the demographics of the prospective buyer, the buyer’s location, the buyer’s previous history with the seller or the seller’s partners, etc.

commerce systems are not necessarily aware of the products and services offered to other potential buyers, or their prices (as they typically would be in a brick-and-mortar store).[10]

The ecosystem that supports surveillance pricing has been enabled by the convergence of multipl C

1.) Data is routinely collected at a volume that is well beyond what was possible previously.[17] Moreover, data collection is not confined to a single technology or industry but is collected by numerous sources across industries.[18]
2. (Data is generated by many interactions between a seller and potential buyer, with numerous online services embedding multiple trackers.[19] Such data can be used both to "segment," or group individuals based on shared interests or traits, and to enable targeting online down to the individual.[20]
3. Data can be collected continually, and potential buyers are often unaware of the data that has been collected and often have not explicitly consented to that data collection.[21]

This section covers how academic researchers and practitioners have approached understanding these foundations, focusing first on the *inputs* (*i.e.*, the data that is collected) before turning to the *outputs* (*i.e.*, where the effect may be observed).

3.1 Surveillance pricing inputs

Sellers across industries have realized that knowing more about buyers—their preferences and how they use different products or services—can provide additional avenues for maximizing profits. Early examples of technologies for gathering such information on consumers include "auditmeters"[22] which recorded the radio stations to which homes were listening. Later examples included more passively collected data, such as the newspapers and magazines to which consumers subscribed,[23] the products that buyers purchased at grocery stores,[24] and the movies that consumers rented from video rental stores.[25] At the same time, credit reporting agencies and other financial firms[26] began to use data they collected to offer data analytics and targeting information to advertisers. Today, sellers and various third parties collect large amounts of data on consumers—including their purchase histories, their physical movements, their communications, their

[17] Cong. Research Serv., *Online Consumer Data Collection and Data Privacy* (No. R47298, Oct. 31, 2022), https://crsreports.congress.gov/product/pdf/R/R47298.

[18] *See* Michael Kwet, *In Stores, Secret Surveillance Tracks Your Every Move*, N.Y. Times, June 14, 2019, https://www.nytimes.com/interactive/2019/06/14/opinion/bluetooth-wireless-tracking-privacy.html; R.J. Cross, *How Mastercard sells its 'gold mine' of transaction data*, U.S. PIRG Education Fund (Sept. 23, 2023), https://pirg.org/edfund/resources/how-mastercard-sells-data/; Kashmir Hill, *Automakers Are Sharing Consumers' Driving Behavior with Insurance Companies*, N.Y. Times, Mar. 11, 2024, https://www.nytimes.com/2024/03/11/technology/carmakers-driver-tracking-insurance.html.

[19] Steven Englehardt & Arvind Narayanan, *Online Tracking: A 1-million-site Measurement and Analysis*, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security 138

Retailers are now directly monetizing the data they collect, with some even0 o v that circumvent these controls. Commonly called “browser fingerprinting,”[44] these approaches identify unique combinations of characteristics of the user’s device, including the web browser and extensions, screen resolution, local fonts, battery information, and hardware and software implementation of different browser APIs.[45] Recent studies have identified browser fingerprinting code on more than 25% of top websites.[46]

: Firms that provide internet access (commonly called “internet service providers”) can monitor the traffic that customers send to collect information on consumers. While much web traffic is encrypted today, requests such as domain name system (“DNS”) queries[47] and the setup of transport layer security (“TLS”) connections (via Server Name Indication[48]) can reveal the domains to which consumers are connecting even if the traffic itself is encrypted. Even this limited view into traffic can be used to develop profiles of customers.

3.1.3 Data from mobile devices

The widespread popularity of mobile devices, including phones and tablets, has provided additional ᵀ ! ḥ ȯ opportunities for collecting data on users.

A distinguishing characteristic of mobile devices is that they allow third-party developers to run applications (“mobile apps”) on users’ devices. Commonly distributed through “app stores” run by the mobile operating system provider, s.

3.1.5 Data from data brokers or other intermediaries

There are also companies whose primary business model is collecting, aggregating, sharing, and reselling consumers' personal data.[71] Frequently called "data brokers," these companies have existed since well before the advent of the internet.[72] Today, they collect data from a wide variety of sources—including government, commercial, and private records—and use it for a variety of purposes including identity verification, advertisement targeting, fraud detection, and credit estimation. Little direct study of data brokers exists, largely because they offer few user-facing tools that researchers can study; much m m hl ȯ am

common sets of participants who have already been “on boarded” and can choose to participate in different studies. Models for such an approach exist in other domains[130] and could be tailored to understand the impacts of surveillance pricing as well.

: : Sellers agreeing to use the same algorithm to determine prices may lead to reduced rivalry and/or the ability to increase prices,[131] but the limits of when the use of common *data* to independently build algorithms might result in similar effects are not yet known. This is especially relevant as the behavior of modern AI systems is often heavily driven by the data they are trained on. In other words, what happens if multiple sellers use the same data to develop separate algorithms for deciding prices, and those algorithms all end up with substantially similar behavior?

Additionally, more research is needed to understand the extent to which use of surveillance pricing strategies commonly involves or promotes the use of either common pricing algorithms and/or common data sets, across competing firms. As discussed above, sellers in the surveillance pricing ecosystem can amass granular customer data from a variety of common sources, presumably without resorting to potentially anticompetitive information exchanges. Existing studies on the competitive impact of algorithmic pricing suggest that its adoption may lead to supra-competitive pricing, even without agreements to use common algorithms or data; but this body of research is in its infancy. As sellers become increasingly sophisticated in how they gather customer information and deploy pricing algorithms in the surveillance pricing ecosystem, additional research is needed into the impacts of autonomous pricing algorithms on pricing and the competitive process.

[130] *Helping researchers understand online behavior.*, National Internet Observatory. https://nationalinternetobservatory.org/ (last visited Jan. 9, 2025).

[131] Hannah Garden-Monheit & Ken Merber. *Price fixing by algorithm is still price fixing,* FTC Business Blog (Mar. 1, 2024), https://www.ftc.gov/business-guidance/blog/2024/03/price-fixing-algorithm-still-price-fixing.