A Appendices

A.1 Data appendices

A.1.1 Data compilation workflow

Figure 2 summarizes the steps to create the data using artificial example data. Note that not all columns in the data are presented in the figure. I begin with the set of all 24,781,046 non-commodity shipments from PIERS from 1990 to 2015. I then refine the sample to only shipments of automation capital using the algorithm described above. There are 919,131 such shipments. I then match those shipment-level records to the NETS manufacturing sample, which allows me to associate multiple shipments with noisy recipient business names and addresses to individual establishments with unique D&B identifiers. The process for this match is described in Appendix A.1.3.

Next I collapse the data to the establishment-by-year level, summing the value of automation capital shipments to get that year's flow of automation capital investment. Then I merge on the establishment-by-year data from NETS to include sales values; I add rows where an establishment was in business, i.e., present in NETS, but had no automation capital shipments. There are 949,410 unique establishments in the data representing 8,555,627 establishment-years.

A.1.2 Included industries

A.1.3 Description of matching procedure

The greatest challenge for building these data is a process called "entity resolution." The PIERS data are not matched to entities, and variations in spelling and shortening of names or addresses can make it hard to tell whether two observed shipment recipient names correspond to the same latent entity. Matching those latent entities to NETS establishments is a second challenge.

To resolve entities, I use a machine-learning algorithm from a class of algorithms introduced by Fellegi and Sunter (1969) and much improved upon since then. These algorithms are supervised-learning techniques

Table A1: Included Industries (NAICS 6-digit)

Included NAICS 6-digit	Unique establishments	Total automation capital investment ($ k)	Unique markets	Included NAICS 6-digit	Unique establishments	Total automation capital investment ($ k)	Unique markets
311211	722	207414.7	158	332510	4271	38968.42	298
311511	1414	5719.45	235	332618	4506	205000	356
311612	2592	13940.25	293	332710	60375	541354.1	867
311811	3825	1511.14	367	332721	2656	18184.81	209
311812	11245	93540.77	582	332722	2202	58567.99	201
311991	753	2447.34	123	332811	1360	15583.65	144
311999	5432	30734.26	367	332812	6408	41750.89	407
312111	4106	48702.64	436	332813	13552	172426.5	506
321113	5604	10444.46	1206	332911	1552	139051.1	176
321114	1248	17527.67	234	332999	9861	42273.99	525
321214	963	757.67	185	333511	4522	36637.25	274
321911	26378	30639.6	1389	333515	4565	826988.6	330
321912	1584	3653.84	342	333911	2492	158622.3	259
321918	6833	20891.38	555	333994	1342	6072.91	163
321920	8723	8760.93	802	333999	10537	386508	506
321992	3360	1402.78	519	334220	10474	36577.63	442
321999	15811	15407.59	1126	334290	3810	4221.19	259
322211	3165	170754.2	431	334412	4462	61460.87	225
322212	413	6417.24	77	334416	1437	20688.73	165
322299	2327	68433.98	206	334418	384	2900.15	53
323113	19238	36671.31	942	334419	9624	123342.6	377
323117	1376	19106.69	153	334513	4788	126607.4	296
324121	3015	3429.71	363	334515	5639	45487.01	279
324122	1053	1885.68	137	335313	2180	132277.5	202
325199	5318	23247.56	375	335314	4614	84726.51	305
325314	847	47.62	182	335929	466	28677.36	94
325411	2413	7329.72	214	335931	1594	35306.52	174
325510	5290	14735.8	328	335999	14565	120418.1	575
325520	3100	22336.18	226	336211	2856	36560.89	337
325910	1739	13612.13	151	336360	1602	18558.76	202
325991	305	2614.02	71	336370	1595	212679.5	159
326111	1230	59574.13	137	336413	4553	424297.8	280
326113	1907	78594.01	198	337110	24507	65169.18	1244
326122	617	15114.14	140	337127	6729	16760.85	417
326140	1479	21630.78	182	337215	8297	14997.85	531
326150	1527	49570.53	193	337910	2079	16282.73	223
326160	453	38044.27	93	337920	4398	2801.75	260
326191	703	20655.94	123	339112	9409	39552.18	361
326199	24349	955908.6	1078	339113	13004	41213.12	593
326299	2811	80980.76	266	339920	23844	62575.34	1181
327215	7392	61781.39	447	339950	67716	17287.5	875
327310	1143	39234.52	162	339991	1350	24962.98	157
327320	12221	23423.54	824	339999	96370	544249.5	2457
327331	2001	19967.66	441	511110	31331	47738.75	883
327332	255	10982.52	78	511120	33734	389029.9	818
327390	11594	78859.85	767	511140	4851	3782.14	332
327420	4210	1690.96	282	511199	72716	112297.9	1485
327991	5094	30418.82	413	517110	17797	8031.24	820
327999	852	10419.65	119	517210	11230	233.01	747
331222	261	36.26	57	541219	979	2710.95	208
332111	2028	117837.1	232	541310	729	962.21	150
332311	3910	20617.72	470	541380	2613	3894.1	402
332312	15071	88004.4	964	541620	50	165.21	20
332313	2764	68255	284	541712	5508	22318.43	303
332321	4721	10250.66	352	541860	1104	15882.01	135
332322	13093	117221.5	750	541910	1048	448.84	107
332323	7740	12261.42	398	561450	381	2154.34	73

Example PIERS shipments

Recipient	Address	Year	Commodity	Tariff value
The Widget Manufacturing Company	1 Widget Circle	2012	Automatic stamper	85000
Widget Manuf	1 Wdiget	2012	Molding machine	42000
Widget M Co	1 Wdiget Cr	2013	robot laser cutter	5800
Industrial Widget Partners LLC	456 Highway 66 Business	1999	CNC mach	30000
Ind wid part	#456 Hwy 66	2002	Face masks	600
Wdiget's Burgers and Fries	123 Main St	2004	Plastic cups	350
Widgets Burgers	123 Main	2007	Toy robots	1250

Example PIERS shipments refined to automation capital

Recipient	Address	Year	Commodity	Tariff value
The Widget Manufacturing Company	1 Widget Circle	2012	Automatic stamper	85000
Widget Manuf	1 Wdiget	2012	Molding machine	42000
Widget M Co	1 Wdiget Cr	2013	robot laser cutter	5800
Industrial Widget Partners LLC	456 Highway 66 Business	1999	CNC mach	30000

NETS establishment data

Company name	Address	ID number
Widget Manufacturing Company	1 Widget Circle	12345
Industrial Widget Partners	456 Highway 66 Business	6789

PIERS-based measure of automation capital investment

ID number	Year	Automation capital investment value
12345	2012	127000
12345	2013	5800
6789	1999	30000

NETS establishment-by-year data

ID number	Year	Sales
12345	2012	12000000
12345	2013	13000000
6789	1999	8000
6789	2000	10000

Final data

ID number	Year	Automation capital investment value	Sales
12345	2012	127000	12000000
12345	2013	5800	13000000
6789	1999	30000	8000
6789	2000	0	10000

Figure 2: Workflow for data compilation

in which one first constructs the Cartesian product of records, yielding data made up of the full set of candidate matches. One then calculates distance metrics between the candidates for each pair. A sample of the candidates are scored as matches or non-matches, and an empirical model is estimated of the likelihood of a match as a function of the distance between the observations.[28] Those weights are used to classify the distance measures in the unscored data by estimating the probability of a match. For shipments with a positive probability of match to multiple NETS records, the maximum likelihood record is selected. Finally, a confidence cutoff is established, and those above the cutoff are classified as matches.

In my implementation, I began by hand-correcting city and state names to match valid geographic identifiers from US TIGER data. This allows blocking the Cartesian product of records on city and state, without which the project would be computationally infeasible. Then, *firm name* from the NETS data, *recipient name* from the PIERS data, and addresses from both are used to compute two distances: string distance between

[28] Steorts et al. (2019) provide an excellent overview of recent advances in techniques to improve the computability of matches in large data samples.

names and string distance between the street components of the addresses. The particular string distance is bigram distance with simple weight penalties assigned to bigrams based on their baseline prevalence in the block.[29] For as much flexibility as possible, similarity scores were each divided into 10 buckets, each representing 10 percent (i.e., scores from .1 to .2). Let $name\_\widehat{sim}\_score$ represent a vector of indicator variables for each bucket of the name and address similarity scores and similarly define $street\_\widehat{sim}\_score$. The empirical model used to create a univariate probability a pair i is a match is as follows:

$$\begin{aligned} I[Match_i] = \alpha &+ \hat{\beta}_1(name\_\hat{sim}\_score) + \hat{\beta}_2(street\_\widehat{sim}\_score) \\ &+ \hat{\beta}_3(name\_\widehat{sim}\_score \times street\_\widehat{sim}\_score) + \epsilon \end{aligned} \tag{1}$$

The cutoff for the main results was set at a probability of .7, though results are robust to a wider range of cutoffs.

A.1.4 Geographic distribution of measured automation expenditures

Figure 3 presents maps showing the distribution of NETS manufacturing establishments, NETS measured sales, establishment size in terms of employment, and measured automation capital expenditure. The fact that the measured investment in automation capital is centered in areas traditionally known to be US manufacturing centers adds confidence in the data.

A.1.5 Categories of NETS firms removed before matching

Figure 4 presents the regular expressions used to remove trading companies, distributors, and shippers who may not be using the purchases themselves. Universities are also removed because manufacturing production is not central to their business model.

[29] String distance is computed as $\frac{1}{(min[s1,s2])}\sum_{m\in M}\frac{1}{f_m}$, where M is the set of shared grams between the two strings, s_i is the number of grams in the relevant string, and f_m is the empirical frequency of the focal gram.

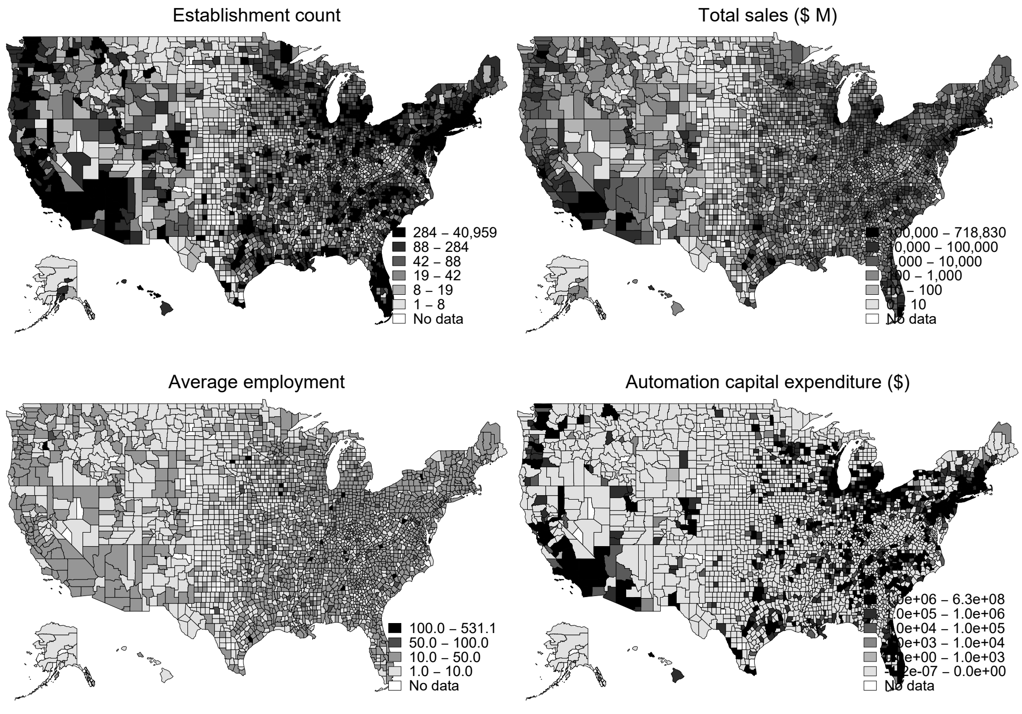

Figure 3: Geographic distribution of automation capital expenditure

A.1.6 Papers assessing the match between NETS and external data sources

Below I summarize primary findings from the comparison of NETS to administrative sources for studying business dynamics and entrepreneurship.

Business dynamics Barnatchez, Crane and Decker (2017) and Crane and Decker (2020) compare the NETS to Census data for the purposes of studying business dynamics.

These two studies are primarily interested in rates of foundings and closures and seeing if results from NETS correspond to findings from administrative data. A primary consideration of the two studies is the count of establishments. Crane and Decker (2020) note that one source of deviation between the Longitudinal Business Dynamics data series (LBD) and NETS is that the LBD is intended to capture only employer establishments. They also note that NETS marketing material suggests that it would also capture non-employer

```
drop if regexm(name,"DI?ST(RIBUT)?(ION|ER|OR)?S?")
drop if regexm(name,"IMPO(RT)?(ER)?S?")
drop if regexm(name,"EXPO(RT)?(ER)?S?")
drop if regexm(name,"(AIR )?EXP(RESS)?(\sLINE)?R?S?")
drop if regexm(name,"FO?R?WA?R?D(I?N?G|E?RS?)?")
drop if regexm(name,"LOGIST?I?C?S?")
drop if regexm(name,"INTE?R?CONTINENTAL")
drop if regexm(name,"TRA?NSP?O?R?T?A?T?I?O?N?")
drop if regexm(name,"SH[I]?P(PIN)?G( LINE)?[S]?")
drop if regexm(name,"CO?NTA?I?NE?R( ?LINE)?S?")
drop if regexm(name,"(AIR )?FRE?I?G?H?T\s?(LINE)?R?S?")
drop if regexm(name,"TRA?DI?N?G")
drop if regexm(name,"CU?STO?MS?( BRO?KE?RS?)?")
drop if regexm(name,"BRO?KE?R(S|A?GE)?")
drop if regexm(name,"CONSOLIDAT(O|E)RS?")
drop if regexm(name,"TRA?DI?N?G?")
drop if regexm(name,"EXPEDIT(O|E)RS?")
drop if regexm(name,"CA?RGO")
drop if regexm(name,"(INTER)?MARINE")
drop if regexm(name,"((AIR ?(AND|&)? ?SEA)|(SEA ?(AND|&)? ?AIR))")
drop if regexm(name,"INTERMODAL")
drop if regexm(name,"MARITIME")
drop if regexm(name,"AD[UA]*NA(S|AL)?E?S?")
drop if ustrregexm(name,"UNIV(\b|E?RSI?TY)?")
drop if strmatch(name,"PANALPINA")
drop if regexm(name,"OVERSEAS")
drop if regexm(name,"OCEAN SERVICE")
drop if regexm(name,"TRUCKING")
```

Figure 4: Regular expressions used to remove NETS firms before matching

establishments. To compare the establishment counts between NETS and the LBD, they construct a synthetic panel that they believe would represent the NETS criteria from the Census County Business Patterns (CBP) data series. Crane et al. find that the count of NETS firms falls between the Census employer establishment count and their count from their synthetic panel. This variation suggests that there is noise in the count of establishments in the NETS data. I address that possibility in three ways.

Barnatchez, Crane and Decker (2017) and Crane and Decker (2020) provide some suggestions for how to transform the NETS data to improve the match of establishment counts to the CBP data. Barnatchez, Crane and Decker (2017) propose aggregating D&B establishments within a geography. However, their recommended aggregation often aggregates establishments in different industries.[30] Given the importance of industry classification to the phenomenon of interest here, the results presented follow the spirit of the suggestion and aggregate establishments with the same name, city, and state.

Crane and Decker (2020) suggest that the counts are actually biased toward an erroneous presence of too many small firms. Erroneous NETS establishments would not actually receive shipments, suggesting that the measured proportion of establishments receiving automation capital is biased downward. Furthermore, because those records would never be matched to shipments, they could not be marked as investing in response to competition. Real establishments in the same market would not be induced to invest by these establishments erroneously included in NETS. This implies a bias toward zero in the present results.

Finally, Crane and Decker (2020) recommend validating effects using a sample in which the population is not defined through NETS. They give the specific example of Compustat, which encompasses the universe of publicly traded firms. As such, I conduct an additional validation exercise correlating both revenue and capital expenditure measures for a sample of NETS headquarters matched to Compustat data in Section A.4.1.

Entrepreneurship Echeverri-Carroll and Feldman (2019) manually match NETS records to state-level Secretary of State (SOS) data (compiled by Guzman and Stern, 2016) on startups from the Austin, TX, and the North Carolina Research Triangle Park (RTP) regions of the US. They focus on corroborating founding years for the purpose of counting firm foundings by year. They find mixed results and suggest reasons for

[30] Barnatchez, Crane and Decker (2017, section 3.6) provide a particular list of industries in which NETS' employment counts deviate particularly strongly from administrative data, though these are all outside of the manufacturing industries studied here.

the deviations. Note that the focus of the present paper is not entry years and that year-level observations of firms that were technically incorporated but not operating or making capital expenditures would bias the results toward zero.

A.2 Additional analysis details

A.3 Detail of main estimation equation

I sum the reported value of the shipments classified as automation capital investments received by establishment i in year t. The $Automation\ capital\ stock_{it}$ variable is the discounted sum of the values of automation capital investments to date.

I estimate Equation 3 separately for each market k. If both coefficients, $\hat{\beta}$ and $\hat{\gamma}$, are greater than zero, then the $Demand\ creating_k$ indicator variable is set to one.

$$Market\ share_{it} = \alpha + \beta_1 Automation\ capital\ stock_{it} + \beta_2 \widehat{Demand\ creating}_k + \beta_3 Automation\ capital\ stock \times Demand\ Creating_{it} + \zeta_i + \eta_t + \nu_{it} \quad (4)$$

Because the $\widehat{Demand\ creating}_k$ indicator is estimated, there is noise in the estimate. That noise needs to be accounted for in calculating the standard errors from Equation 4. To do that, I calculate ν_{it} with 100 replications of a bootstrap routine that includes both: (a) the set of all of the market-level estimates of Equation 3 and (b) Equation 4.

Figure 5: Workflow for analysis for testing for Strong Increasing Dominance (SID)

A.3.1 First stage results from Section 4.1 IV analyses

Table A2 presents the results of the first stage estimates for the analyses in Table 5.

Figure 6 shows the weight of annual shipments of microwave ovens (HS: 851650) to the Port of Long Beach from the Republic of Korea for our example in Section 4.1.

Table A2: First stage results from Section 4.1 IV analyses

VARIABLES	(1) Stock	(2) Stock × DC
Non-automation shipments	-0.000001***	-0.000001***
	(0.000)	(0.000)
– × DC market [dummy]	-0.000000	0.000001***
	(0.000)	(0.000)
R^2	0.612015***	0.537274***
Observations	3,367,207	3,367,207
Constant	0.0104***	0.00391***
	(0.000)	(0.000)
Market classification	Elasticity	Elasticity
Model	OLS	OLS
Establishment FE	Yes	Yes
Year FE	Yes	Yes
Standard errors	Bootstrap	Bootstrap

A.4 Additional data validation exercises

A.4.1 Validating automation capital estimates by comparing to Compustat

As a confirmatory check that the PIERS-based measure captures some automation capital investment, I compare it to self-reported accounting measures for public firms in Compustat. I first take a random sample of 300 firms present in Compustat during the time period that were listed on American exchanges and classified as being in manufacturing industries (NAICS 31–33). Of those, I found 200 with corresponding DUNS numbers listed in the Mergent Intellect Database. I summed all establishment-year PIERS records to the matching headquarters DUNS number and compared my PIERS-based automation capital flow and stock values with the two accounting line items that would represent automation capital: “Capital Expenditure" and “Property, Plant, and Equipment—Machinery and Equipment at Cost." Table A3 shows raw correlations between the accounting measures and the PIERS-based measure.

Table A3: Raw correlations between PIERS-based measure and Compustat accounting measures

	[1]	[2]	[3]	[4]
[1]Automation capital (flow) ($M)	1			
[2]Automation capital (Stock) ($M)	0.660***	1		
[3]Property, Plant, and Equipment—Machinery and Equipment at Cost ($M)	0.157***	0.301***	1	
[4]Capital expenditure ($M)	0.139***	0.248***	0.877***	1
Observations	609			

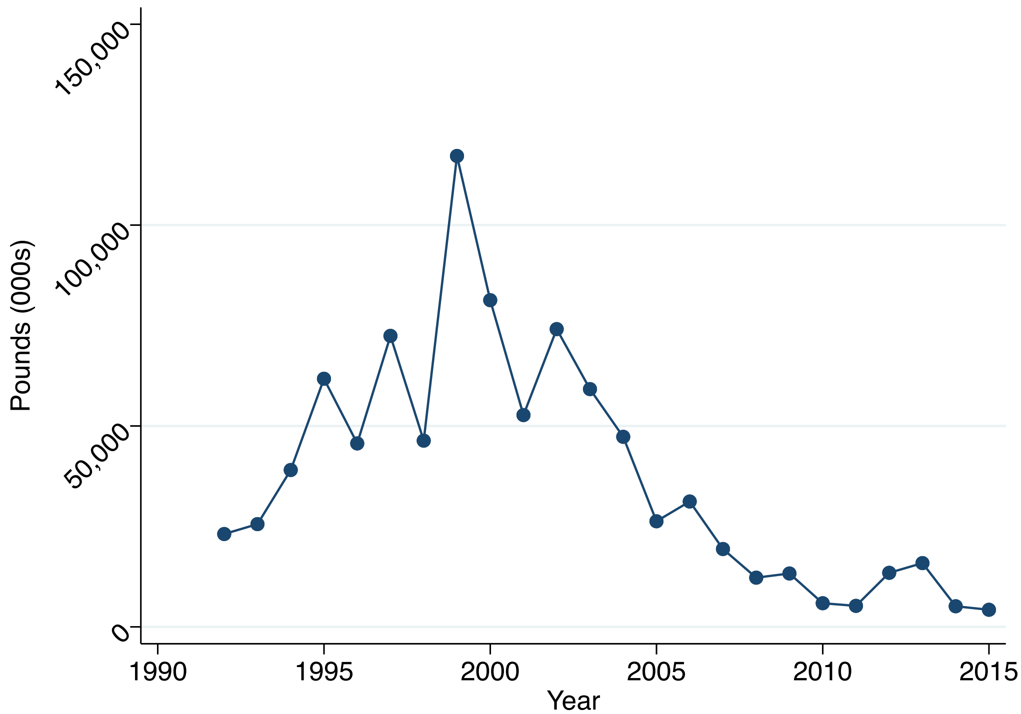

Figure 6: Weight of shipments of microwave ovens (HS: 851650) from the Republic of Korea through the Port of Long Beach by year

In Table A3, the correlations indicate that one dollar of additional capital expenditure reported in financial statements is associated with 13.9 cents of automation capital spending captured in my data. Capital expenditure and PP&E include investments that would not be included in automation capital—for example, non-automation machinery like wind tunnels and air-filtration devices. Given that, even perfect automation capital measures would not perfectly correlate with these values. A statistically significant correlation suggests that automation capital is sufficiently well matched to individual firms that we can reject the null that they are randomly allocated. Although this is not dispositive, I interpret the significant associations on such limited data as a small piece of support for the PIERS-based measure, including at the firm level. Figure 7 presents the results of Lowess regressions corresponding to the correlations from Table A3, showing that the positive association is present for the entire range of the accounting values from the Compustat matched sample.

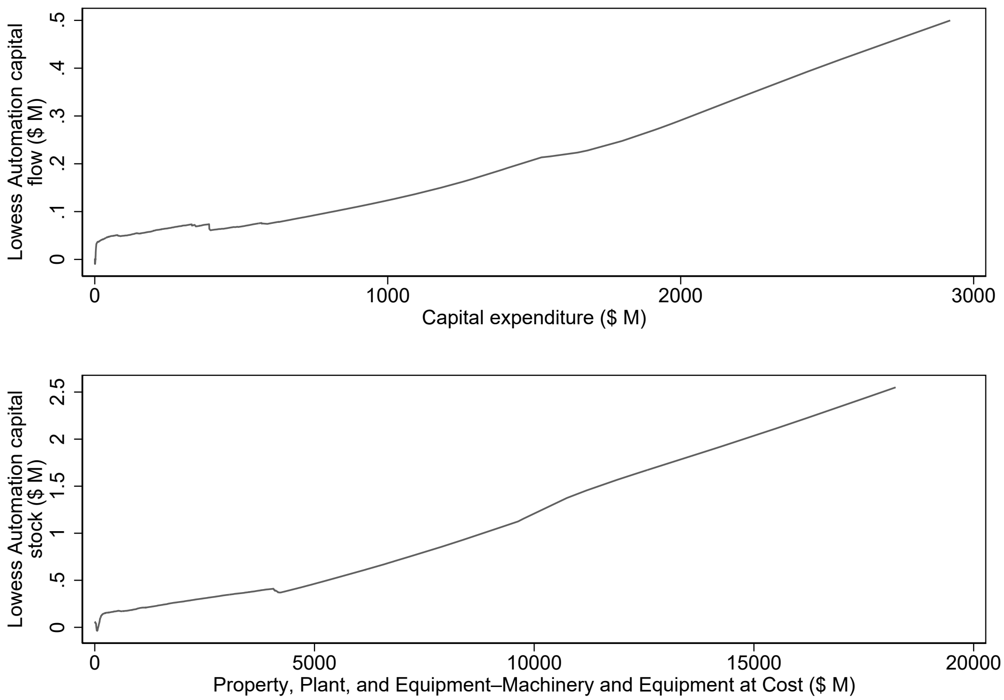

Figure 7: Lowess associations between PIERS-based automation data with accounting measures for matched Compustat sample

A.4.2 Industry-level elasticity of substitution (σ)

At least since Hicks (1932), economists have recognized the importance of substitutability of labor for capital for production decisions. That construct was formalized in the production function by Arrow et al. (1961) as σ. The literature on σ has focused on aggregate economy-level substitution (summarized in Chirinko, 2008), but some work has recently focused on estimating industry-level differences in substitutability (e.g., Chirinko and Mallick, 2017) and their implications (e.g., Belenzon, Bennett and Patacconi, 2019).

Theoretically, one might think of σ as representing the latent "automatability" of labor tasks in a production model, suggesting that investment in automation equipment should be greater in industries with higher labor–capital substitutability. This proposition provides an opportunity to further validate the PIERS-based measure using an outside data source.

In Table A4, I present the raw summary statistics for automation capital stock at establishments in discretized buckets of σ values. Automation capital stock is increasing monotonically in σ. I then graph the coefficients and standard errors for each bucket indicator.

Table A4: Raw summary statistics of automation capital stock value by industry-level σ bucket from Chirinko and Mallick (2017)

σ bucket	Automation capital stock Mean	Std. dev.	Observations	Unique NAICS4
.05–.1	1,065.585	98,993.663	421,799	3
.1–.15	2,920.940	320,835.8	469,082	6
.15–.2	4,747.048	212,192.06	774,843	8
.2–.25	5,167.954	246,275.94	2,246,647	11
>.25	5,758.253	583,432.32	1,141,797	177
Whole sample for which σ is available	4,685.868	348,044.96	5,054,168	45

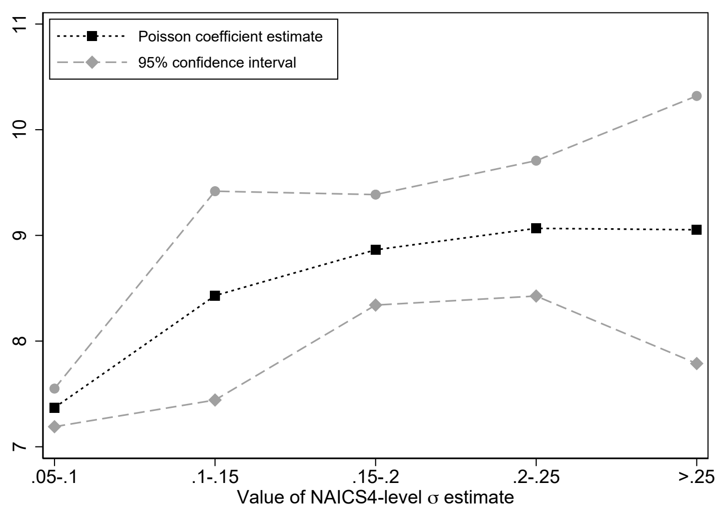

Figure 8: Validating the PIERS measure by correlating automation capital stock to industry-level σ from Chirinko and Mallick (2017)

Figure 8 shows the correspondence between the industry-level measure of σ from Chirinko and Mallick (2017) and the PIERS-based automation capital measure. To generate Figure 8, I estimate a Poisson regression of automation capital stock on indicators for being in a NAICS4 with a corresponding σ estimate from

Chirinko and Mallick (2017) and year fixed effects. Because σ estimates vary at the NAICS4 level, standard errors are allowed to cluster at that level.

Because Chirinko and Mallick (2017) estimate σ using factor share shifts from BEA and BLS' KLEMS data, [31] this is a completely independent measure of the phenomenon, so the correspondence between the two values supports the assertion that there is "signal" in the PIERS-based measure.

A.4.3 US Census Quarterly Survey of Plant Capacity

Another independently collected measure of investment in automation can be derived from the US Census Quarterly Survey of Plant Capacity (QSPC). Beginning in 2008 (with the exception of 2012), respondents were asked to select from a list of candidate reasons why they had excess plant capacity. Reasons include depressed demand as well as "Machinery Capital Expenditures." Publicly available data are available at the industry-by-quarter level, so I was able to aggregate the percentage of respondents who noted "Machinery Capital Expenditures" to the industry-by-year level and correlate that with the industry-by-year value of automation capital shipments derived from the PIERS-based measure.

Table A5: Relationship between NAICS3-by-year reported machinery capital expenditure in US Census Quarterly Survey of Plant Capacity and PIERS-based measure

VARIABLES	Automation capital investment			
	Flow ($M)		Stock ($M)	
	(1)	(2)	(3)	(4)
% reporting "Machinery and Capital Expenditure"	7.277***	197.677***	4.502***	636.674***
	(1.176)	(32.558)	(0.278)	(40.293)
Constant	-3.864***		0.550**	
	(1.174)		(0.277)	
Observations	36	36	36	36
Model	Poisson	Poisson	Poisson	Poisson
Marginal effects		Y		Y

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

[31] `https://www.bea.gov/data/special-topics/integrated-industry-level-production-account-klems`, accessed 2/22/2024.

Table A5 shows that the correlations between the QSPC-based measure and the PIERS-based measure are economically important and significantly different from zero. Because the QSPC measure includes all machinery expenditure and not only maritime shipments, it increases confidence in the PIERS-based measure.

A.5 Generalized model of weak and strong market dominance

This appendix describes a dynamic oligopoly model and, more specifically, a modified version of the Athey and Schmutzler (2001) incremental "investment precedes competition" model. In each of t periods, a firm plays a two-stage game in which it has the option to invest in marginal-cost–reducing automation and then compete in a generalized form of competition. I describe, under the maintained assumptions, the conditions under which two outcomes obtain: weak increasing dominance and strong increasing dominance. Weak increasing dominance is an outcome in which leading firms—those with a higher state variable—invest more. Strong increasing dominance refers to equilibria in which leading firms increase their market-share lead.

A.5.1 Setup

I firms, $i = 1, ..., I$, interact over T periods, $t = 1, ..., T$, where $T \leq \infty$. Each firm-by-period is fully characterized by a state variable $Y_i^t \in \mathcal{Y}_i$. $\mathcal{Y}_i$ is a partially ordered subset of $\mathbb{R}$. Unsubscripted $\mathcal{Y} = \times_i \mathcal{Y}_i$. Without loss of generality, I denote firm i having a larger market share than firm j in period t by $Y_i^t \geq Y_j^t$.

The state variable may, however, represent not market share itself but rather a variable that maps monotonically to market share. For our purposes, I interpret it as accumulated cumulative marginal-cost reduction achieved by firm i at time t from its investment in automation capital. In each period, firm i chooses its investment in automation capital $a_i^t \in \mathcal{A}_i$, based only on $\mathbf{Y}^{t-1}$. Let $\mathbf{a}^t = (a_1^t, ..., a_i^t)$ and $\mathcal{A} = \times_i \mathcal{A}_i$.

The vector of state variables at the start of the game is $\mathbf{Y^0} = (Y_1^0, ..., Y_I^0)$. From there, they evolve according to $Y_i^t = Y_i^{t-1} + y_i^t$. The vector of state variables across all firms in time t is denoted $\mathbf{y}^t = (y_1^t, ..., y_I^t)$. Each firm faces the same investment cost function k, which enters the firm's objective function as follows:

$$\Pi^i(\mathbf{a}^t, \mathbf{Y}^{t-1}) = \pi^i(\mathbf{a}^t, \mathbf{Y}^{t-1}) - k(a_i^t, Y_i^{t-1}). \tag{1}$$

For parsimony, I will continue with the assumption that $y_i^t = a_i^t$, which defines what Athey and Schmutzler (2001) refer to as an "incremental investment" model.[32]

In each period, each firm i chooses its investment level a_i^t and bears the cost $k(a_i^t, Y_i^{t-1})$ in the first stage. All the firms' choices combined yield $\mathbf{y}^t$. In the second stage, firms compete in the product market. Results are general to a wide range of competition models, so I don't need to model the competition explicitly; it suffices to make the following assumptions:

1. An equilibrium in the second-stage competition game exists and yields firm i equilibrium profits $\hat{\pi}^i(\mathbf{Y}^{t-1} + \mathbf{a}^t)$. The equilibrium profit function is the product of two variables: $\hat{\pi}^i(\mathbf{Y}) = D^i(\mathbf{Y}) \cdot M^i(\mathbf{Y})$. The first is firm i's equilibrium demand, which itself is a mapping of state vectors to the reals, $\mathbf{D}^i : \mathcal{Y} \to \mathbb{R}$. The second is the equilibrium markup, also a mapping from the state vectors to the reals, $\mathbf{M}^i : \mathcal{Y} \to \mathbb{R}$.

2. All relevant functions are differentiable.

3. Conditional on firms $z \neq i, j$ playing equilibrium actions $\mathbf{a}_{ij}^*(\mathbf{Y}^t)$, if firms i and j have different state variables, there is a unique equilibrium.[33] This is a weaker assumption than there being a unique equilibrium in the I-player game. This "conditional uniqueness" condition can be stated in math as follows: For each period t, firms i and j, and state vector $\mathbf{Y}^t$, if the firms choose different actions such that $Y_i^t \neq Y_j^t$ and there exist two vectors of equilibrium actions $\mathbf{a}^*(\mathbf{Y}^t)$ and $\tilde{\mathbf{a}}^*(\mathbf{Y}^t)$ such that $\mathbf{a}_{-ij}^*(\mathbf{Y}^t) = \tilde{\mathbf{a}}_{-ij}^*(\mathbf{Y}^t)$, then $a_i^*(\mathbf{Y}^t) = \tilde{a}_i^*(\mathbf{Y}^t)$ and $a_j^*(\mathbf{Y}^t) = \tilde{a}_j^*(\mathbf{Y}^t)$.

4. I assume Athey and Schmutzler's (2001) "exchangeability" condition. The name comes from the requirement that firm i's profits are what firm j's profits would be were firm j in firm i's position. Slightly more formally, the requirement is that any focal firm i's choices depend on the set of actions and state variables of opponents, but not on which action corresponds to which state variable or to which opponent.

[32]The incremental investment model is a special case of the more general model in which y_i^t can be some function $h(\mathbf{a}^t, \mathbf{Y}^{t-1})$ including realizations of random variables. That flexibility allows for a more general model allowing for network externalities or learning-by-doing. I direct interested readers to Athey and Schmutzler (2001) for the relevant discussion.

[33]The condition is implied by the "dominant-diagonal" conditions summarized in Tirole (1988), which themselves are implied by a weaker set of uniqueness conditions summarized by Vives (1999).

Exchangeability holds in many common models of competition, as described below. Intutively, it is related to the "anonymity" requirement in many of the value-based strategy models based on Brandenburger and Stuart Jr (2005) and Brandenburger and Stuart Jr (2007).

Mathematically, the exchangeability condition is that firms' profit functions are exchangable as functions of $\mathbf{a}^t$ and $\mathbf{Y}^{t-1}$, where exchangeability is defined as follows: A set of I functions $f^i : \times_i \mathcal{X}_i \rightarrow \mathbb{R} \forall i \in \{1, ...I\}$ is exchangable if $\forall i \neq j \neq k \in \{1, ...I\}$ the following three conditions hold:

(a) $\mathcal{X}_i = \mathcal{X}_j$

(b) $f^i(\mathbf{x}) = f^j(T_{ij}(\mathbf{x}))$

(c) $f^i(\mathbf{x}) = f^i(T_{jk}(\mathbf{x}))$

A.5.2 Weak increasing dominance

As noted in the introduction, a central concern for scholars studying competition is whether the availability of marginal-cost–reducing technology leads to a self-reinforcing spiral of the kind depicted in Klepper (1996). In this vein, firms invest to lower their costs and get bigger, and then have more incentive to lower costs—and the cycle repeats until they have dominant market share. A primary argument of my paper is that such a model conflates two properties that can—but need not—go together. The tendency for firms with more investment in marginal-cost–reducing technology to invest more is called "weak increasing dominance." In the formalism of this model, it refers to a situation in which equilibrium actions are increasing in the focal firm's state variable: $Y_i > Y_j \Rightarrow a_i^*(\mathbf{Y}) \geq a_j^*(\mathbf{Y})$. But that condition does not necessarily imply "strong increasing dominance," whereby a focal firm's *market share* is increasing in its state variable. Intuitively, for two reasons, it need not be the case that more investment results in increased market share. First, competitors may also be spending on marginal-cost–reducing capital, and the marginal returns to the leader's expenditure may be decreasing. This is true of many models of cost-reducing investment, particularly because it has the appealing feature of representing reality. When the cost associated with reducing marginal production costs is not decreasing, firms either don't invest at all or, if able, invest infinitely.

Given the plausibility that strong and weak decreasing dominance are not necessarily coupled, despite their frequent conflation, it is worth discussing the conditions in which the former implies the latter.

Proposition 1. *When firms are myopic in the incremental investment game,*[34]

$$\hat{\pi}^i_{Y_i,Y_i} \geq 0 \text{ and } \hat{\pi}^i_{Y_i,Y_j} \leq 0 \tag{2}$$

and

$$\pi^i_{a_i,Y_i} \geq k_{a_i,Y_i}$$

then

$$Y_i > Y_j \Rightarrow a^*_i(\mathbf{Y}) \geq a^*_j(\mathbf{Y})$$

Proof. If $f : \mathbb{R}^2 \to \mathbb{R}$ is smooth and $f_{x,y} \geq 0$, it satisfies increasing differences, as defined by Topkis (1979). As $\mathcal{Y} = \times_i \mathcal{Y}_i$ is a partially ordered product set, $f : \mathcal{Y} \to \mathbb{R}$ is supermodular in $\mathbf{y}$, then—given (a) the assumptions from above of conditional uniqueness, exchangeability, and the game being one of strategic substitutes, (b) the supermodularity of the function, and (c) the profit functions $\pi^i(\mathbf{a}^t, \mathbf{Y}^{t-1})$ satisfying increasing differences—Theorem 1 from Athey and Schmutzler (2001) shows that $y_i > y_j \Rightarrow a^*_i(y) \geq a^*_j(y)$. $\square$

In words, the listed conditions are sufficient for weak increasing differences. Athey and Schmutzler's (2001) proof follows from proving weak dominance for a two-player game and then using exchangeability to expand that result to the general case of $N > 2$ firms. The usefulness of this result, however, depends on how often the assumptions actually hold.

First, I address the assumption in Equation 2 regarding the cross-partials of the profit functions with regard to state variables. In their appendix, Athey and Schmutzler (2001) prove that the condition depicted in this manuscript's Equation 2 applies in a wide set of frequently studied competitive models in which, as here, the state variable represents the negative of marginal cost. Such models include (a) Either Bertrand or Cournot models in which firms produce differentiated goods at constant marginal cost for linear demand as long as firm i's demand is more sensitive to its own price than to that of its opponent, and (b) models of "horizontal"

[34] The model does not require myopia. Analogous results obtain with far sighted firms in both Open Loop and Markov Perfect equilibria under some additional assumptions. Athey and Schmutzler (2001) provide a full characterization.

competition with quadratic transportation costs, whether on a line (e.g., d'Aspremont, Gabszewicz and Thisse, 1979) or a circle (e.g., Salop, 1979).

Exchangeability and Conditional Uniqueness cannot be observed, but I can test the reduced-form prediction that the assumption depicted in Equation 2 correspond with weak increasing dominance. Markets whose mode of competition does not exhibit either of Exchangeability or Conditional Uniqueness would bias against finding a result.

As mentioned above, weak increasing dominance does not necessarily imply strong increasing dominance—a cycle of increasing market share for firms that already have higher market share.

The following section describes those theoretical predictions.

A.5.3 Strong increasing dominance

Here we continue with the incremental investment game under the relatively weak assumption that firm i's market share is decreasing in the state variables of all firms $j \neq i$. For total aggregate demand D, the formal condition under which strong increasing dominance would obtain is that

$$\left.\frac{\partial}{\partial \alpha} \frac{D^i(\mathbf{Y} + \alpha \mathbf{1})}{D(\mathbf{Y} + \alpha \mathbf{1})}\right|_{\alpha=0} \geq 0$$

where $\mathbf{1} = (1, ..., 1)$ and α is a scaling factor indicating how much is being added to the state variable vector $\mathbf{Y}$.

As the leading firm has the highest state variable,$Y_i \geq Y_j \forall i \neq j$, and $\frac{D^i}{D} > \frac{1}{I}$, this is equivalent to a restriction on the implications of a hypothetical change in the state variables for aggregate demand—essentially, whether investments are demand-stealing. Given the assumption of exchangability and that demand is linear in states, we can write

$$\frac{\frac{\partial}{\partial \alpha} D^i(\mathbf{Y} + \alpha \mathbf{1})}{\frac{\partial}{\partial \alpha} D(\mathbf{Y} + \alpha \mathbf{1})} = \frac{1}{I}. \tag{3}$$

Plugging Equation 3 back in above yields a very intuitive condition. Weak increasing dominance implies strong increasing dominance when investments are demand-stealing:

$$\frac{\partial}{\partial \alpha} D(\mathbf{Y} + \alpha \mathbf{1}) \bigg|_{\alpha=0} \leq 0.$$

Thus, the final implication is that weak increasing dominance holds for a commonly studied set of models of competition, but does not always imply strong increasing dominance. Given the assumptions above, weak increasing dominance implies strong increasing dominance when investments are sufficiently demand-stealing.

References

Arrow, Kenneth J, Hollis B Chenery, Bagicha S Minhas, and Robert M Solow. 1961. "Capital-labor substitution and economic efficiency." *Review of Economics and Statistics*, 43(3): 225–250.

Athey, Susan, and Armin Schmutzler. 2001. "Investment and Market Dominance." *RAND Journal of Economics*, 32(1): 1–26.

Barnatchez, Keith, Leland Dod Crane, and Ryan Decker. 2017. "An assessment of the National Establishment Time Series (NETS) database." *FEDS working paper*.

Belenzon, Sharon, Victor Manuel Bennett, and Andrea Patacconi. 2019. "Flexible Production and Entry: Institutional, Technological, and Organizational Determinants." *Strategy Science*, 4(3): 175–249.

Brandenburger, Adam M, and Harborne W Stuart Jr. 2005. "Value-based Business Strategy." *Journal of Economics & Management Strategy*, 5(1): 5–24.

Brandenburger, Adam M, and Harborne W Stuart Jr. 2007. "Biform Games." *Management Science*, 53(4): 537–549.

Chirinko, Robert S. 2008. "σ: The long and short of it." *Journal of Macroeconomics*, 30(2): 671–686.

Chirinko, Robert S, and Debdulal Mallick. 2017. "The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model." *American Economic Journal: Macroeconomics*, 9(4): 225–53.

Crane, Leland D, and Ryan A Decker. 2020. "Research with private sector business microdata: The case of NETS/D&B." *Federal Reserve Board Working Paper*.

d'Aspremont, Claude, J Jaskold Gabszewicz, and J-F Thisse. 1979. "On Hotelling's "Stability in competition"." *Econometrica: Journal of the Econometric Society*, 1145–1150.

Echeverri-Carroll, Elsie L, and Maryann P Feldman. 2019. "Chasing entrepreneurial firms." *Industry and Innovation*, 26(5): 479–507.

Fellegi, Ivan P, and Alan B Sunter. 1969. "A theory for record linkage." *Journal of American Statistical Association*, 64(328): 1183–1210.

Guzman, Jorge, and Scott Stern. 2016. "The State of American Entrepreneurship: New Estimates of the Quality and Quantity of Entrepreneurship for 32 US States, 1988-2014." National Bureau of Economic Research Working Paper 22095.

Hicks, John. 1932. *The Theory of Wages.* Macmillan.

Klepper, Steven. 1996. "Entry, Exit, Growth, and Innovation over the Product Life Cycle." *American Economic Review*, 86(3): 562–583.

Salop, Steven C. 1979. "Monopolistic competition with outside goods." *The Bell Journal of Economics*, 141–156.

Steorts, Rebecca, Brenda Betancourt, Andee Kaplan, Neil Marchant, and Beidi Chen. 2019. "Record Linkage Tutorial."

Tirole, Jean. 1988. *The theory of industrial organization.* MIT press.

Topkis, Donald M. 1979. "Equilibrium points in nonzero-sum n-person submodular games." *SIAM Journal on control and optimization*, 17(6): 773–787.

Vives, Xavier. 1999. *Oligopoly pricing: old ideas and new tools.* MIT press.