. This generates an HHI of 0.234, nearly identical to the first market. However, the aggregate profit is 10.77 and average profit per firm is 0.347.

This simple example shows that two markets with the same HHI have different aggregate profits and different profits per firm. Thus, this example demonstrates that HHI is not a sufficient statistic in a Cournot model to identify profitability.

I.B. Analytic Relationships

One can show that a firm's profit in this basic Cournot model can be expressed as $\pi_i = \frac{P^\star}{\varepsilon} q_i m_i$, where ε is the price elasticity of demand at the equilibrium quantity $Q^\star$ and price $P^\star$, and m_i is firm i's market share. Therefore, the aggregate profit in this simple Cournot model is related to HHI as follows:

$$\sum^{n} \pi_i = \frac{Q^\star P^\star}{\varepsilon} HHI \tag{A.1}$$

This shows a direct relationship between aggregate profit and HHI. However, the same forces that shape HHI, richness and evenness of firm cost structures, also influence elasticity and the equilibrium quantity and price. Therefore, it is possible to have the same HHI for different profit levels.

We can rewrite the equation for sum of profit using the fundamentals of the market. First, normalize $A = 1$ for simplicity, then,

$$\sum^{n} \pi_i = \frac{1}{b} \frac{n - \sum^{n} c_i}{n+1} HHI \tag{A.2}$$

Notice that the first part of this expression is determined solely by the richness of the market. Under mild assumptions about costs, the unevenness of costs does not influence the total profit through the market equilibrium and elasticity channel. Instead, the HHI component captures both richness and evenness of costs.

In a symmetric Cournot model, where all firms have equal costs, we can rewrite the equation for aggregate profits as

$$\sum^{n} \pi_i = \frac{(1-c)^2}{b} \left(\frac{n}{n+1} \right)^2 \frac{1}{n} \tag{A.3}$$

where $\frac{1}{n}$ is the value of HHI when market shares are equal. Thus, for symmetric Cournot, only richness is related to aggregate profit.

A2. The Trade off Between Richness and Evenness: Elasticity

There is a trade-off between richness and evenness such that a reduction in true diversity can be obtained by reducing richness or reducing evenness. In particular, for any order q, when evenness is maximized, true diversity equals to the number of species, n. If we maintain perfect evenness, but remove one species, true diversity will equal $n - 1$. Alternatively, we can reduce diversity by transferring abundance and making the distribution uneven. Starting from a perfectly even assemblage, in which each species has a relative abundance of $\frac{1}{n}$, the transfer that provides the most unevenness for the same amount of transfer is to increase the relative abundance of one species by $\Delta > 0$ through transfers of equal amounts from all other species equal to $\frac{\Delta}{n-1}$. We denote the relative abundance of the species with higher abundance as $p_{max} = \frac{1}{n} + \Delta$. For each order q, there is a value of Δ such that both methods (decreasing richness or decreasing evenness) will generate the same reduction in true diversity.[10]

We can write this equivalence as follows:

$$\left[\left(\frac{1}{n}+\Delta\right)^q + (n-1)\left(\frac{1}{n}-\frac{\Delta}{n-1}\right)^q\right]^{\frac{1}{1-q}} = n-1. \tag{A.4}$$

When $q = 1$, the above equation is undefined. In the limit, as $q \to 1$, the equation is:

$$e^{-\left(\frac{1}{n}+\Delta\right)\ln\left(\frac{1}{n}+\Delta\right)-(n-1)\left(\frac{1}{n}-\frac{\Delta}{n-1}\right)\ln\left(\frac{1}{n}-\frac{\Delta}{n-1}\right)} = n-1. \tag{A.5}$$

The left-hand sides of these equations reflects a change in true diversity by decreasing evenness while holding fixed the richness n. From the perfect evenness benchmark in which each species has a share $\frac{1}{n}$, we add Δ to one species and subtract the same amount from the other species, split equally across the other $n-1$ species. The right-hand side reflects a change in true diversity by decreasing richness while holding evenness fixed. With perfect evenness, true diversity simply decreases by one. Solving this equation for Δ explains how large Δ must be for a given n and q, to reduce diversity by one.

To convert Δ into a more intuitive concept, we calculate the elasticity of substitution, ε, between p_{max} and n. This represents the percentage increase in p_{max} (a measure of evenness) that is required to maintain the same diversity for a one percent decrease in richness.

$$\varepsilon = \frac{\frac{\Delta}{\frac{1}{n}}}{-\frac{1}{n}} = -\Delta n^2 \tag{A.6}$$

[10] A trivial solution is $\Delta = -\frac{1}{n}$ which holds for all q. It is also the only solution if $q = 0$.

To our knowledge, a generalized closed-form solution for ε as a function of n and q does not exist. However, closed form solutions exist for $q = 0.5, 1$, and 2 as follows:

$$\varepsilon = \begin{cases} 4 - 3n & \text{if } q = 0.5 \text{ and } n > \frac{4}{3} \\ 1 - (e-1)n & \text{if } q = 1 \\ -n & \text{if } q = 2 \end{cases} \tag{A.7}$$

First, note that the elasticity depends on n. This is consistent with Jost (2010), who shows that the decomposition of diversity into independent evenness and richness components is impossible. In other words, the elasticity of substitution between evenness and richness that maintains diversity depends on the level of richness, n.

Second, note that for a fixed level n, the elasticity becomes larger in magnitude as q increases. This means for the same decrease in diversity caused by a reduction in the number of firms, the percentage decrease in evenness must be larger for larger q. For example, a change from a diversity of 8 to 7 represents a 12.5% decrease in richness. To achieve the same change through evenness requires an increase of 100% in p_{max} if $q = 2$, an increase of 159% if $q = 1$, and an increase of 250% if $q = 0.5$. This reflects that lower orders of q are more sensitive to richness and less sensitive to evenness. Therefore, to maintain the same level of diversity as richness changes requires a larger change in evenness when q is low.

Next, for a fixed level q, the magnitude of the elasticity of substitution increases as n increases. For example, if $q = 1$, to maintain constant diversity, a 1% decrease in richness requires a 16.2% increase in p_{max} if $n = 10$ and a 33.3% increase if $n = 20$. This reflects that the tradeoff between richness and evenness requires larger changes in unevenness when richness is large.

A final note of interest is that when $q = 2$, the solution for Δ is $\frac{1}{n}$. This means that to decrease the true diversity of a perfectly even market by one, one firm must have exactly double the share of the share in an even market. For example, in a market with four equally sized firms, where each firm has a share of 25%, the true diversity is four. If we double the share for one firm, taking away equally from the others, so we have shares of 50%, 16.7%, 16.7%, and 16.7%, true diversity is exactly three. Likewise, a market with five equally-size firms with shares of 20% has a true diversity of five, while a market with five firms with market shares of 40%, 15%, 15% 15%, and 15%, has a true diversity of four, when $q = 2$.

In sum, this exercise shows that the trade-off between evenness and richness depends on both the level of n and the order q. When q is small, to maintain diversity as richness falls requires a larger increase in unevenness than when q is larger. Because the large majority of research on concentration in economics assumes an order of $q = 2$, which is relatively more sensitive to unevenness than to richness, to find equal diversity between two different assemblages requires larger changes in richness to offset smaller changes in unevenness. This will tend to under-emphasize the entrance and exit of firms relative to changes in market shares of dominant firms. However, if the same markets were analyzed using smaller orders of q, the emphasis would be reversed.

A3. Diversity in Lognormal and Pareto Distributions

Hart (1975) shows that many common measures of concentration in economics, including HHI and entropy, can be expressed in terms of moments of moment distributions. We extend his analysis to provide closed form expressions of balanced diversity and dominance diversity for the two most common distributions of market shares, log-normal and Pareto.

Following Hart (1975), assume N firms have sizes distributed according to $F(z)$, with $\gamma_r = \sum z_i^r/N$ equal to the rth moment about zero. Let $\alpha_0 = \sum z/N$ be the arithmetic mean of $F(z)$ and log $\alpha_{0g} = \sum \log z/N$ be the geometric mean of $F(z)$. The first moment distribution of $F(z)$ is defined as $F_1(z) = \frac{\int_0^z zf(z)dz}{\int_0^\infty zf(z)dz}$. The arithmetic mean of $F_1(z)$ is $\alpha_1 = \frac{\gamma_1}{\gamma_2} = \sum z^2/\sum z$ and its geometric mean is log $\alpha_{1g} = \sum z \log z/\sum z$.

Using these moment distributions, Hart (1975) shows that for any distribution, HHI is equivalent to

$$HHI = \frac{\alpha_1}{N\alpha_0} \tag{A.8}$$

and entropy is equivalent to

$$Entropy = -\log \alpha_{1g} + \log N\alpha_0. \tag{A.9}$$

Using these relations, we can derive true diversity as a function of q for any distribution.

III.A. True diversity with lognormal or Pareto-distributed firm sizes

First, assume that firm size is distributed log-normally with parameters μ and σ. Using the moments of the moment distributions, we find that diversity of order q when shares are distributed lognormally (${}^qD_L$) is

$${}^qD_L = Ne^{-\frac{\sigma^2}{2}q}. \tag{A.10}$$

The analytic form of ${}^qD_L$ are independent of the location parameter (μ). Diversity represents a scaling of the number of firms in the market N, where the shape parameters σ creates a shrinkage factor. If $\sigma = 0$ in the lognormal distribution, then $e^{-\sigma^2} = 1$ and market shares are perfectly even. As σ increases, market shares become less even and true diversity shrinks. The size of q determines the strength of the shrinkage factor, where $q = 0$ is equivalent to richness.

Hart (1975) does not derive a closed-form solution for Pareto distributed data. To derive the expression for true diversity for Pareto-distributed data, we first derive α_0, α_1, and $\log \alpha_{1g}$ for Pareto distributions with scale parameter k and shape parameter α. First, α_0 is the first moment of the distribution, which is $\alpha_0 = k\left(\frac{\alpha}{\alpha-1}\right)$. Second, the first moment distribution of the Pareto distribution is $F_1(z) = 1 - \left(\frac{k}{z}\right)^{\alpha-1}$. α_1 is the arithmetic mean of the first moment distribution, which is equivalent to $\frac{\gamma_2}{\gamma_1} = k\left(\frac{\alpha-1}{\alpha-2}\right)$. Next, $\log \alpha_{1g} = \int \log(z)dF_1 = \log k + \frac{1}{\alpha-1}$. The same relations can be used for higher moment distributions necessary for larger q.

Substituting the moments from the Pareto distribution into Equations A.8 and A.9, and converting to true diversities yields:

$$
\begin{aligned}
{}^qD_P &= N \cdot \frac{\alpha(\alpha-q)^{\frac{1}{q-1}}}{(\alpha-1)^{\frac{q}{q-1}}} \quad \text{if } q \neq 1 \text{ and } \alpha > q, \\
{}^1D_P &= N \cdot \left(\frac{\alpha}{\alpha-1}\right) e^{-\frac{1}{\alpha-1}} \quad \text{if } q = 1 \text{ and } \alpha > 1.
\end{aligned} \tag{A.11}
$$

Note that the expression for ${}^qD_P$ requires that $\alpha > q$. This is because the higher moments of the Pareto distribution do not exist unless α is larger. In particular, the variance of the Pareto distribution is only defined if $\alpha > 2$. Because HHI and ${}^2D$ are based on the second moment of the firm size distribution, they are not analytically defined in Pareto distributions with $\alpha \leq 2$. However, much of the empirical evidence on the firm size distribution estimates that α is less than two. In contrast, the expression for ${}^1D_{Pareto}$ requires that $\alpha > 1$, which is consistent with empirical literature, though some studies estimate that $\alpha = 1$ in a Zipf distribution.

Equivalent to lognormally-distributed market shares, true diversity in Pareto data represents a shrinkage of N, based on the shape parameter α. In the Pareto distribution, as α approaches infinity, the distribution approaches perfect evenness.

A.1. Substitution and Elasticity

Given the analytic solutions for true diversities in lognormal and Pareto distributions, we can calculate the marginal rate of substitution between richness (N) and the shape parameter of the distributions. These are:

$$MRS({}^qD_L) = N\sigma q$$
$$MRS({}^qD_P) = \frac{-Nq}{\alpha(\alpha-1)(\alpha-q)}$$

These represent the trade-off between N and the shape parameter. The MRS of ${}^qD$ is always increasing in q, which means that ${}^2D$ is more sensitive than ${}^1D$.

We can also calculate the elasticities of substitution between N and the shape parameters. These are the elasticities of the shape parameter with respect to N,

$$Elasticity({}^qD_L) = \frac{1}{q\sigma^2}$$
$$Elasticity({}^qD_P) = \frac{-(\alpha-1)(\alpha-q)}{q}$$

These elasticities reflect the percentage change in the shape parameter required to hold true diversity constant for a 1% increase in N (the number of firms).

A.2. *Which q is best?*

Given that the lognormal and Pareto distributions are the most likely distributions for the firm size distribution, we can use the analytic expressions for diversity in each distribution to evaluate the relative merits of different orders of q, in particular, balanced diversity ($q = 1$) and dominance diversity ($q = 2$).

First, in the Pareto distribution, true diversity is only defined analytically when the shape parameter α is greater than q. Nearly all estimates of α in empirical research on the firm size distribution are close to one. In particular, HHI and ${}^2D$ are based on the second moment of the distribution of firm sizes. Because the second moment of a Pareto distribution is undefined if $\alpha < 2$, HHI and ${}^2D$ are not mathematically valid measures of concentration if firm sizes are distributed following Pareto distributions with $\alpha < 2$, as is commonly found in the data. Though it is possible to calculate HHI using data that is drawn from a Pareto distribution with $\alpha < 2$, the estimates will be biased.

In contrast, Balanced Diversity with $q = 1$ can accommodate nearly all distributions of market shares, except when $\alpha = 1$, which is equivalent to a Zipf's law. While any true diversity with $q < 1$ will provide reliable statistics for Pareto distributions, because we are generally interested in distributions where $\alpha \geq 1$, there is no advantage from using lower values of q because $q < 1$ gives excess importance to firm counts over unevenness of market shares.

Second, in log-normally distributed data, there are no pathological diversity measures, as in Pareto-distributed data. However, $q = 1$ represents a unique value for diversity of lognormally distributed data. In particular, when $q < 1$, the elasticity of inverse variance with respect to firm counts is magnified; when $q > 1$, the elasticity is attenuated. Instead when $q = 1$, the elasticity is exactly equal to the inverse variance of the distribution. The specialness of $q = 1$ is also apparent in the marginal rate of substitution (MRS) for lognormal data. Holding richness (N) and standard deviation (σ) constant, when $q < 1$, MRS is attenuated; when $q > 1$, MRS is magnified. At $q = 1$, MRS is $N\sigma$, which means that rate of substitution between richness and evenness is the product of the standard deviation and richness.

Balanced diversity ($q = 1$) is also especially suited to log-normal data because it is based on logarithms of moments, as shown in Equation A.9. By taking the log of the first moment, the entropy-based measure ${}^1D$ transforms the log-normally distributed data into symmetric, Gaussian-distributed data. In contrast, ${}^2D$ is based on an arithmetic mean of skewed data, which provides a less valid measure of central tendency.

In summary, in data that follows either a Pareto or lognormal distribution, ${}^1D$ has superior statistical properties compared to ${}^2D$.

A4. Small-Sample Bias of ${}^{q}D$

As discussed above, ${}^{q}D$ is only defined in Pareto distributions if $q < \alpha$. However, ${}^{q}D$ can always be calculated using sample data. This means that the empirical estimates of ${}^{q}D$ when $q < \alpha$ are biased upwards. Similarly, HHI is biased downward in such samples, because the variance is infinite. Intuitively, as a distribution becomes more skewed, the likelihood of an extreme tail observation is smaller in smaller samples. This biases the estimated concentration downward, and the estimated true diversity upward. This is true for both Pareto and lognormally-distributed data.

In this section, we consider the small-sample properties of ${}^{1}D$ and ${}^{2}D$ in lognormal and Pareto distributed data. First, we consider deviations between population and sample statistics in our NETS data. Second, we consider difference between population and sample statistics in simulated data.

IV.A. Small-Sample Bias in NETS Data

Using the analytic forms of ${}^{1}D$ and ${}^{2}D$, we can match the summary statistics in Table I. In particular, we substitute q, N, and ${}^{q}D$ from Table I into Equations A.10 and A.11 and solve for the shape parameter values. Alternatively, for a given N, we find the shape parameter such that the expected diversity value across 10,000 simulated distributions of employment size matches the observed diversity values from NETS in Table I. Thus, the first method finds parameter values under the assumption that the sample data equals the population data, while the second method finds parameter values under the assumption that the sample data does not reflect the population. The closer are these values for a given q, the smaller is the small-sample bias in NETS.

Internet Appendix Table I presents the results of the calibration exercise. The parameter estimates from the first method are presented under the columns labelled 'Analytic,' and

Internet Appendix Table I
Calibration to Lognormal and Pareto Distributions

	lognormal (σ)		Pareto (α)	
Year	Analytic	Data	Analytic	Data
Panel A: Balanced Diversity ($q = 1$)				
1990	0.8609	0.8576	1.8900	1.8505
2000	1.0411	1.0386	1.6994	1.6206
2010	0.9240	0.9213	1.8143	1.7570
2020	0.9668	0.9638	1.7690	1.7041
Panel B: Dominance Diversity ($q = 2$)				
1990	1.7325	2.0510	2.0258	1.1070
2000	1.8092	2.1868	2.0195	1.0625
2010	1.8580	2.3101	2.0162	1.0204
2020	1.9174	2.4549	2.0129	0.9752

Notes: This table presents estimated shape parameters for lognormal and Pareto distributions, calibrated to observed number of firms and diversity as reported in Table I of the main paper. Analytic values are calibrated to the analytic solutions for ${}^1D$ and ${}^2D$ in Equations A.10 and A.11. Data values are calibrated to match the observed data.

the estimates from the second method are presented under the columns labelled 'Data.' The analytic values are much closer to the values inferred from the data when $q = 1$ than when $q = 2$ for both lognormal and Pareto distributions. In particular, the deviation between data-derived estimates of the shape parameters versus analytically-derived values is less than 1%, on average, compared to more than 50% for dominance diversity. This implies that balanced diversity calculated using sample data better reflects population-level diversity than does dominance diversity. As mentioned above, the limitation of dominance diversity for Pareto distributions is severe. As shown here, dominance diversity does a poor job of matching a Pareto distribution because the Pareto distribution is only defined for $\alpha > q$, though to match the observed values in the data, $\alpha < 2$.

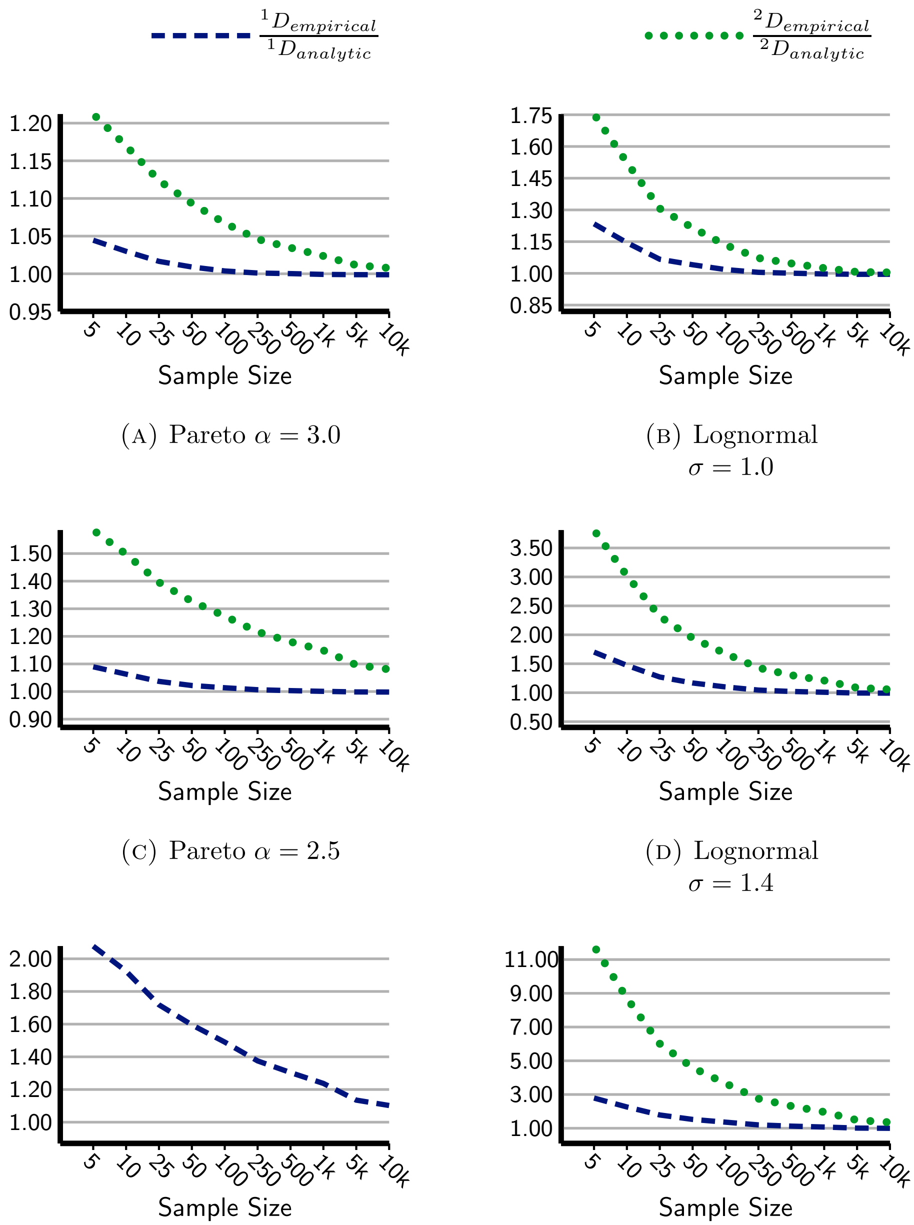
As a side note, these results also indicate that neither Pareto nor lognormal distributions are fully supported by the data. If the data were drawn from a lognormal or Pareto distribution, the shape parameters estimated from the data should be identical for $q = 1$ as $q = 2$. The results in Internet Appendix Table I suggest that when more weight is placed on the right-tail values using dominance diversity the distribution must have higher variance to match the data.

IV.B. Small-Sample Bias in Simulated Data

To investigate the small sample properties of ${}^1D$ and ${}^2D$ in Pareto and lognormal distributions, we compare their analytic values to simulated sample estimates. In particular, we calculate the mean value of ${}^1D$ and ${}^2D$ across 10,000 random samples of a given sample size N and shape parameter α or σ. We then calculate the ratio of the empirical estimate of ${}^1D$ and ${}^2D$ to their analytical values from Equations A.10 and A.11. The results are plotted in Internet Appendix Figure I.

The figures indicate that ${}^2D$ exhibits more severe small-sample bias than ${}^1D$. For example, in Pareto-distributed data with $\alpha = 3$, the empirical estimate of ${}^2D$ in a sample of 25 firms is 12.3% larger than the true analytical value. In contrast, the empirical estimate of ${}^1D$ is 1.6% larger. As α gets smaller, and the distribution becomes more skewed, the small-sample biases increase. When $\alpha = 2.5$, in a sample of 25 firms, the empirical estimate of ${}^2D$ is 39.6% larger than the analytical value, while ${}^1D$ is 3.7% larger. Even in a sample of 5,000 firms, ${}^2D$ is 9.6% above the analytical value when $\alpha = 2.5$. When $\alpha < 2$, the analytical value of ${}^2D$ is undefined. As α approaches 2, the ratio of the empirical estimate of ${}^2D$ to the analytical value goes to infinity. For $\alpha = 1.5$, ${}^1D$ is defined but biased in small samples. In a sample of 250 firms, the empirical estimate ${}^1D$ is 37.5% larger than its analytical value.

The small sample biases are larger in lognormally-distributed data, though the biases are larger for ${}^2D$ than ${}^1D$. For example, when $\sigma = 1.0$ in a sample of 100 firms, the empirical estimate of ${}^1D$ is 1.8% larger than the analytical value compared to 13.3% for ${}^2D$. As the distribution becomes more skewed, the biases increase. When $\sigma = 1.8$, in a sample of 100 firms, the empirical estimate of ${}^{1}D$ is 35.9% larger than the analytical value, whereas the empirical estimate of ${}^{2}D$ is 265% larger than the analytical value.



(A) Pareto $\alpha = 3.0$

(B) Lognormal $\sigma = 1.0$

(C) Pareto $\alpha = 2.5$

(D) Lognormal $\sigma = 1.4$

(E) Pareto $\alpha = 1.5$

(F) Lognormal $\sigma = 1.8$

Internet Appendix Figure I
Small Sample Bias of Diversity Estimates

This figure presents the ratio of simulated diversity estimates to their analytical values for ^1D (blue dashed lines) and ^2D (green dotted lines) in data that follows either a Pareto distribution (subpanels A, C, and E) or a lognormal distribution (subpanels B, D, and F). Simulated observations are the mean values of ^1D and ^2D across 10,000 samples, each with size N and the distribution parameter indicated in the subpanel. For Pareto distributions with $\alpha < 2$, ^2D is infinite.

A5. Using Analytic Formulas to Estimate Diversity from 1963 to 1992

Given the analytic formulations of diversity in Equations 5 and 6 in the main paper, it is possible to estimate diversity if the total number of firms N is known and if we have parameter estimates of σ assuming lognormally-distributed firm sizes or α assuming Pareto-distributed firm sizes. For example, given estimates of HHI and N, we could back out the distribution parameters, then calculate diversity for any value of q under each firm-size distribution assumption. This analysis depends on the validity of the distributional assumptions, but if these assumptions are reasonably close to reality, then the analytic formula provide a reasonable estimate of diversity. Unfortunately, it is rare to observe all necessary parameters in public data. For example, though the U.S. Economic Census reports the number of firms per industry, it only reports HHI for the top 50 firms in the industry. To estimate the parameters, we require HHI for the entire population of firms, not just the top 50.

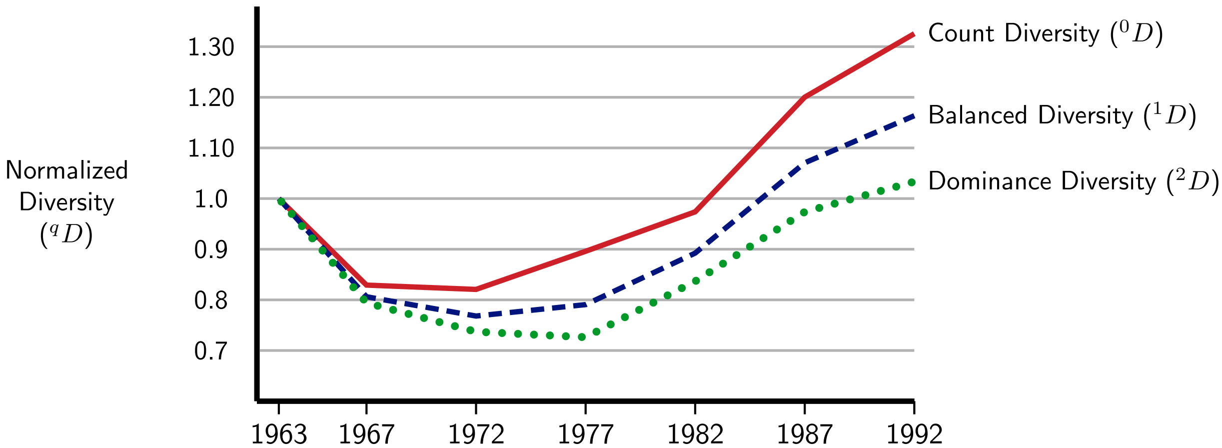
To address this limitation, we estimate σ and α using the 4-, 8-, 20-, and 50-firm concentration ratios reported in the Economic Censuses. These concentration ratios represent estimates of the probability mass in the upper tail of the firm-size distribution. Under the assumption of lognormal or Pareto distributions, we can estimate the entire distribution using these summary statistics. Because concentration ratios are based on rank orderings, we cannot directly identify the sales value that corresponds to the probability mass from the distributions CDF. Instead, we simulate the expected concentration ratios for a grid of paramater pairs (N,σ) and (N,α) in a Monte Carlo simulation. For a given value of N in the Census reports, we find the parameter value that produces an estimated concentration ratio from the Monte Carlo simulations nearest the concentration ratio reported by the Census. In validation tests, we find that the most accurate match for the lognormal distribution is the average value of σ across the four concentration ratios. For Pareto, it is the 50-firm concentration ratio.

We find that the estimates of σ assuming lognormal distributions are better behaved than the α estimates from Pareto estimates. The Pareto estimates vary considerably and many of the estimates are less than one, which indicates a more extreme distribution than a power law and we cannot estimate dominance diversity or balanced diversity for $\alpha < 1$. Therefore, we present results using the lognormal estimates.

An advantage of this approach is that we can exploit the Census data back to the 1960s because the Census reports numbers of firms and concentration ratios in the Economic Census. We use the data generously provided by Jan Keil on his website that tabulates the Census reports as used in Keil (2017). The Census only reports enough information to estimate sales diversity at the 4-digit SIC code-level for the manufacturing sector. The number of industries prior to 1963 are sparse, and in 1997 the Census switched to NAICS codes which are difficult to harmonize between 1992 and 1997, so we limit our analysis to 1963 to 1992.

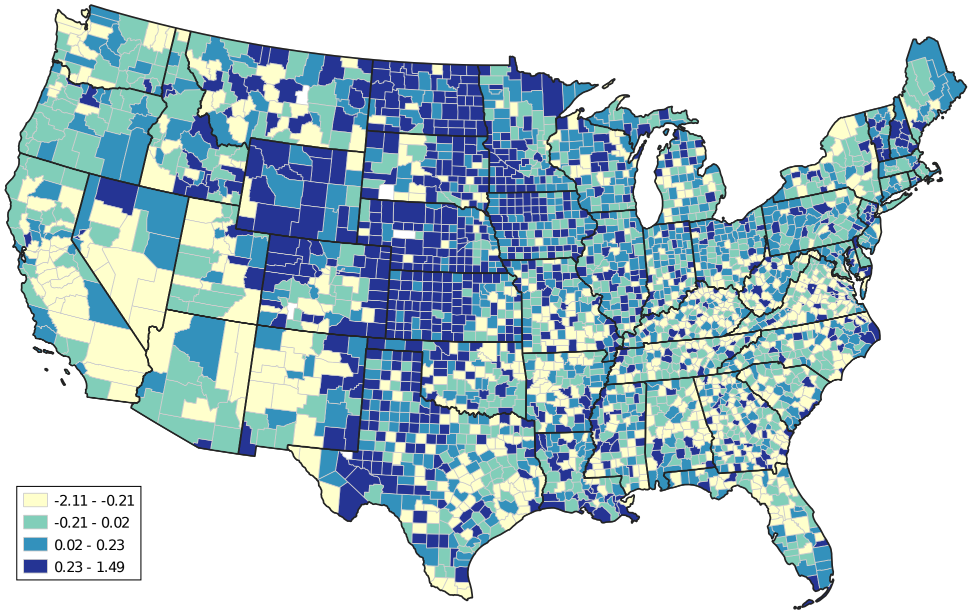
Internet Appendix Figure II reports the results of the simulation. We find that diversity for each diversity order q declined from 1963 to the mid 1970s, but then increased through 1992. Dominance diversity was roughly equivalent in 1992 as it was in 1963. The increase in diversity during the 1970s and 1980s may reflect the general trend towards deregulation. The figure also reveals that the three diversity measures comove closely over time. This indicates that the driving force is the number of firms in the sector, not the evenness of their market shares. Indeed, the estimate of σ is highly persistent over the sample period.

This analysis illustrates the benefit of the analytical formulation of diversity under an assumed distribution of firm sizes. If a distributional assumption is reasonable, to calculate the full set of diversity measures only requires an estimate of the shape parameter and the number of firms in the market.

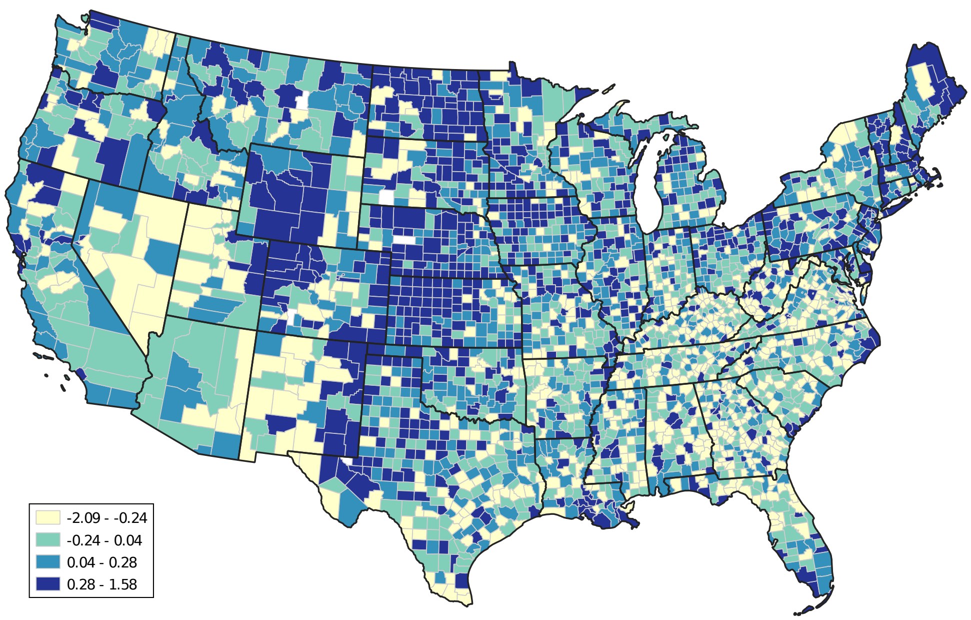


Internet Appendix Figure II
Sales Diversity in the Manufacturing Sector Assuming Lognormally Distributed Market Shares

This figure presents the time-series of sales diversity at the SIC 4-digit nationwide industry level for Manufacturing industries, from 1963 to 1992, normalized by 1963 values, for diversity order $q = 0$, 1, and 2. Original data are from the Economic Census of the United States, following Keil (2017), generously provided on Jan Keil's website. Diversity is estimated using Equation 6, where σ is estimated from the best fit of a lognormal distribution to the 4-, 8-, 20-, and 50-firm concentration ratios provided by the Census.

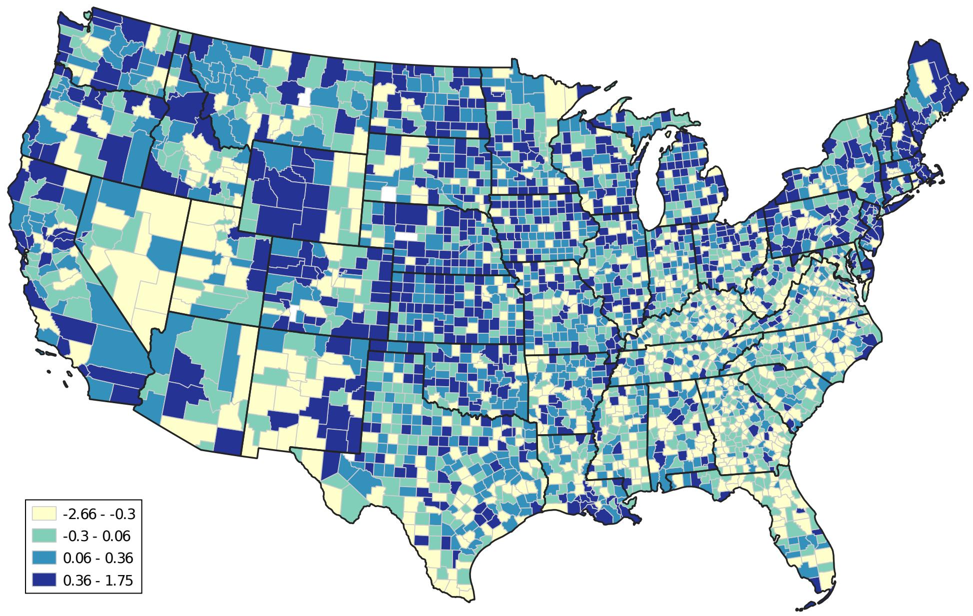


(a) Count Diversity



(b) Balanced Diversity

Internet Appendix Figure III

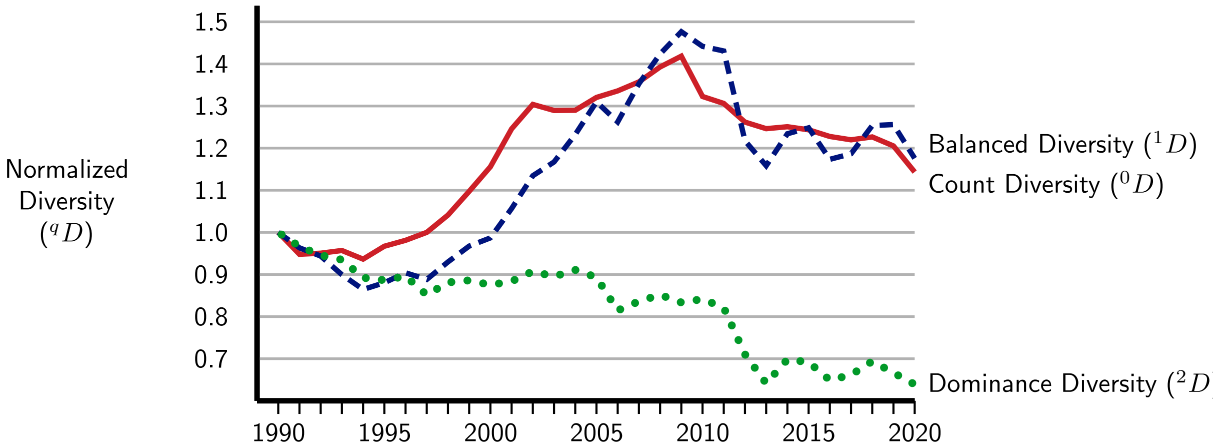


(c) Dominance Diversity

Internet Appendix Figure III
Abnormal Diversity per Capita in the U.S. in 2020

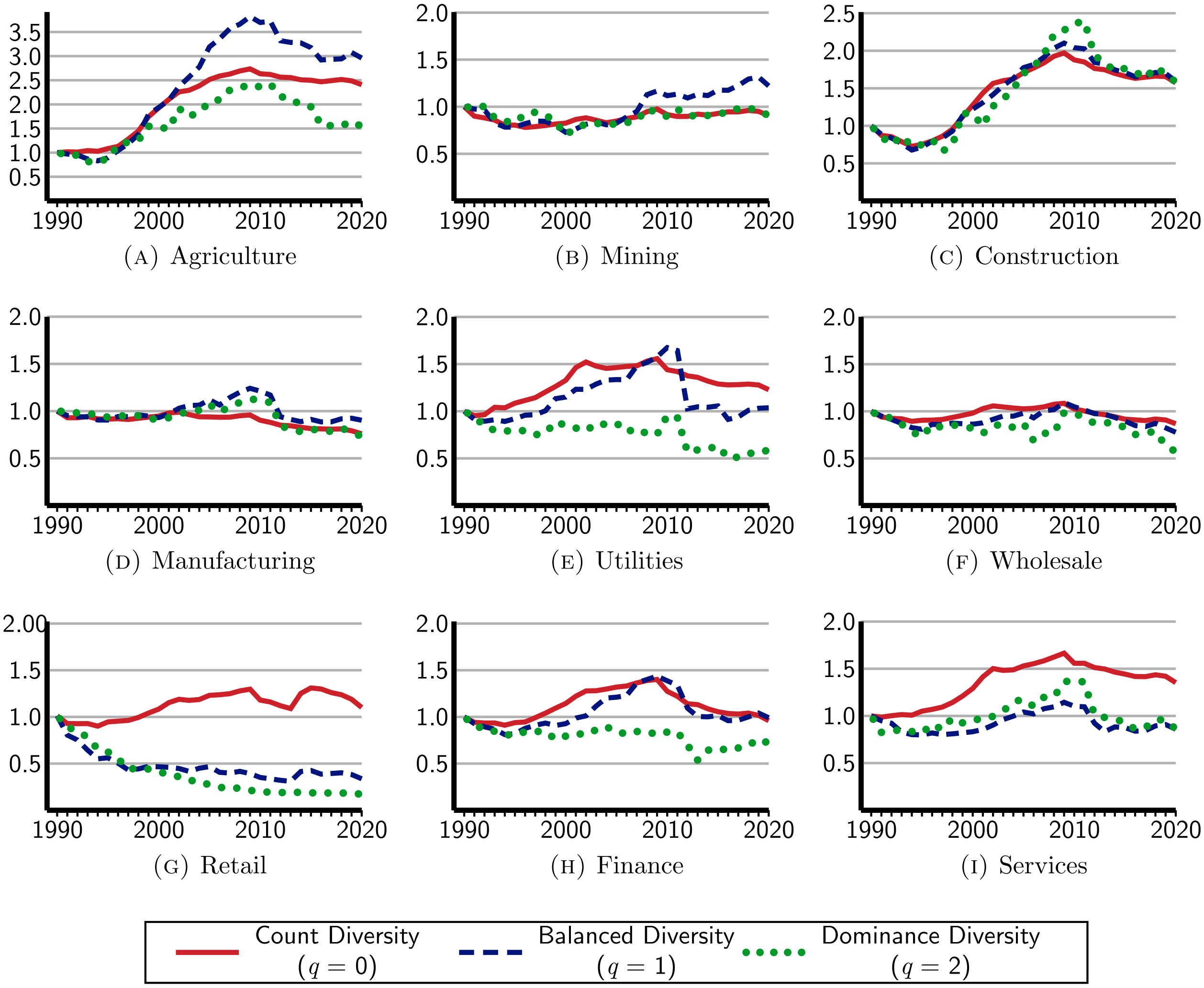
This figure presents abnormal diversity for $q = 0$ (Panel A), $q = 1$ (Panel B), and $q = 2$ at the county level in 2020. Abnormal diversity is the county-level residual in a regression of $\log(1 +^q D)$ on a constant and log(1+population). Darker values represent higher abnormal diversity per capita. Categories of abnormal diversity are based on quartile values across all counties in the US. Data are from NETS and include firms with at least 20 employees.

A6. Results Using Non-Imputed Data



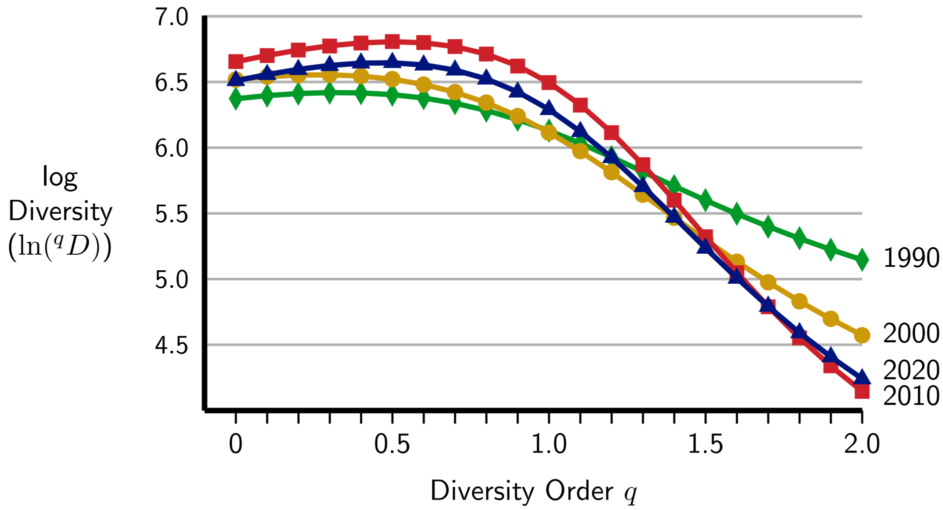
Internet Appendix Figure IV
Weighted Average Industry Diversity: Non-Imputed Data

This figure presents the time-series of employment diversity at the SIC 4-digit nationwide industry level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Data are from NETS and include firms with at least 20 employees and only include non-imputed data.



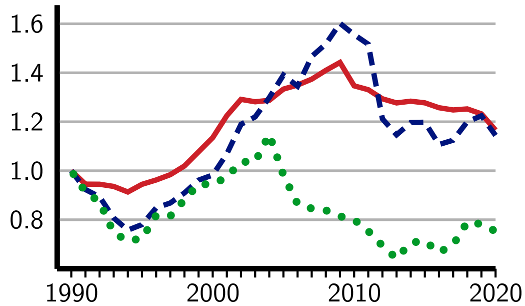
Internet Appendix Figure V
Diversity Decomposition: Within and Between Industries: Non-Imputed Data

This figure presents the time-series of alpha, beta, and gamma diversity of employment at the SIC 4-digit nationwide industry level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within an industry (within diversity). Beta diversity reflects the number of effective industries (between diversity). Gamma diversity reflects the total diversity as the product of alpha and beta diversity. Data are from NETS and include firms with at least 20 employees and only include non-imputed data.

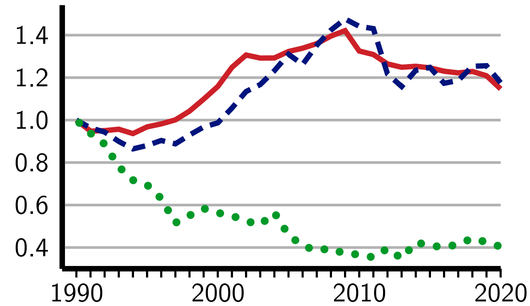


Internet Appendix Figure VI
Employment Diversity Profiles: Non-Imputed Data

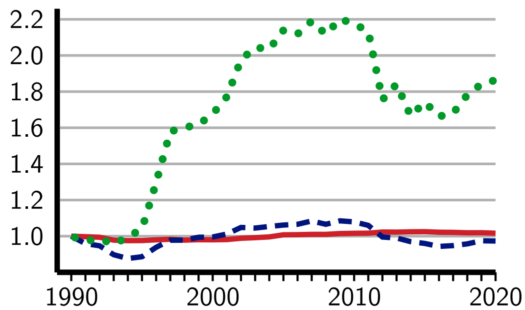
This figure presents the profile of diversity for industry-level employment for years 1990, 2000, 2010, and 2020. Diversity reflects the number of effective firms within an industry. The diversity order q controls how much weight is given to unevenness versus richness. Data are from NETS and include firms with at least 20 employees and only include non-imputed data.



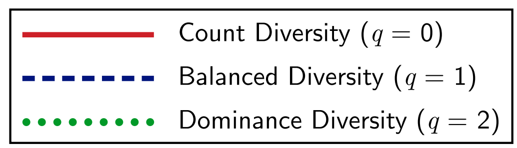
(A) Total Diversity (${}^{q}D_{\gamma}$)



(B) Effective Firms in Average Industry (${}^{q}D_{\overline{\alpha}}$)

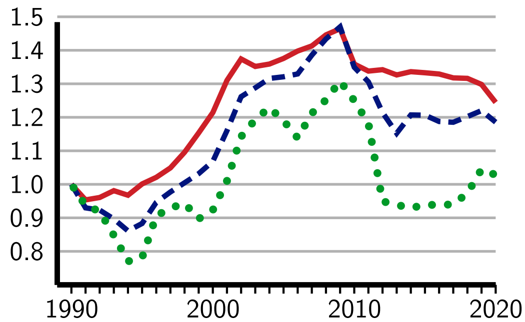


(C) Effective Industries (${}^{q}D_{\beta}$)

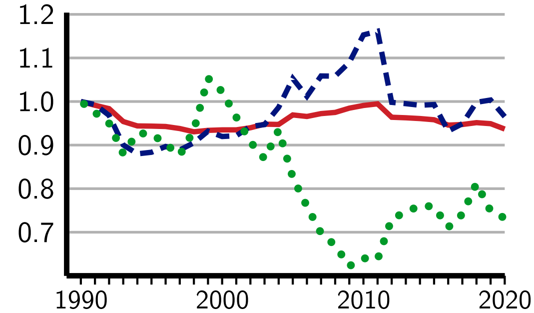


Internet Appendix Figure VII
Diversity Decomposition: Within and Between Industries:
Non-Imputed Data

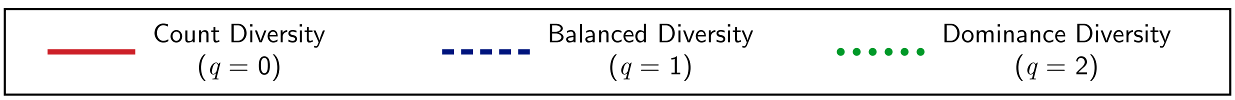
This figure presents the time-series of alpha, beta, and gamma diversity of employment at the SIC 4-digit nationwide industry level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within an industry (within diversity). Beta diversity reflects the number of effective industries (between diversity). Gamma diversity reflects the total diversity as the product of alpha and beta diversity. Data are from NETS and include firms with at least 20 employees and only include non-imputed data.



(A) Effective Firms in the Average County (${}^{q}D_{\overline{\alpha}}$)

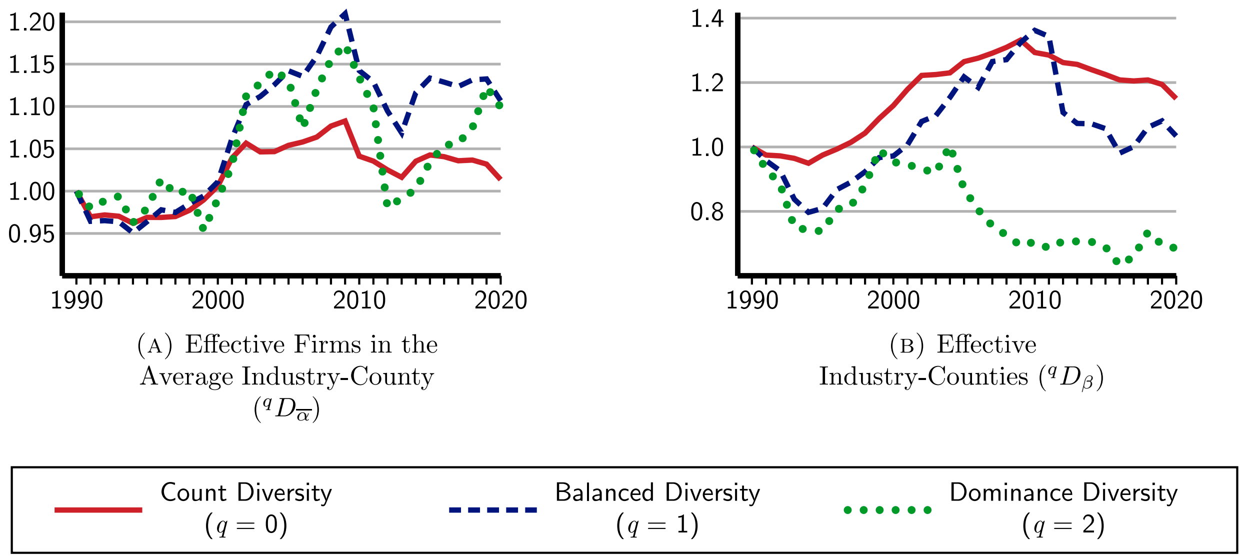


(B) Effective Counties (${}^{q}D_{\beta}$)



Internet Appendix Figure VIII
Diversity Decomposition: Within and Between Counties: Non-Imputed Data

This figure presents the time-series of alpha and beta diversity of employment at the county level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within a county. Beta diversity reflects the number of effective counties. Data are from NETS and include firms with at least 20 employees and only includes non-imputed data.

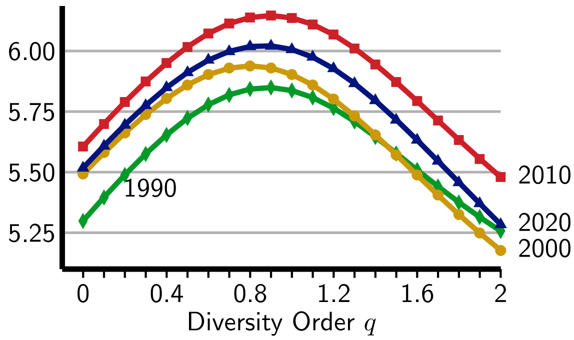


(A) Effective Firms in the Average Industry-County (${}^{q}D_{\overline{\alpha}}$)

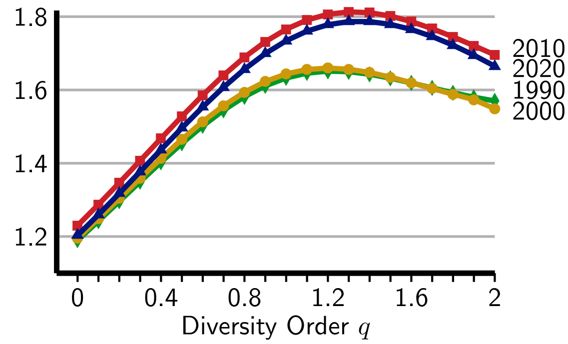
(B) Effective Industry-Counties (${}^{q}D_{\beta}$)

Internet Appendix Figure IX
Diversity Decomposition: Within and Between Industry-Counties: Non-Imputed Data

This figure presents the time-series of alpha and beta diversity of employment at the industry-county level, from 1990 to 2020, normalized by 1990 values, for diversity order $q = 0$, 1, and 2. Alpha diversity reflects the number of effective firms within an industry-county. Beta diversity reflects the number of effective industry-counties. Data are from NETS and include firms with at least 20 employees and only include non-imputed data.



(A) County Level



(B) Industry-County Level

Internet Appendix Figure X
Diversity Profiles at County and Industry-County Levels:
Non-Imputed Data

This figure presents the profile of log alpha diversity at the county level (Panel A) and industry-county level (Panel B) for years 1990, 2000, 2010, and 2020. Alpha diversity reflects the number of effective firms within a county or industry-county. The diversity order q controls how much weight is given to unevenness versus richness. Data are from NETS and include firms with at least 20 employees and only include non-imputed data.

References

Hart, P.E., "Moment Distributions in Economics: An Exposition," *Journal of the Royal Statistical Society. Series A (General),* 138 (1975), 423–434.

Jost, Lou, "The Relation Between Evenness and Diversity," *Diversity,* 2 (2010), 207–232.

Keil, Jan, "The trouble with approximating industry concentration from Compustat," *Journal of Corporate Finance,* 45 (2017), 467–479.