Valuable Words: The Price Dynamics of Internet Domain Names

Thies Lindenthal
Massachusetts Institute of Technology, Center for Real Estate, 120 Massachusetts Avenue, Cambridge, MA 02139. E-mail: thilin@mit.edu

This article estimates the first constant quality price index for Internet domain names. The suggested index provides a benchmark for domain name traders and investors looking for information on price trends, historical returns, and the fundamental risk of Internet domain names. The index increases transparency in the market for this newly emerged asset class. A cointegration analysis shows that domain registrations and resale prices form a long-run equilibrium and indicates supply constraints in domain space. This study explores a large data set of domain sales spanning the years 2006 to 2013. Differences in the quality of individual domain names are controlled for in hedonic repeat sales regressions.

Internet domain names bring location back to the otherwise location-less Internet economy. A domain name provides a virtual street address for any website or service on the Internet. It is comparable with a tract of land on which a business or just a private home page can be built. This space–network analogy is as old as the World Wide Web, and numerous terms related to the Internet exhibit a spatial connotation: Labels for technical network addresses, for instance, are called *domains*, users are *visitors*, Internet browsers have been baptized *Navigator* or *Explorer*, web*sites* are *home* pages, users communicate in chat *rooms*—the list can be easily extended.

Understanding domains as a novel form of land offers the opportunity to transfer established theoretical and empirical frameworks for the pricing of land into virtual space. Theoretically, Alonso-Muth-Mills models of urban layouts (Alonso, 1964; Mills, 1972; Muth, 1969) explain differences in land rents by differences in the distance to jobs or amenities. Applying this reasoning to domains, the price of a domain is hypothesized to depend on its "proximity" to potential users. Because a voyage on the World Wide Web usually begins with the user entering the domain name of the desired website into his or her web browser, distance to the user can be seen as the effort a user is required to make to correctly remember and type a domain name. An appealing domain name such as Apple.com is easy to recall and quickly entered. In this sense, an intuitive domain name is like a convenient downtown address linked to excellent transportation systems. Long or cryptic domain names are more burdensome, which is comparable with a longer commute to a location somewhere in the outskirts.

Ieong, Mishra, Sadikov, and Zhang (2012) provide a similar, albeit nonspatial explanation for the value of domains. They show that domains help users to evaluate the reliability of search results from online search engines. Domains serve as *brands* for the displayed information. Based on this line of thought, differences in domain prices could also stem from the brand potential inherent in the domain name. Again, catchy and easy-to-remember names sell at a premium above registration costs.[1]

Differences in "location" and "brandability" fuel a heated race for the shortest and most memorizable domain names that sprang up since the very first domain was created in March 1985. By now, more than 240 million unique domain names are registered (Verisign, 2012), with no end of growth in total numbers in sight. An active secondary market facilitates investments in domains. Exclusive domains oftentimes trade for five- or six-figure dollar amounts, and some for even more. The current record in reported sales prices is the widely covered $13 million transaction of www.sex.com in 2010.[2] Trading of and investing in domains has quickly evolved from a geeky pastime of a few to the serious bread and butter industry feeding hundreds of professionals today.

Received May 10, 2012; revised April 9, 2013; accepted May 2, 2013

 • Published online 22 January 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23012

[1]Similar to the real economy, successful brands are plagued by imitators. Edelman and Moore (2010) investigate brand infringements by so-called typo squatters, who intentionally register misspellings of popular website addresses. Whenever a user makes a typing error anticipated by a squatter, he or she is forwarded to a page that is usually cluttered with advertisements.

[2]"Sex.sells." *The Economist*. October 27, 2010. Retrieved from http://www.economist.com/blogs/dailychart/2010/10/domain-name_prices

The focus of this article is on estimating the first constant quality price index for Internet domain names. It adapts an empirical framework borrowed from real-estate research that is suitable for valuing infrequently traded and nonstandardized assets like houses, antiques, pieces of art—or, by analogy, virtual locations. The linguistic nature of domains causes substantial heterogeneity in their quality, which makes domain names comparable with traditional asset classes where the intrinsic value of an asset is not directly observable as well. The index spans 7 years and is updated on a monthly basis.

Despite rapid growth in the past decade, domain names are still a relatively small investment class that lacks any information on its inherent risk and return profile.[3] Market participants and investors simply do not know whether domain names are a "good" investment. Did domain holdings deliver positive returns in past years? How big were any returns?

Rational investors evaluate return and risk of their holdings simultaneously. The market index is the best proxy fundamental risk of domain names. The price volatility of a market portfolio hypothetically containing all domain names cancels out any domain-specific volatility. The fundamental or market risk of domains can be compared with the risk of an investor's portfolio, which puts any return on this portfolio into a risk-adjusted perspective. Furthermore, the fundamental risk of domains can be compared with the risk of other investment classes such as stocks, bonds, or real estate.

The massive increase in virtual space and new extensions (ICANN Internet Corporation for Assigned Names and Numbers, 2012a) scheduled for 2013 and 2014 by the Internet Corporation for Assigned Names and Numbers (ICANN) is, among other reasons, motivated by a perceived scarcity of domains. This article provides the first evidence that supply of domain space is currently constrained.

Finally, any cautious economist will surely ask the following: Are domain names for real or just another fad? Or does an economic rationale justify the prices paid? A comparison of domain prices with share price indices for information technology (IT) companies shows that domains are not a totally detached "new economy." The value of locations on the web is closely correlated with, for example, the NASDAQ-100.

The remainder of this article first presents a primer on Internet domain names, including a brief introduction on the nature of domain markets. The Method section provides an overview of the data this study analysis relies on. This motivates the choice of an adequate index estimation method in the subsequent section. The Conclusion section discusses the empirical results leading to general conclusions.

[3]A first glance at the price trends in domain sales are offered by industry sources such as Sedo.com (2011).

A Primer on Internet Domain Names

Domain names make the Internet navigable for users. All servers, computers, and other devices linked to the Internet use a unique Internet protocol (IP) number as their address within the network. Typing "74.125.226.201" into the address bar of a web browser leads to the website of a well-known search engine from California. Because IP numbers are difficult to remember for Internet users, the domain name system provides shortcuts, assigning more memorizable domain names to the underlying IP numbers.

Each domain name consists of at least two components: Top-level domains (TLDs) structure the overall name space into a limited number of subsets that either have global scope (gTLD), such as COM, NET, ORG, or that are country specific (ccTLD), such as DE, FR, NL.[4] Each TLD is subdivided into second-level domains (SLDs) that can consist of letters, numbers, and hyphens only. Spaces are not allowed. The combination of a TLD and SLD, separated by a colon, is commonly referred to as an Internet domain name. It cannot exceed 67 characters in total.[5]

From a lawyer's perspective, it is not resolved yet whether a domain name is an intangible asset or just a contractual right (Burshtein, 2005); in practice, the case is clear: Domains are treated and traded like assets.

The primary market for domain names resembles the land rushes in 19th-century America (Lindenthal, 2011). The general rule is the following: Whoever files a registration for a domain first receives this domain. No costs but a modest registration fee to the TLD registry are incurred. As a result of this liberal first-come-first-serve approach, the number of domain names grew rapidly in the past decade. Figure 1 shows the growth in the total number of registered .com/.net/.org domains since 1998.

A domain must be renewed at periodic intervals, again incurring a fee. As long as the renewal fees are taken care of, the domain owner enjoys full ownership benefits, including the possibility to rent out the domain or to sell it on the secondary market. Domains that do not get renewed become available for registration again. About 27% of .com/.net domains were not continued in 2011 (Verisign, 2012) but handed in again. Alternatively, owners can sell domains they do not wish to develop themselves in secondary markets.

The majority of resales of domains are closed privately or through competing trading platforms that offer marketing and transaction handling services to buyers and sellers. No trading licenses or other formal requirements restrict access to the market.

Estimating the size of secondary markets is difficult because no central registry explicitly keeps track of domain sales. ICANN reports 8 million transfers of .com/.net domains for January 2011, which is the upper bound of total

[4]Some countries subdivide their name spaces again, such as CO.UK for commercial United Kingdom domains and the AC.UK name space reserved for academic usage.

[5]More hierarchical layers such as mail.SLD.TLD or www.SLD.TLD can be added if needed.

DOI: 10.1002/asi

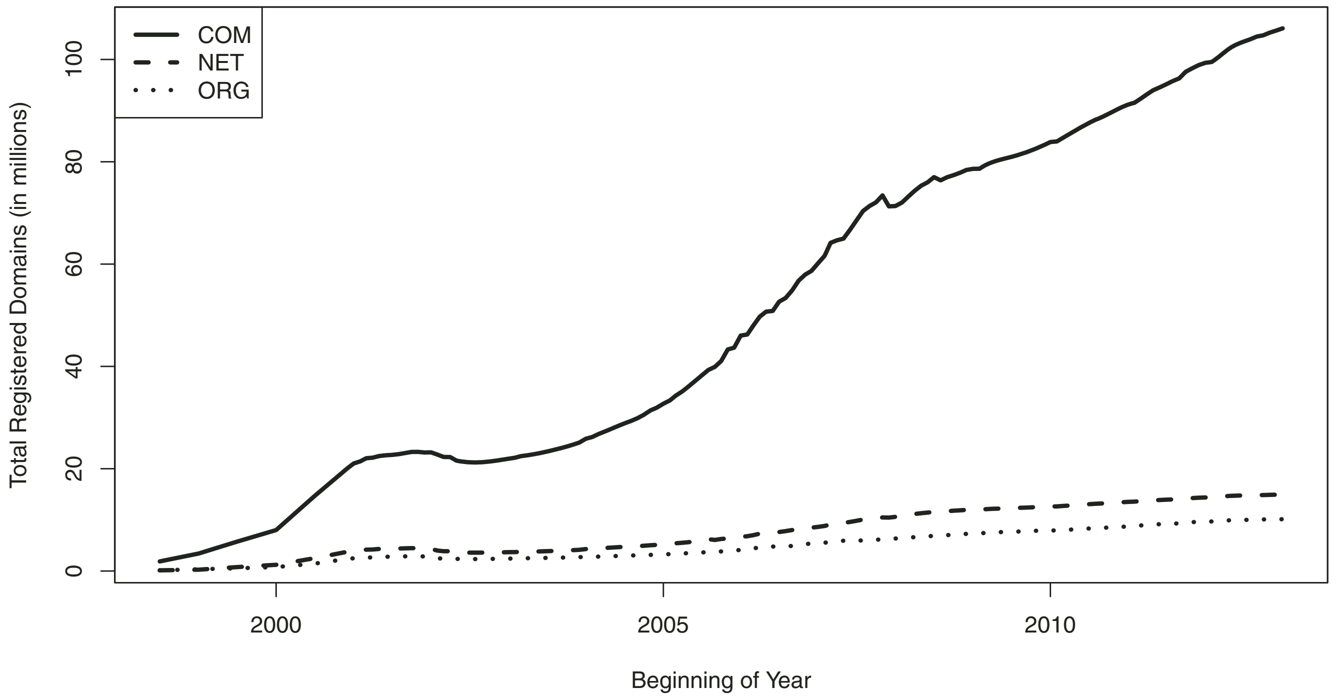

FIG. 1. Number of registered domain names for selected TLDs. A heated race for the best domain names sprung up since the first domain name in history, symbolics.com, was created in March 1985. By now more than 246 million unique domain names are registered (Verisign, 2012), with no end of growth in total numbers in sight. (Data from Zooknic [http://www.zooknic.com] and ICANN [http://www.icann.org/en/resources/registries/reports].)

sales as this number contains many not-for-sale transfers. The number of domain transfers across whole-sale registration companies could serve as a better proxy for domain transactions. In January 2011, 0.65 million .com or .net domains were transferred from one registrar to another (ICANN Internet Corporation for Assigned Names and Numbers, 2011).

Scenarios under which domains trade are manyfold. Entrepreneurs who registered a domain to host a new website sell the domain when the intended project does not materialize.[6] Owners of an operating website might move its content to a different location in case they receive a favorable purchase offer from other developers who can utilize the domain name more effectively. Finally, domain investors have registered wide swathes of address space hoping to resell domains at a profit to end users later.

In all of these cases, buyers and sellers negotiate freely, and it is reasonable to assume that prices are distributed around fair market value. However, domain speculation also has a dark side in which so-called typo squatters register misspellings of popular domains (Banerjee, Rahman, & Faloutsos, 2011; Edelman & Moore, 2010) or trademarks to benefit from advertisements placed at these domains. Allthough legal remedies[7] exist to protect domain or trademark owners against these registrations in bad faith, oftentimes the safest and easiest solution is simply to purchase the domain from the perpetrator. Such transactions do not represent fair market values. Fortunately, the domain marketplace providing for the data for this article screens the domains listed for sale and rejects typosquatters. This ensures fair transactions and a clean data set.

Data

This study relies on a data set of all domain transactions facilitated by the trading platform Sedo.com, which is one of the largest domain marketplaces in the world based on completed transactions and sales volume. The sample comprises 243,291 transactions for the period January 2006 through January 2013. Figure 2 presents the number of observations per quarter. The sample is not uniformly distributed in time, with more transactions taking place in later quarters.

Completed transactions are a small subset of all domains offered for sale. In 2010, about 14 million domains were listed at Sedo.com, in contrast with just 44,000 sales in that year. The low initial registration fee of just a few dollars per domain and the potentially high returns in case of a

[6]Verisign (2012) estimates that 44% of the 118.5 million domains registered under the .com or .net extensions in 2012 were used merely extensively: 21% led to minuscule online presences of one page only, and 13% did not resolve to any content at all. Extrapolating these numbers to the universe of all domains, about 100 million domain names are still waiting to be developed by their current owners.

[7]See AntiCybersquatting Consumer Protection Act (ACTA, 1999) or ICANN's Uniform Domain Name Dispute Resolution Policy (ICANN Internet Corporation For Assigned Names and Numbers, 2012b).

DOI: 10.1002/asi

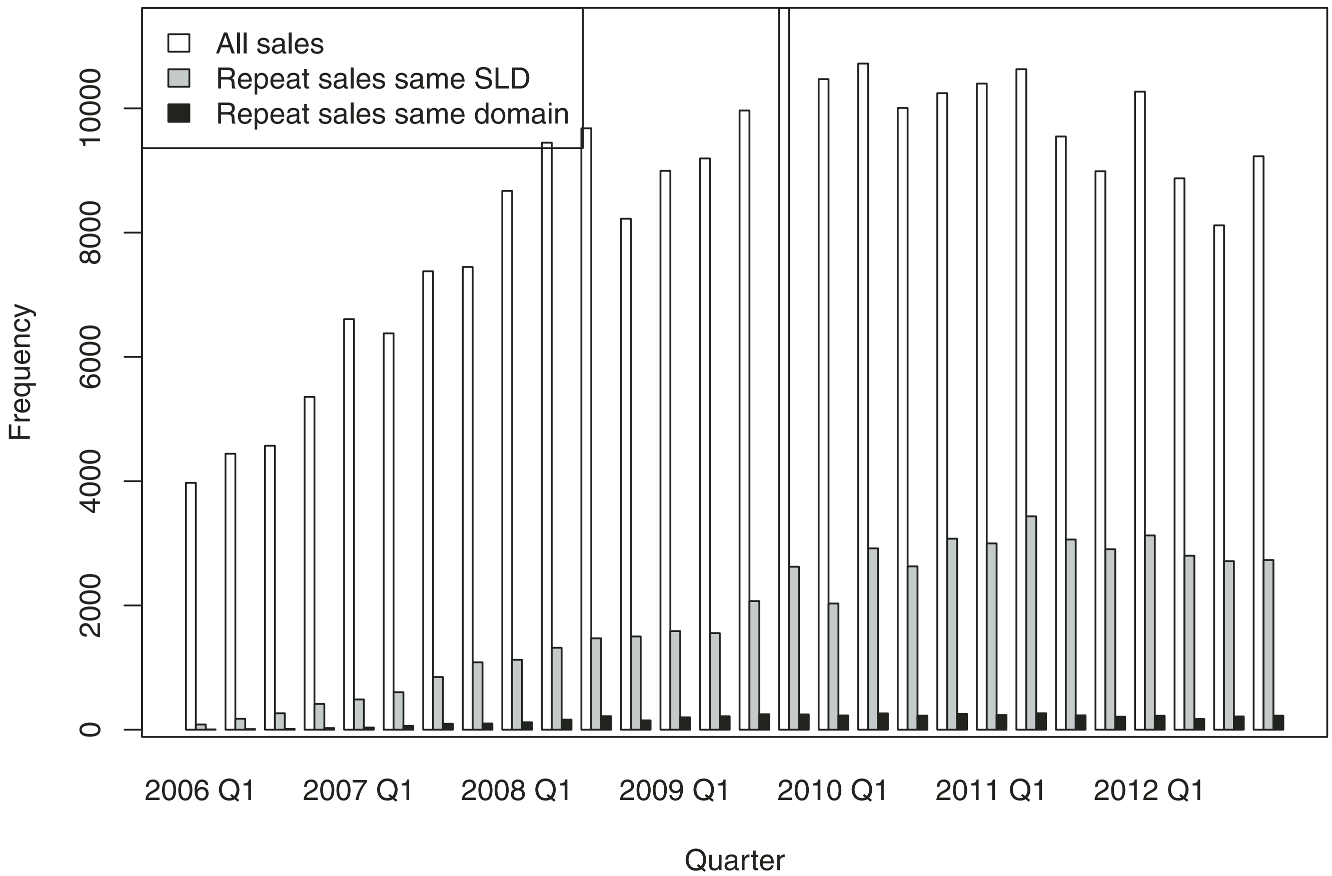

FIG. 2. Distribution of sales per quarter (2006Q1–2012Q4). The sample is not uniformly distributed in time, with more transactions taking place in later quarters. The number of pairs of sales sharing the same SLD is about one fifth of the number of sales. For sales pairs of exactly the same domain, the count is only 2% of the number of sales.

successful sale warrant a lottery-like market where domains are registered and put up for sale on a large scale. A back-of-an-envelope calculation illustrates that domain trading is profitable at the aggregate level: Multiplying the median transaction price in 2010 ($500) by the number of transactions (44,000) gives a total turnover that is several times bigger the sum of all registration and renewal fees ($6 to $10 per year and domain). For the data at hand, Sedo.com's fee structure is likely to induce substantial cheating as sellers first use the platform for marketing purposes but close the deal privately to avoid the fees due when domains change hands through the platform's official channels, causing low sales figures.

In May 2011, 28% of all sales are closed at a fixed price set by the seller in advance, whereas for 42% of trades, the price is set in bilateral price negotiations, where the buyer has to give a first price quote. Thirteen percent of transactions are initiated by professional brokers affiliated with Sedo.com. Seventeen percent of sales were closed in auctions. The selling mechanism is not a true hedonic characteristic because it is not necessarily linked to the quality of the domain name. Still, including this information in a hedonic framework improves the model fit, as Bulow and Klemperer (1996) show that auctions lead to higher transaction prices than negotiations with one less bidder.

All transactions that include a website or other content beside the naked domain name are excluded from the sample because disentangling the value of the domain from other factors such as an existing user base, software, or other content is impossible.

The sample does not contain any information on the characteristics of buyers and sellers.

Figure 3 displays a histogram of prices in the sample. The median transaction price is $500. The sales price distribution has an enormous right tail, with 12 domains exceeding $1 million. The "biggest fish" in the sample is sex.com, which was sold for $13 million in November 2010. The average SLD has 9.87 characters, 1.8% of the sample's SLDs contain diacritic characters, 5.68% include at least one digit, whereas 9% are split by dashes (Table 1).

The most senior COM/NET domain in the sample was registered in early 1989. The average year of registration for sold domains slowly deteriorates throughout the years. Domains sold in 2006 on average originated from 2003, compared to 2005 for sales in 2010.

DOI: 10.1002/asi

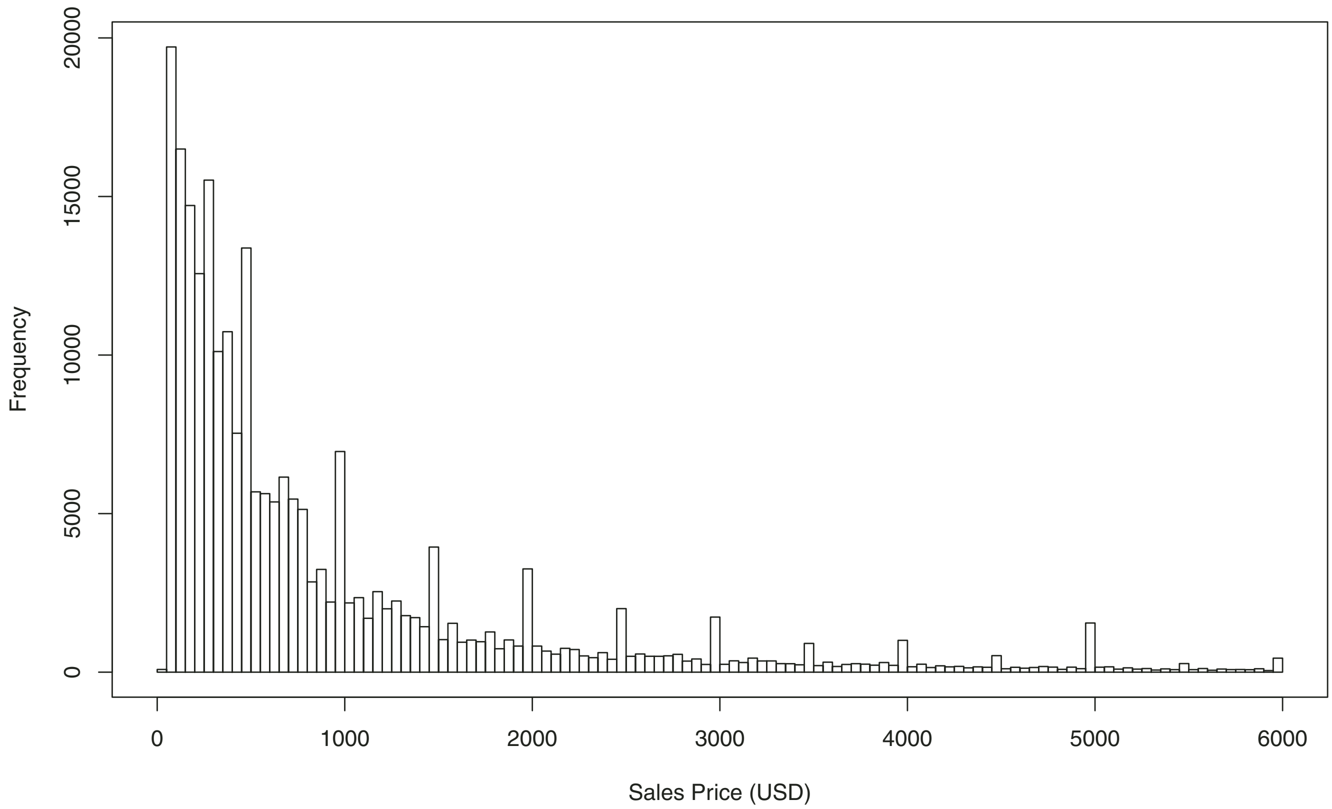

FIG. 3. Distribution of transaction prices. The top 5% of transaction prices is not shown. The median sales price is $500, whereas the maximum price in the databases is $13 million.

TABLE 1. Summary statistics.

Sample	N	Length in characters	Share of domains containing		
			Digits	Dashes	Diacritics
All COM domains	107,000,000	14.36	7.00%	11.52%	0.88%
Domains listed for sale	16,000,000	12.59	5.30%	11.47%	1.76%
Domains sold (Sedo.com)	243,291	9.87	5.68%	9.00%	1.80%
Repeat sales same SLD	52,811	7.62	5.44%	3.81%	1.56%
Repeat sales, same SLD and TLD	4729	7.48	14.65%	4.67%	1.18%

Note. Domain listings and sales are taken from http://Sedo.com (2006–2013) and are therefore a subsample of all sales and listings. Sold domains are, on average, shorter than both listed domains and domains registered and contain less digits and dashes than registered domains. This suggests a filtering based on domain quality: "Better" domains trade more frequently.

In total, 156 different TLDs are present in the database. Figure 4 compares the share of the nine most frequent TLDs within the sample with their share in the universe of registered domains. COM accounts for almost half of all domains, followed by DE, NET, ORG, and UK. The weights in the sample do not divert much from the true weights based on total registrations for most TLDs. The country-code TLD DE is overrepresented, probably because of Sedo's strong roots in the German market. Less popular domains that are subsumed under the label "other" are less prominent in the sales sample. The wide distribution across international TLDs is a unique feature of sales data from Sedo.com, besides the outstanding size of the sample.

Method

Domain names are unique: By definition, it is not possible to have the same combination of TLD and SLD twice. When estimating the value of a domain name, one cannot draw a direct comparison with recent transactions of exactly the same (or at least very similar domains), as it is common for pricing standardized goods like stocks or bonds.

DOI: 10.1002/asi

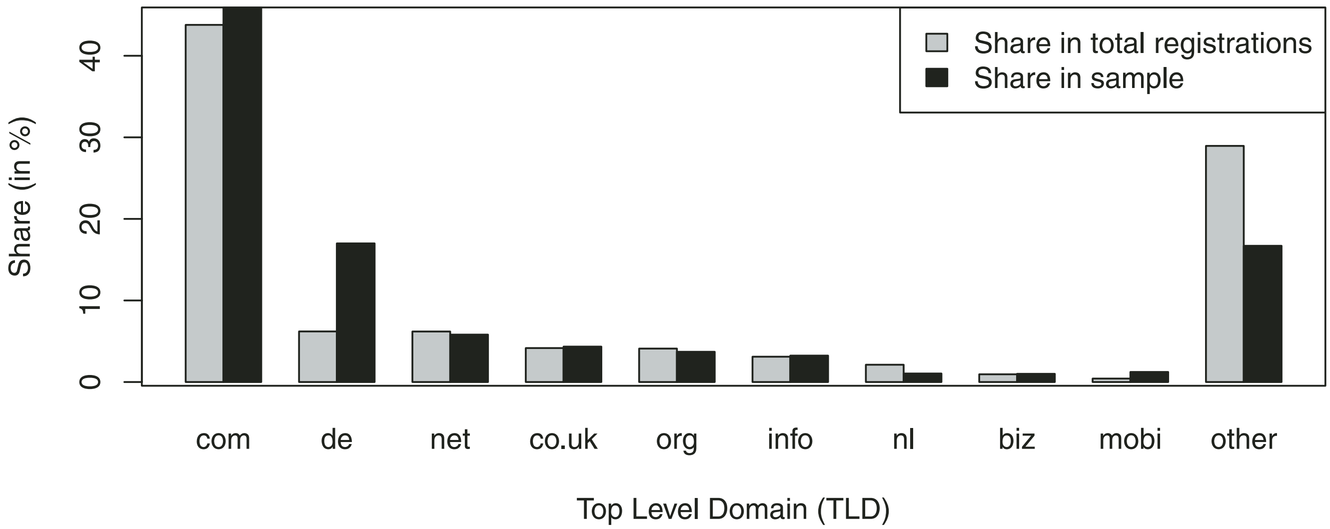

FIG. 4. Distribution of TLDs. The grey bars represent the market share of selected popular TLDs. Domain registration numbers provided by Denic (2011) and Verisign (2012). The black bars show the weight of a TLD in the sample. In total, 156 different TLDs are present in the database (top nine are shown). The country-code TLD .DE is overrepresented, probably because of Sedo's strong roots in the German market.

Taking the average of transaction prices gives a first indication of price developments in the market for domains but does not control for quality differences in the domains sold. If there is a time period in which many high-quality domains are sold, the average transaction price will increase, regardless of any true trend in prices. Median-based indices are less sensitive to extreme values but suffer from the same systematic shortcoming of disregarding the characteristics of the underlying transactions. Putting it differently, one is comparing apples with oranges.

Two general methods are commonly used to price nonstandardized assets such as real estate or art. The first one, the so-called hedonic regression analysis, explains the price of an asset by a set of quality variables that describe the characteristics (hedonics) of that asset. For example, when investigating the sales prices of cars, hedonic characteristics one might think of are mileage, year of production, manufacturer, and so on. Differences in quality are captured in the regression coefficients of the hedonic variables, whereas a general price trend can be separated from potential changes in the underlying quality of the traded assets. The overall explanatory power of this approach depends on how well the hedonic variables capture the price-relevant characteristics of the asset. For domains, this method is not feasible because only very few dimensions of the quality of domains can be described by quantitative variables.

Alternatively, the repeat sales method (Bailey, Muth, & Nourse, 1963) controls for quality by tracing individual assets in time, comparing each transaction with previous transactions for the very same asset. It assumes that the quality of the asset does not change between the two sales. Although this is probably the most direct form of making sure to "compare apples with apples," it disregards all information on transactions that are sold only once. Furthermore, repeat sales can cause a sample selection bias because domains that trade more than once could be systematically different from domains that trade only once.[8] In practice, this approach is suitable for samples of nonstandardized goods that provide sufficient numbers of repeat sales.

This article follows the hedonic repeat sales (HRS) method, first suggested by Shiller (1993). HRS combines the advantages of having individual dummies for each SLD and a hedonic classification of the highly standardized TLD. It uses the simplicity of repeat sales estimation while alleviating the problem of sample selectivity. Using repeat sales avoids the problem of identifying explanatory variables on domain quality and circumvents the omitted variables problem that is pronounced for domains.[9] Including the hedonic characterization for the TLD extends the number of observations entering the regression more than 10 times from 4,729 repeat sales of exactly the same domain to 52,811 pairs with the same SLD. To give an example, a repeat sales regression only considers repeat transactions of exactly the same domain (like XYZ.COM), whereas a HRS approach includes single sales that share the SLD but have different TLDs (such as XYZ.COM and XYZ.NET).

Conceptually, the sales price P of domain i is split into three components:

$$P_{i,t} = SLD_i + TLD_i + D_t \tag{1}$$

[8]See Gatzlaff and Haurin (1997) for a discussion of sample selection bias in repeat sales indices for residential real estate.

[9]When testing a hedonic regression approach for the data at hand, only 35% of the variation in domain prices in the sample can be accounted for. Furthermore, a theoretical framework on which factors to include as independent variables is missing.

DOI: 10.1002/asi

where *SLD* and *TLD* capture the quality of each SLD and TLD. Time dynamics enter the equation through D_t. Differencing transactions that share the same SLD cancels out the *SLD* component:

$$P_{i,t1} - P_{j,t2} = (SLD_i - SLD_j) + (TLD_i - TLD_j) + (D_{t1} - D_{t2}) \\ = TLDCombi_{i,j} + D_{t1-t2} \quad (2)$$

which leads to an equation that can be empirically tested in a regression:

$$ln(P_{i,t2}) - ln(P_{j,t1}) = \sum_{k=1}^{k=K} \gamma_k TLDCombi_{k,i,j} + \sum_{t=2}^{t=T} \beta_t D_{t,i,j} + \varepsilon_{i,j} \quad (3)$$

where the difference in the natural logarithm of prices from transaction *i* and *j* is the dependent variable, explained by a set of time dummies D_t and TLD combination dummies. For each pairwise combination *k* of TLDs in the sample, $TLDCombi_k$ is defined to be 1 if TLD_i equals the first element and TLD_j equals the second element of *k*, and zero otherwise. *K* is the squared number of unique TLDs in the sample. The time dummy D_t is defined in a Bryan and Colwell (1982) way, as described in Geltner (1997). For each time period between *t*1 and *t*2, D_t is set to 1. When only parts of a given time period fall between the two sales dates, the value for D_t is scaled down accordingly. If *t*1, for instance, is June 30, 2009, the annual dummy for 2009 will be set to 0.5; for November 31 it will be adjusted to 1/12. The default value is 0. This method provides end-of-time-period return estimates and avoids averaging within each time period (Geltner).

γ_k and β_t are regression coefficients. The relative price differences between TLDs are kept constant in time. For each pair of TLDs (*a*, *b*), the coefficients $\gamma_{k=(a,b)}$ are restricted to be $-\gamma_{k=(b,a)}$. The error term $\varepsilon_{t,i,j}$ is assumed to be independently and identically distributed.

The goal of this article is the estimation of an index that can be updated frequently and that still provides reliable index figures not subject to excess volatility. When estimating a high-frequency index, the number of observations per time period will be low. This causes the resulting index to be sensitive to noise, resulting in excessively volatile index estimates with low signal-to-noise ratios. Goetzmann (1993) further shows that repeat sales regressions suffer from spurious negative autocorrelation in the estimated return series and an excess return volatility, especially at the beginning and at the end of the series where the data thin out.

Imposing a structure on the time coefficients β_t reduces the effect of transaction price noise in thin markets. Goetzmann (1993) therefore suggests a Bayesian shrinkage technique where the log levels are modeled as a random walk with drift process. Francke (2010) generalizes this approach and extends the random walk with drift model to a structural time-series model.

This article follows a novel frequency conversion approach suggested by Bokhari and Geltner (2010). In a first step, lower frequency indices are estimated staggered in time. For the domain name index, we estimate 12 annual

TABLE 2. Regression estimates for 1 of 12 first-stage regressions.

Variable	Estimate	Standard error	*t* value	*P*(>|t|)
Year 1	0.210	0.088	2.396	.017
Year 2	0.291	0.060	4.853	.000
Year 3	–0.335	0.051	–6.525	.000
Year 4	0.133	0.047	2.835	.005
Year 5	0.127	0.047	2.700	.007
Year 6	0.053	0.053	0.994	.320
Year 7	0.005	0.075	0.073	.942
TLD_com_net	–1.358	0.053	–25.584	.000
TLD_com_org	–1.751	0.071	–24.534	.000
TLD_com_info	–2.282	0.078	–29.098	.000
TLD_com_nl	–1.510	0.194	–7.769	.000
TLD_com_me	–4.395	0.183	–24.062	.000
TLD_com_de	–0.528	0.046	–11.504	.000
TLD_com_co_uk	–1.330	0.075	–17.728	.000
TLD_com_es	–2.034	0.148	–13.743	.000
TLD_com_eu	–1.734	0.070	–24.670	.000
TLD_net_org	–0.522	0.071	–7.384	.000
TLD_net_info	–0.953	0.075	–12.714	.000
TLD_net_nl	0.450	0.290	1.552	.121
TLD_net_me	–2.127	0.170	–12.515	.000
TLD_net_de	0.711	0.074	9.670	.000
TLD_net_co_uk	0.497	0.108	4.615	.000
TLD_net_es	–0.693	0.165	–4.187	.000
TLD_net_eu	–0.509	0.081	–6.314	.000
TLD_org_info	–0.673	0.083	–8.135	.000
TLD_org_nl	0.883	0.290	3.042	.002
TLD_org_me	–1.936	0.216	–8.959	.000
TLD_org_de	1.057	0.090	11.717	.000
TLD_org_co_uk	0.773	0.139	5.576	.000
TLD_org_es	–0.025	0.184	–0.136	.892
TLD_org_eu	–0.165	0.088	–1.888	.059
TLD_info_nl	0.656	0.240	2.733	.006
TLD_info_me	–0.750	0.154	–4.880	.000
TLD_info_de	1.392	0.081	17.158	.000
TLD_info_co_uk	1.852	0.121	15.278	.000
TLD_info_es	0.209	0.163	1.288	.198
TLD_info_eu	0.283	0.081	3.481	.001
TLD_nl_me	–0.472	0.395	–1.194	.233
TLD_nl_de	0.343	0.180	1.908	.056
TLD_nl_co_uk	–0.084	0.258	–0.325	.745
TLD_nl_es	–0.104	0.250	–0.417	.677
TLD_nl_eu	–0.420	0.250	–1.679	.093
TLD_me_de	1.983	0.171	11.573	.000
TLD_me_co_uk	2.108	0.188	11.204	.000
TLD_me_es	1.094	0.218	5.008	.000
TLD_me_eu	0.786	0.160	4.918	.000
TLD_de_co_uk	0.396	0.098	4.043	.000
TLD_de_es	–0.746	0.152	–4.915	.000
TLD_de_eu	–1.015	0.066	–15.330	.000
TLD_co_uk_es	–1.127	0.203	–5.542	.000
TLD_co_uk_eu	–1.063	0.117	–9.111	.000
TLD_es_eu	–0.236	0.178	–1.325	.185

Note. The estimation of Equation 3 contains seven annual dummy variables. In total, 12 annual indices are estimated, each starting at a different month of the year (table shows first estimation only). A second step converts the staggered annual indices to an index at monthly frequency (Bokhari & Geltner, 2010). The lower panel contains coefficients for the pairwise TLD comparisons *TLDCombi*. The adjusted R^2 is 0.32.

DOI: 10.1002/asi

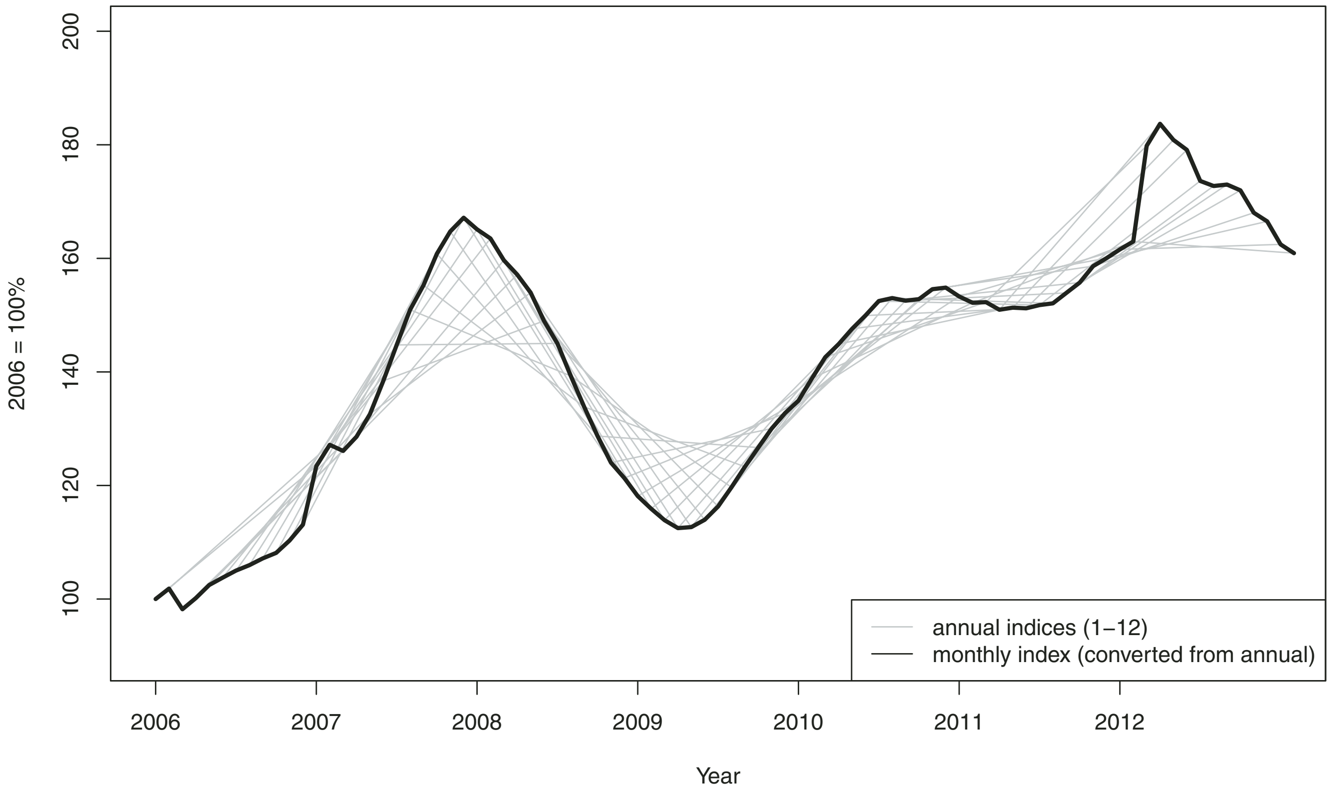

FIG. 5. Staggered annual indices for all TLDs and resulting monthly index. Twelve annual indices based on Equation 3 are estimated, each starting in a different month. Applying a generalized inverse estimator converts these annual indices to a monthly return series (black line).

indices based on Equation 3, each starting in a different month.[10] Applying a generalized inverse estimator converts these annual indices to a monthly return series.

The Bokhari–Geltner approach relies on few assumptions. Any index estimation technique can be used in the first stage. No time structure on the high-frequency time coefficients need to be formalized. The resulting higher frequency index has a better signal-to-noise ratio than directly estimated high-frequency indices. In addition, no time lag is introduced. Bokhari and Geltner (2010) provide a very detailed step-by-step explanation of the method.

Results

Table 2 presents the coefficients of 1 of the 12 first-stage regressions based on Equation 3. In total, 12 annual indices are estimated based on the seven time dummies (years 1–7), each index starting at a different month of the year (the table shows first estimation for index starting in January only). The lower panel contains coefficients for the pairwise TLD comparisons *TLDCombi*. The adjusted R^2 for the first-stage regressions is 0.32.

In a second step, the 12 annual indices are converted to one index at monthly update frequency. All 13 indices are displayed in Figure 5. The solid black line represents the resulting aggregate price trend at a monthly frequency, whereas the first-stage indices are depicted by thin lines. Overall, the domain price index increased by 63% from January 2006 through January 2013. The average historic return is 0.72% per month with a standard deviation of 2.6%.

Table 3 presents the Internet Domain Name Index (IDNX) estimates on a monthly frequency for the years 2006 through 2013. Internet domain names gain in value in 2006 through 2007, with prices peaking in November 2007 (+67%) before losing one third of their value in the subsequent 16 months. Since then, domain prices regained their strength, climbing to an all-time high in March and April 2012 (Figure 6), again followed by a sharp decline. The pronounced cycle demonstrates that domain names offer attractive investment opportunities to the skilled (or lucky) investor who can identify boom and bust phases in advance.

These results also show that domain names are risky investments—the observed bust phase in 2008 wiped out a

[10]Case and Shiller (1987) expect the variation in the error terms of a repeat sales regression to increase in the time between sales. Any heteroscedasticity in the error terms is a violation of the assumptions underlying ordinary least squares regressions. They therefore suggest to first run an auxiliary regression, following the method of Bailey et al. (1963). In a second step, they regress the error terms derived from the first step and regress them on the time between sales. Based on fitted values from this regression, they construct weights for a final re-estimation of step one by generalized least squares. The annual indices do not exhibit significant coefficients in the second step, which indicates that the Case–Shiller correction is not needed for this data.

DOI: 10.1002/asi

TABLE 3. IDNX, 2006–2013.

Date	IDNX	Date	IDNX
2006-01-01	100.0	2010-01-31	138.8
2006-01-31	101.8	2010-02-28	142.6
2006-02-28	98.2	2010-03-31	144.9
2006-03-31	100.2	2010-04-30	147.6
2006-04-30	102.5	2010-05-31	149.9
2006-05-31	103.8	2010-06-30	152.5
2006-06-30	105.0	2010-07-31	153.0
2006-07-31	106.0	2010-08-31	152.5
2006-08-31	107.2	2010-09-30	152.8
2006-09-30	108.1	2010-10-31	154.6
2006-10-31	110.3	2010-11-30	154.8
2006-11-30	113.1	2010-12-31	153.3
2006-12-31	123.4	2011-01-31	152.2
2007-01-31	127.2	2011-02-28	152.3
2007-02-28	126.1	2011-03-31	150.9
2007-03-31	128.6	2011-04-30	151.3
2007-04-30	132.6	2011-05-31	151.2
2007-05-31	138.4	2011-06-30	151.8
2007-06-30	144.8	2011-07-31	152.1
2007-07-31	150.9	2011-08-31	153.9
2007-08-31	155.3	2011-09-30	155.7
2007-09-30	160.8	2011-10-31	158.6
2007-10-31	164.7	2011-11-30	160.0
2007-11-30	167.2	2011-12-31	161.6
2007-12-31	165.1	2012-01-31	163.0
2008-01-31	163.5	2012-02-29	179.8
2008-02-29	159.7	2012-03-31	183.7
2008-03-31	157.1	2012-04-30	180.9
2008-04-30	154.0	2012-05-31	179.1
2008-05-31	149.0	2012-06-30	173.6
2008-06-30	145.0	2012-07-31	172.7
2008-07-31	139.3	2012-08-31	173.0
2008-08-31	133.8	2012-09-30	172.0
2008-09-30	128.7	2012-10-31	168.0
2008-10-31	124.0	2012-11-30	166.5
2008-11-30	121.3	2012-12-31	162.5
2008-12-31	118.1	2013-01-31	160.9
2009-01-31	115.9		
2009-02-28	113.9		
2009-03-31	112.5		
2009-04-30	112.7		
2009-05-31	114.0		
2009-06-30	116.3		
2009-07-31	119.7		
2009-08-31	123.3		
2009-09-30	126.7		
2009-10-31	130.0		
2009-11-30	132.7		
2009-12-31	134.9		

Note. All index figures can be downloaded at http://idnx.com/idnx.csv.

large share of market value before bouncing back. There are simply no risk-free gains to be made, no "free lunch" waiting to be consumed. In addition, liquidity in the market dries up just when it is needed most. Investors trying to liquidate their domain name holdings in 2008 had to find a buyer in relatively thin markets, indicated by reduced total transaction numbers during the bust period (Figure 2).

Domain prices have an economic foundation. They are not detached from the economy in general. On the contrary, Figure 6 plots the NASDAQ-100 index, which covers the 100 largest technology companies listed on the NASDAQ stock market along IDNX. The close resemblance of both lines indicates that domain price changes are very similar to changes in the IT economy. The estimated correlation on monthly changes in both indices is 0.19. When the Internet economy is expanding, domain names as one production factor are in high demand as well.

Furthermore, one can observe a close link between prices paid for domain names and revenues from online advertisement, which form an important source of income for the majority of Internet enterprises (Evans, 2009). This dependency is visible in the close comovement of advertising revenues tracked by IAB (2012) and domain prices in the boom years before 2008. Subsequently, the financial crisis paused any growth in advertisement spendings and sent domain prices on a downward trajectory. In the time period from 2009 through 2010, advertising revenues and domain prices jointly recovered: From June 2009 through January 2011, advertisement revenues expanded by 37% and domain prices by 32%. However, starting in the second half of 2010, domain prices exhibit a lower growth rate and cannot keep pace with rapidly expanding advertising money.

The number of registered domain names and the level of resale prices are two different views on the overall demand for domain names—the first tracking the primary domain market and the latter the secondary or resale market. Figures 6 and 7 suggest a comovement of changes in prices and domain registrations, indicating an equilibrium between registrations and prices. Whenever prices are high, more domains get created, and in times of falling prices, the growth of registrations also comes to a halt.

We conduct a cointegration analysis[11] to formally test for any long-run equilibrium between the primary and secondary market. Following Pfaff (2008) in notation, first a general vector autoregressive (VAR) model is defined as:

$$y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + BX_t + u_t \tag{4}$$

where y_t is a vector of the quarterly changes in prices in the secondary domain market (estimated by IDNX) and changes in the number of registrations in the primary domain markets for COM, NET, and ORG domains and p the number of lags. The A_i are (2×2) coefficient matrices for $i = 1, \ldots, p$ and u_t is a two-dimensional white noise process. Quarterly changes in revenue from online advertising and seasonal dummies enter the equation as exogenous variables (X_t). Based on this VAR, a transitory vector error correction model (VECM) can be specified as:

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} y_{t-p+1} + \xi X_t + u_t \tag{5}$$

[11]See Johansen (1988).

DOI: 10.1002/asi

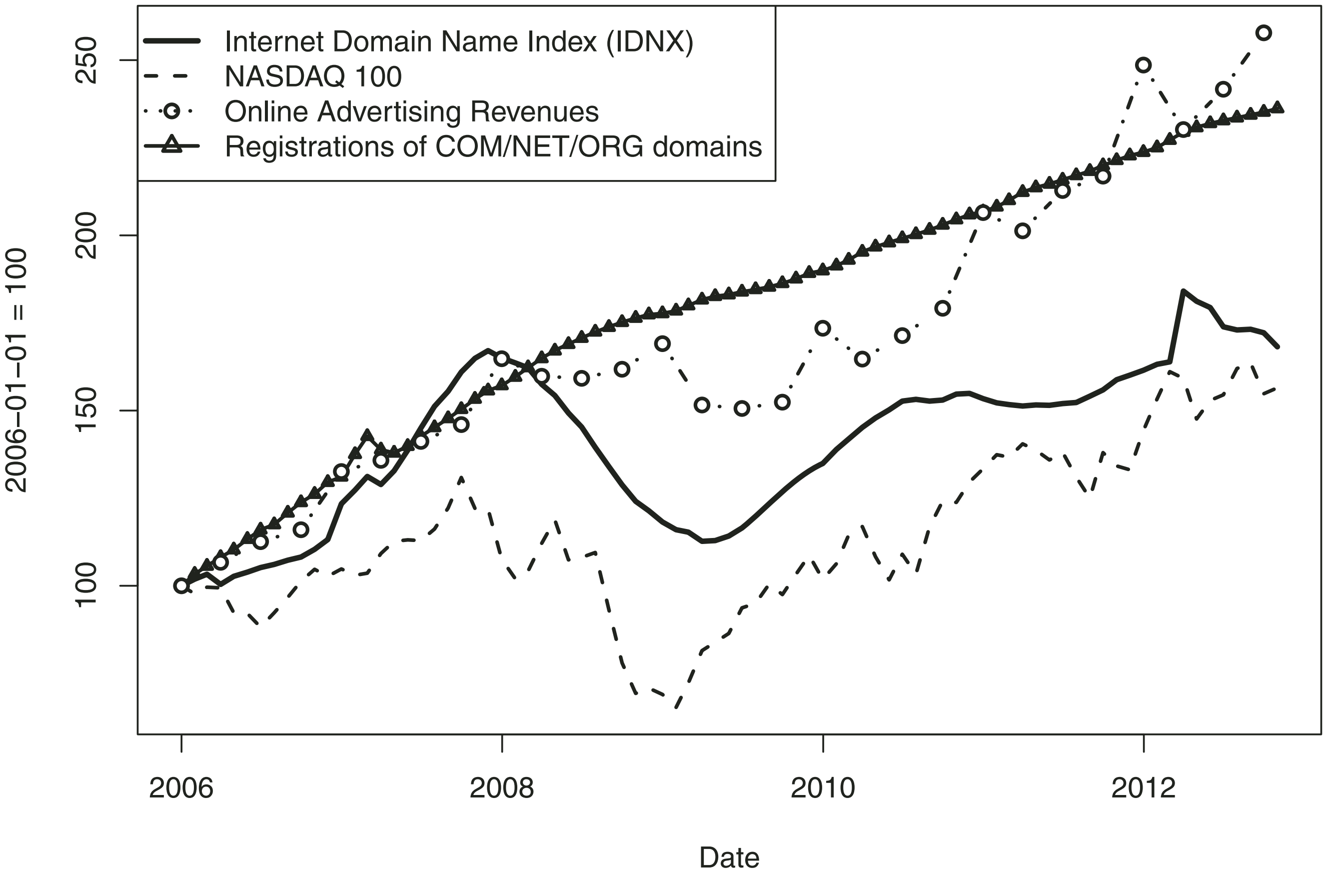

FIG. 6. Price dynamics of Internet domain names compared with NASDAQ-100. Domain prices have an economic foundation. This figure suggests that domain price changes are very similar to changes in the NASDAQ-100 index, which covers the 100 largest technology companies listed on the NASDAQ stock market. When the Internet economy is expanding, domain names as a production factor are in high demand as well. The most important source of income for the majority of Internet enterprises is revenue from online advertisements (time series for online advertisement revenues in the United States are based on the Interactive Advertising Bureau [IAB, 2012]). In 2006–2007, both domain prices and ad revenues move at the same speed. After the dip caused by the global financial crisis, the recovery takes place at similar rates again. However, from 2010/2011 onward, ad revenues outgrow domain prices, which indicates that most domain owners do not benefit from the incoming funds as much as before.

where

$$\Pi = \alpha\beta' = -(I - A_1 - \ldots - A_p) \tag{6}$$

and

$$\Gamma_i = -(A_{i+1} + \ldots + A_p) \tag{7}$$

and $i = 1, \ldots, p-1$. In case the series are cointegrated, the matrix Π will have a reduced rank (Johansen, 1988).

Table 4 features the test statistics and critical values for Johansen's cointegration test.[12] Based on these values, the hypothesis of Π having rank 0 can be rejected at all levels of confidence, whereas a rank of 1 cannot be rejected. In other words, changes in prices and registrations are linked. In addition, the regression coefficients presented in Table 5 indicate that deviations from the long-run equilibrium between primary and secondary markets are corrected by adjustment in registrations, not prices, and that the secondary market responds to changes in overall demand for domains faster than the primary market.

Finally, the empirical data confirm a domain trader's mantra: COM is the most valuable TLD. Table 6 shows pairwise price differences for identical SLDs under the 10 most frequently traded TLDs. The discount for NET is 75%, ORG is worth less than a fifth of COM, BIZ less than a tenth (please refer to Table 6 for all TLD combinatios). In the current econometric setup, the TLD differences are assumed to be constant in time. DE domains trade for 37% less than their COM counterparts. This estimate might be inflated as the language of the SLD is not considered. This could lead to lower estimated discounts for ccTLDs from non-English-speaking countries versus COM whenever the SLD carries some meaning in the local language.

Conclusion

This article estimates the first constant quality price index for Internet domain names. The suggested index provides a

[12]Augmented Dickey–Fuller test statistics with p values <.01 show that the monthly changes y_t are stationary.

DOI: 10.1002/asi

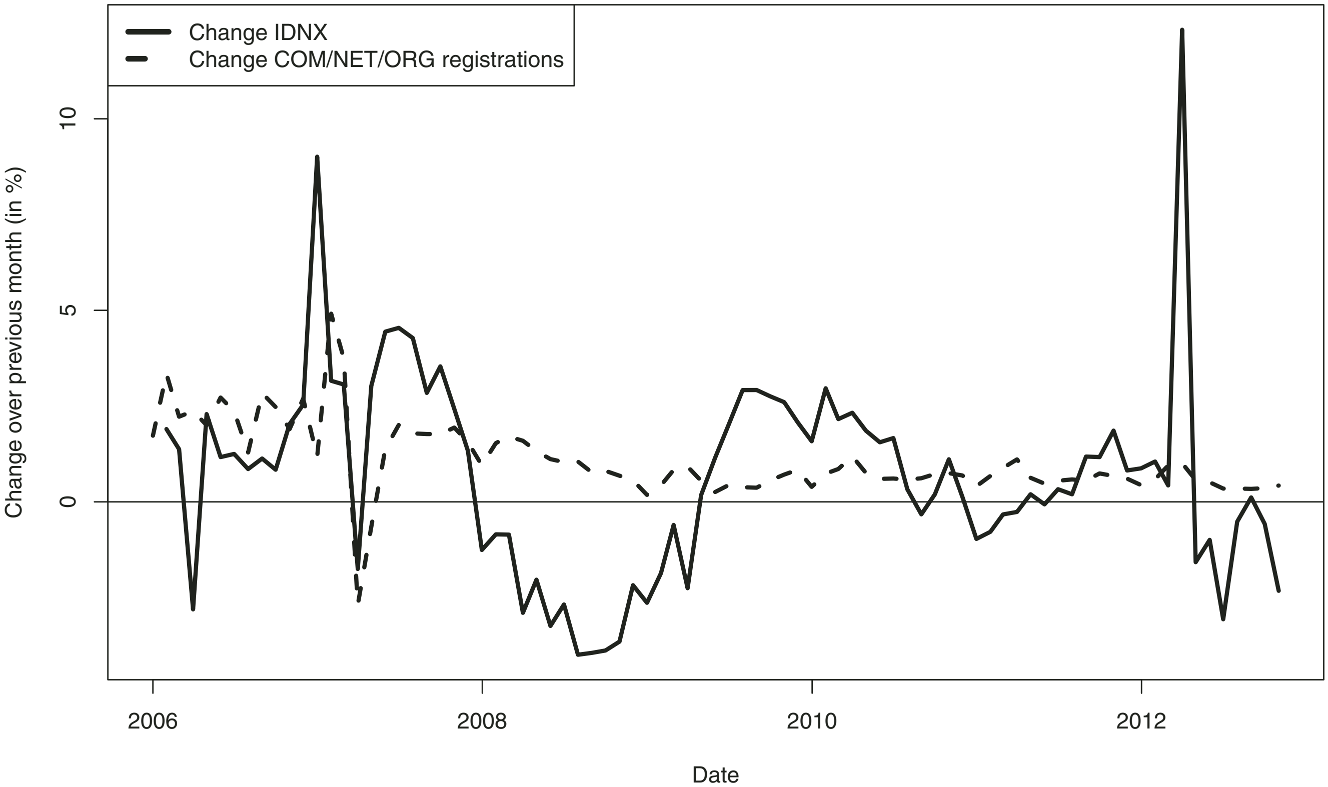

FIG. 7. Cointegration of primary and secondary domain markets. The figure shows the monthly change for resale prices of domain names as estimated by IDNX (solid line) and the growth in total registrations for COM, NET, and ORG domains (dashed line). Overall, the variance in secondary market price growth is higher than the variance for the primary market of registrations. Furthermore, price swings precede registrations, indicating that secondary markets respond faster to changes in general demand for domains than the primary market.

TABLE 4. Values of test statistic and critical values of Johansen's test.

Rank cointegration vector	Test statistic	Critical values 10%	5%	1%
$r <= 1$	2.43	7.52	9.24	12.97
$r = 0$	24.58	13.75	15.67	20.20

Note. The hypothesis of the cointegration matrix having a rank of 0 can be rejected at all levels of confidence, whereas a rank of 1 cannot be rejected. This indicates that both series are cointegrated.

TABLE 5. Regression coefficients for unrestricted VECM.

Variable	$\Delta IDNX_t$	$\Delta Reg.COM/NET/ORG_t$
Constant	0.007	0.006*
Δ Ad. Rev.	–0.389	–0.028
Season 1	–0.059	0.002
Season 2	–0.022	–0.011
Seadon 3	–0.059	–0.003
$\Delta_2 IDNX_{t-1}$	–0.379	0.069**
$\Delta_2 Reg.COM/NET/ORG_{t-1}$	–3.178	–0.605**
$\Delta_2 IDNX_{t-2}$	0.333	0.026
$\Delta_2 Reg.COM/NET/ORG_{t-2}$	–0.549	–0.449*
$\Delta IDNX_{t-1}$	–0.432	0.089**
$\Delta Reg.COM/NET/ORG_{t-1}$	0.016	–0.337***
Adj. R^2	0.12	0.73

Note. Significance levels are indicated by $***p = .01$, $**p = .05$, and $*p = .1$. None of the coefficients for $\Delta IDNX_t$ is statistically significant at any level.

benchmark for domain name traders and investors looking for information on price trends, historical return, and fundamental risk of Internet domain names. It thereby increases transparency in the market for this newly emerged asset class and allows for comparisons with other investments. On average, domain prices grew by 6.6% per year in the last 7 years, exhibiting a boom and bust pattern that closely resembles the path of the overall IT industry. The strong correlations of domain prices with the high-tech economy and online advertisement revenues show that domain name buyers and sellers make economically motivated price decisions. Domain markets are not a cloud-cuckoo-land where dreamers trade esoteric goods at fantasy prices.

From a general asset pricing view, the price for an Internet domain name is the discounted future cash flow that can be generated from this domain. Domain prices are therefore forward looking, giving an indication of not only the current income but including expectations about future opportunities as well. Domain name prices can therefore serve as a fever curve capturing the well-being of Internet companies,

DOI: 10.1002/asi

TABLE 6. Pairwise price comparison of selected TLDs.

	net	org	info	mobi	biz	de	co.uk	es	eu
com	−75	−84	−90	−97	−96	−37	−74	−86	−81
	(−77; −72)	(−86; −81)	(−92; −89)	(−97; −96)	(−97; −95)	(−42; −31)	(−78; −70)	(−90; −81)	(−83; −78)
net		−39	−62	−80	−81	+90	+63	−36	−39
		(−47; −30)	(−67; −56)	(−84; −75)	(−85; −75)	(66;118)	(31;102)	(−55; −9)	(−47; −29)
org			−48	−64	−63	+232	+124	−16	−10
			(−55; −39)	(−72; −54)	(−71; −53)	(179;295)	(73;189)	(−42; 22)	(−24; 6)
info				−32	−48	+342	+532	+60	+26
				(−45; −17)	(−59; −35)	(281;412)	(397;705)	(17;120)	(9;47)
mobi					−14	+223	+467	+73	+95
					(−34; 14)	(154;311)	(320;666)	(23;143)	(51;151)
biz						+529	+643	+149	+227
						(386;715)	(401;1002)	(51;310)	(159;314)
de							+64	−58	−68
							(35;101)	(−68; −45)	(−72; −64)
co.uk								−68	−61
								(−79; −52)	(−69; −51)
es									−29
									(−51; 4)

Note. The matrix shows the estimated price differences (in %) across TLDs based on the regression coefficients from Equation 3; 95% confidence bounds in parenthesis. A NET domain, for instance, is estimated to be 75% more affordable than the COM equivalent, controlling for time trends and the SLD. In the current econometric setup, the TLD differences are assumed to be constant in time. Furthermore, the language of the SLD is not considered. This might lead to lower estimated discounts for ccTLDs from non-English-speaking countries versus COM whenever the SLD carries some meaning in the local language.

online media providers, and start-ups that base their business model on income from selling advertisement space. This sheds light on the prospects of small- and medium-sized online enterprises that are currently excluded by traditional stock-price indices.

Comparing the dynamics of the estimated index with fundamental economic variables suggests a regime shift to a new equilibrium between domain prices and advertising revenues after 2010. After a recovery from the financial crisis, domain prices did not grow as fast as advertising spending. This shift is possibly caused by changes in the user flows caused by the growing importance of search engines and their algorithms. Ieong et al. (2012), for instance, find that search engines channel the growing number of searches to relatively fewer domains: In 2010, the share of search results linking to the top 30 domains increased to 38%, which is already 1.5 times higher than 2009 values. The concentration in traffic leads to a funneling of advertising revenues on fewer domains, putting price pressure on the majority of domains. In addition, the exclusion of parked domains from search results (Singhal & Cutts, 2011) put a dent in revenues from owning domains.

A cointegration analysis of primary and secondary domain markets reveals that domain registrations and prices are linked. Price changes observed in the secondary markets and registrations both depend on changes in the general demand for domain names. Unlike the market for "real" land, no zoning or other formal regulations restrict the creation of new locations on the Internet, leaving availability of domains the only limit. The results show that this limitation is actually binding. If supply was unconstrained, resale prices for domains would not deviate far from registration fees, and no statistical link between the two markets would be detectable.

In addition, resale prices respond faster to changes in the underlying demand than registrations, indicating that traders incorporate new information on domain demand more rapidly than registrants.

Finally, the land-domain name analogy on which this article rests can be extended from individual domains to agglomerations of domains. TLDs such as COM or NET are like cities in virtual space. As for "real" metropolitan areas, demand for space determines both the size of the virtual agglomeration measured in total number of domains registered (Figure 1) and, simultaneously, the differences in prices paid for space in the centers of these "cities" (Table 6). Higher demand for a specific extension results in both more "sprawl" and new initial registrations, and in higher prices paid for locations within this agglomeration. This warrants a word of caution about the upcoming increase in available extensions scheduled for 2013 and 2014 by the Internet's governing body, ICANN. It will require millions of "inhabitants" for new online locations to support price levels comparable with existing TLDs. Aspiring virtual land barons, beware!

References

ACTA (1999). Anticybersquatting consumer protection act. Retrieved from http://www.gpo.gov/fdsys/pkg/BILLS-106s1255is/pdf/BILLS-106s1255is.pdf.

Alonso, W. (1964). Location and land use: Toward a general theory of land rent. Harvard University Press.

Bailey, M., Muth, R., & Nourse, H. (1963). A regression method for real estate price index construction. Journal of the American Statistical Association, 58(304), 933–942.

Banerjee, A., Rahman, S., & Faloutsos, M. (2011). Sut: Quantifying and mitigating url typosquatting. Computer Networks, 55, 3001–3014.

Bokhari, S., & Geltner, D. (2010). Estimating real estate price movements for high frequency tradable indexes in a scarce data environment. The Journal of Real Estate Finance and Economics, 1–22.

Bryan, T., & Colwell, P. (1982). Housing price indices. In C.F. Sirmans (Ed.), Research in real estate (Vol. 2). Greenwich: JAI Press.

Bulow, J., & Klemperer, P. (1996). Auctions versus negotiations. The American Economic Review, 86(1), 180–194.

Burshtein, S. (2005). Is a domain name property? Journal of Intellectual Property Law and Practice, 1(1), 59.

Case, K.E., & Shiller, R.J. (1987). Prices of single-family homes since 1970: New indexes for four cities. New England Economic Review, 45–56.

Denic (2011). Domainzahlenvergleich international. Retrieved from http://www.denic.de/hintergrund/statistiken/internationale-domainstatistik.html.

Edelman, B., & Moore, T. (2010). Measuring the perpetrators and funders of typosquatting. In 14th International Conference on Financial Cryptography and Data Security. Tenerife, Spain.

Evans, D.S. (2009). The online advertising industry: Economics, evolution, and privacy. The Journal of Economic Perspectives, 23(3), 37–60.

Francke, M. (2010). Repeat sales index for thin markets: A structural time series approach. The Journal of Real Estate Finance and Economics, 41(1), 24–52.

Gatzlaff, D.H., & Haurin, D.R. (1997). Sample selection bias and repeat-sales index estimates. The Journal of Real Estate Finance and Economics, 14, 33–50. doi:10.1023/A:1007763816289

Geltner, D. (1997). Bias and precision of estimates of housing investment risk based on repeat-sales indices: A simulation analysis. The Journal of Real Estate Finance and Economics, 14(1), 155–171.

Goetzmann, W. (1993). Accounting for taste: Art and the financial markets over three centuries. The American Economic Review, 83(5), 1370–1376.

Interactive Advertising Bureau. (2012). IAB Internet Advertising Revenue Report. Retrieved from http://www.iab.net/media/file/IAB-HY-2011-Report-Final.pdf

ICANN Internet Corporation for Assigned Names and Numbers. (2011). Registry Operator's Monthly Report.COM/.NET January 2011. Retrieved from http://www.icann.org/en/tlds/monthly-reports/

ICANN Internet Corporation for Assigned Names and Numbers. (2012a). New genereic top level domains. Retrieved from http://newgtlds.icann.org/en/

ICANN Internet Corporation for Assigned Names and Numbers. (2012b). Uniform domain-name dispute-resolution policy. Retrieved from http://www.icann.org/en/help/dndr/udrp

Ieong, S., Mishra, N., Sadikov, E., & Zhang, L. (2012, February). Domain bias in web search. In Proceedings of the fifth ACM international conference on Web search and data mining (pp. 413–422). ACM.

Johansen, S. (1988). Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control, 12(2), 231–254.

Lindenthal, T. (2011). On rushes and riches: The "wild west" era for Internet domain names is over as efficient markets for this "virtual land" have emerged. Scientific American. Retrieved from http://blogs.scientificamerican.com/guest-blog/2011/08/24/

Mills, E. (1972). Studies in the structure of the urban economy. Baltimore, MD: The Johns Hopkins Press.

Moore, T., & Edelman, B. (2010). Measuring the perpetrators and funders of typosquatting. In R. Sion (Ed.), Financial Cryptography and Data Security (pp. 175–191). Berlin: Springer.

Muth, T. (1969). Cities and housing: The spatial pattern of urban residential land use. Chicago: University of Chicago Press.

Pfaff, B. (2008). Var, svar and svec models: Implementation within r package vars. Journal of Statistical Software, 27(4), 1–32.

Sedo.com (2011). Market trends. Retrieved from http://www.sedo.com/uk/resources/market-trends/

Shiller, R.J. (1993). Measuring asset values for cash settlement in derivative markets: Hedonic repeated measures indices and perpetual futures. The Journal of Finance, 48(3), 911–931.

Singhal, A., & Cutts, M. (2011). Finding more high-quality sites in search. Google Official Blog. Retrieved from http://googleblog.blogspot.de/2011/02/finding-more-high-quality-sites-in.html

Verisign. (2012). The domain name industry brief, 9(3). Retrieved from http://www.verisigninc.com/assets/domain-name-brief-oct2012.pdf.

DOI: 10.1002/asi