BEA Working Paper Series, WP2024-1a

Capitalizing Data: Case Studies of Driving Records and Vehicle Insurance Claims

Author Rachel Soloveichik, U.S. Bureau of Economic Analysis

Contact Rachel.Soloveichik@bea.gov

Date April 2024

Abstract Early papers on capitalizing data focused on complex digital data that are stored on supercomputers and managed by highly skilled computer scientists (Statistics Canada 2019) (Eurostat 2020) (Coyle 2022) (Calderon and Rassier 2022) (Mitchell et al. 2022). This paper studies two very different types of data: driving records and vehicle insurance claims. Like the tax forms and individual credit reports studied in a previous paper (Soloveichik 2023), these types of data are simple text records that can be stored on paper (U.S. Women's Bureau 1944) and managed by relatively unskilled workers (U.S. Bureau of Labor Statistics 2022). Despite their simplicity, these two data types are expensive to create. This paper estimates that driving records had a creation cost of $229 billion in 2017 and vehicle insurance claims had a creation cost of $153 billion in 2017.

Growth during the early 1990s changes noticeably when vehicle insurance claims are tracked as long-lived intangible assets. Insurers started sharing their claims data via the Comprehensive Loss Underwriting Exchange database in 1992. This new data sharing increased the quality of vehicle insurance claims data substantially and therefore increased real GDP growth between 1989 and 1994 by 0.2 percentage point per year. Furthermore, the real capital deepening associated with real data growth in the early 1990s explains 0.2 percentage point of the 0.4 percentage point acceleration to measured productivity in the late 1990s.

Keywords Data, driving, vehicle, insurance

JEL codes D14, E01, and G14

Bureau of Economic Analysis
U.S. DEPARTMENT OF COMMERCE

The views expressed in this paper are those of the authors and do not necessarily represent the U.S. Bureau of Economic Analysis, or the U.S. Department of Commerce.

Introduction

This paper studies two types of driving data: driving records and vehicle insurance claims. Driving records report underlying risk factors like when an individual obtained their driving license, whether a vehicle is currently licensed, whether a driver has citations for safety issues like speeding or drunk driving, and whether their smartphone detects hard stops. In contrast, vehicle insurance claims report outcomes like accidents or theft. Driving records are created by governments, individuals, lawyers, insurance companies, and sometime employers, while vehicle insurance claims are only created by insurance companies. Even though these two types of data are separate, they are often used together in production. For example, a rideshare company might use both driving records and vehicle insurance claims, along with other information like vehicle type and non-driving criminal record (Agan et al. 2022), to decide whether an individual is eligible to drive for the rideshare company.

In recent years, national accountants have discussed including both long-lived purchased and long-lived own-account data in gross domestic product (GDP) as intangible capital assets (Rassier et al. 2019) (Eurostat 2020). This inclusion is a broadening of the current recommendation to only include long-lived purchased data in GDP (United Nations Statistics 2008, sec. 10.112-10.114). Later sections of this paper show that driving data are long-lived enough to be tracked as intangible assets in the national accounts. This paper includes government driving data and private driving data (both digital and non-digital) in GDP using a recently developed national accounting framework for data investment (Eurostat 2020).

At first glance, one might dismiss the driving data studied in this paper as valueless data exhaust. It is certainly true that neither driving records nor vehicle insurance claims are the primary product of their respective industries. But measured GDP includes many other products that are not the primary product of their industries. For example, cow hides are included in GDP even though they are almost always a byproduct of beef production rather than a primary product. Similarly, one might argue that neither the driving records nor the vehicle insurance claims studied in this paper count as "data" because neither are collected by data specialists. But measured GDP includes many other products that are produced by non-specialists. For example, restaurant meals cooked by waiters who are filling in for absent chefs are included in GDP together with restaurant meals cooked by trained chefs. Accordingly, this paper's inclusion of driving data in the broad category "data capital" is consistent with general national accounting principles even if it is not consistent with some papers' narrow focus on data that are collected by data specialists as a primary product.

Tracking data as capital raises measured GDP because both private business investment and government consumption of fixed capital are components of final output. In addition, personal

consumption expenditures (PCE) by category change when purchased consumer data are tracked separately from purchases of other insurance industry output. Finally, household production changes when consumer data creation and consumer durable services are included in the framework outlined in the U.S. Bureau of Economic Analysis' (BEA's) existing household production satellite account (Bridgman et al. 2022). This paper presents calculations that tracking these two types of data as intangible assets increases measured GDP in 2017 by $128 billion and measured household production in 2017 by $160 billion. This GDP revision is smaller than the $383 billion GDP increase associated with private software investment or the $438 billion GDP increase associated with private R&D—but it is not negligible.

Real creation of vehicle insurance claims data grew rapidly during the early 1990s. Before the Comprehensive Loss and Underwriting Exchange (CLUE) was founded in 1992, claims data were kept in-house and only used by one insurance company to price vehicle policies. After CLUE was founded, claims data were shared broadly and used by multiple insurance companies, vehicle dealers (Huang et al. 2021), rideshare platforms (Boyle 2019), and employers (U.S. Department of Transportation 2023). In other words, insurance claims data quality rose enormously when they switched from an asset with only one user to an asset with many users. Furthermore, the stock of insurance claims data switched from being owned by insurance companies to being owned by the driver described. Finally, claims data services switched from being business asset services that are implicitly included in insurance industry output to being consumer durable services that can be explicitly included in household production (Bridgman et al. 2022). This paper demonstrates that this switch has important implications for measured GDP, measured productivity, and measured household production.

This paper is divided into six sections. Section 1 describes the current and historical system in which driving data are created, distributed, and used. Section 2 uses BEA's published spending estimates, arrest counts from the Federal Bureau of Investigation (FBI) Uniform Crime Report, and academic research to value the driving records created each year. Section 3 uses industry literature and expert opinion to value the vehicle insurance claims created each year. Section 4 presents revisions of nominal GDP and nominal household production when these two types of data are included as intangible capital assets. Section 5 presents price indexes for data and then uses those price indexes to recalculate overall GDP prices, overall household production prices, real GDP, and real household production. Finally, Section 6 presents revisions of private business productivity when both driving records and vehicle insurance claims are included in the production accounts.

1. System of Data Creation, Sharing, and Usage

There are four main entities in the shared data system. First, there is a data creator that records risk data and submits it to a data platform. Second, there is a data platform that organizes, audits, and distributes data to a narrow list of specific users for a narrow list of permitted reasons. Third, there are the data owners who distributed their data to whomever they please for whatever reason they please. Finally, there is the data user. This section discusses each of the four main entities briefly.

Shared Data Creators

Driving records are created jointly by drivers, police officers, courts, and lawyers. To start out, both drivers and vehicles require licenses before they are allowed on the road (Sheldon 2021) (Homer 2022). Once on the road, police officers observe vehicles and decide whether to initiate a traffic stop. During the traffic stop, the police officer and driver talk, and the police officer decides whether to write a ticket for a minor safety violation like speeding and whether to make an arrest for a major safety violation like drunk driving. After a driver is ticketed or arrested, they then face a decision. They can plead guilty immediately or go to court to contest the ticket or arrest. If they choose to go to court, they may represent themselves, use a court-appointed lawyer, or hire their own lawyer. Finally, the court decides whether to convict or acquit the accused.

Driving records are imperfect measures of safety. Some drivers are denied licenses due to non-safety factors like unpaid child support (Congressional Research Service 2011) or immigration status (Johnson 2004) (Amuedo-Dorantes et al. 2020). Similarly, some vehicles are denied licenses due to non-safety factors like missing import records (U.S. Customs and Border Protection 2023). Conversely, many unsafe drivers escape detection because enforcement resources are limited (Becker 1968), because some police officers choose to be lenient (Goncalves and Mello 2023), or because some vulnerable groups are disproportionately subject to traffic enforcement (Aggarwal et al. 2022) (Lofstrom et al. 2022). Innocent people sometimes plead guilty because contesting a ticket or arrest may be too time consuming or too stressful (Rakoff 2021), and guilty people sometimes have their citations dismissed because the police officer does not show up in court (Desind 2015) or for other reasons. Nevertheless, meta-analyses find that negative driving records are correlated with both general accident risk (Barraclough et al 2016) and fatal accident risk (Slikboer et al. 2020). Because of these correlations, individuals without a valid U.S. driving license have difficulty getting vehicle insurance (Brady and Glover 2023), and insurers routinely charge higher premiums to drivers with bad driving records (Chiappori and Salanie 2000).

Vehicle claims data are created by insurers after an accident or theft is reported. Insurance companies generally differentiate between claims where the insured is at fault and claims where the insured is not

at fault (Shuldiner 1968) (ChoicePoint Asset Company 2001). As a result, the exact same payment may have different implications for future premiums depending on who is judged to be at fault. Conversely, insurance companies often record a claim even if the assessed damage is below the deductible and therefore no money is paid out (Mancusi 2018) and may even record a claim when a client asks about the claim-filing process but does not actually file a claim (Marte 2014).

Insurance customers know that shared vehicle claims data are bundled with other insurance services. To start out, both insurance applicants and individuals who apply for driving jobs are generally asked specific questions about their claims data (Megna and Kasperowicz 2023) (Rogers 2023). In addition, popular financial advisors tell Americans that filing a claim raises future insurance premiums (Shuldiner 1968) (Quinn 2009) and may reduce the likelihood of being hired for jobs that involve driving (Boyle 2019). Based on these facts, shared vehicle claims data are clearly not an externality under the official guidelines for national accounting (United Nations 2008, sec. 3.92). This paper tracks shared vehicle claims as purchased data creation services that are bundled together with other insurance services. Tracking shared data creation does not change total output for insurers. Instead, insurance output is split into two components: data creation and other insurance services.

Vehicle claims data are sometimes inaccurate. Minor accidents like parking lot scratches are often handled unofficially without either party reporting the accident to their insurance company (Vallet and McCormick 2023). Individuals involved in an accident often misreport key facts in the hope of getting a larger payout (Crocker and Tennyson 2002). Conversely, insurance companies sometimes may misinterpret their policies to deny potentially legitimate claims (Richmond 1994). Nevertheless, the same meta-analyses cited earlier to demonstrate that driving records are correlated with accident risk also find that prior vehicle insurance claims are correlated with accident risk (Barraclough et al 2016) (Slikboer et al. 2020). This paper calculates a value for the sometimes inaccurate claims data that do exist rather than a hypothetical value for the perfectly accurate claims data that do not exist.

Shared Data Platforms

Data platforms have four data jobs. First, they organize the scattered data supplied by separate data creators into a combined file. Second, they audit the combined file for problems that must be resolved. Third, they store the combined file for future reference. Finally, they distribute the combined file to any entity with a valid request. In addition to these four data jobs, data platforms sometimes sell additional services. These additional services may be important sources of revenue—but they are not directly related to the driving data system.

Data platforms are small relative to the value of the data they hold. This paper estimates that the total budget of all departments of motor vehicles in the United States was only $16 billion in 2017. Similarly, the global RelX group, which owns CLUE and other data systems, earned only £1.6 billion from their North American Risk & Business Analytics division in 2017 (RelX 2017). Their small size is possible because data supplied to data platforms are already formatted by data creators (American Association of Motor Vehicle Administrators 2023). In addition, data creators are responsible for handling disputes about potentially inaccurate information (Kasperowicz and Ohman 2022). Because data platforms are so small and because data creation accounts for only a portion of their output, even a substantial change to their output has little impact on this paper's measures of data creation.

Shared Data Owners

Businesses and individuals can generally distribute their data to any entity they choose. In some cases, the business or individual distributes the data directly. For example, used-car sellers might show a potential buyer a Carfax report (Huang et al. 2021). In other cases, the business or individual authorizes a platform to distribute data to a specified user. Direct distribution and authorized distribution have similar economic impacts, so this paper does not distinguish between them. Because businesses and individuals have so much control over their data, this paper treats them as the primary data owner.

Governments also distribute data to specific entities for specific uses. Examples of permitted driving-record distribution include traffic stop investigations (LaFave 2004) and racial bias research (Lofstrom et al. 2022). Examples of permitted insurance claims data distribution include collection of Medicare liens for health care related to vehicle accidents (Helland and Kipperman 2011) and state insurance regulation (Feinman 2015). Because governments have some control over data, this paper treats them as partial data owners.

Partial ownership by multiple parties likely yields higher social welfare than either complete secrecy or full sharing. On the one hand, driving data are used for many purposes and total output is higher when data are distributed broadly. On the other hand, businesses and individuals value privacy and prefer limited distribution of data (Acquisti et al. 2016). Each business and each individual have their own desired level of privacy, and so it would be difficult for the government to set a data distribution level that works for everyone. Governments resolve these contradictory preferences by distributing a minimum amount of data for purposes that are believed to be important enough to override privacy preferences. Businesses and individuals who desire less than the maximum amount of privacy can then decide to distribute their data to any entity they choose.

"Free" distribution of data likely yields higher output than either selling data or keeping it in-house. Data are nonrival and complementary to other data, and so their contribution to output is maximized when they are shared broadly (Coyle 2022). Furthermore, broad sharing of data requires marginal prices for data that may be too low to cover the cost of data creation (Coyle and Diepeveen 2022). In this respect, data are similar to other products whose creation involves high fixed costs and low marginal costs. Previous researchers have shown that government provision is one method to finance these products (Samuelson 1954). That method is partially used to finance the creation of driving-record data. But recent papers have shown that individual ownership is an alternative financing method that can increase welfare (Jones and Tonetti 2020). That method is partially used to finance the creation of driving-record data and almost entirely used to finance the creation of vehicle insurance claims data. In particular, an individual can finance the creation of their own driving data, give those data for "free" to any entity with a valid reason for its request, and then benefit when that entity uses those data to provide services or employment to the data owner.

Shared Data Users

Governments use driving records constantly. Courts use driving records to determine whether an individual should be fined for minor safety violations or imprisoned for major safety violations (Gershowitz 2011). Departments of motor vehicles use driving records to determine whether an individual's driving license should be suspended (Johnston 1951). Finally, policymakers use both driving records and vehicle insurance claims data to measure current road safety and suggest ways to improve road safety in the future (Carey et al. 2020). This paper neither studies how each government agency uses data in its operations nor studies whether government agencies could improve efficiency by using data better (Harrington et al. 2023). Instead, it simply accepts that driving records are an important production input. This paper assumes that the government share of data usage equals the government share of data creation funding. This assumption is equivalent to treating driving records data created by governments as intangible assets owned by governments.

Private businesses use driving data to make useful predictions (Farboodi and Veldkamp 2022). Insurers use driving data to predict the likelihood of future insurance claims (Barraclough et al 2016) (Slikboer et al. 2020), and car buyers use driving data to predict used-car quality (Chapman 2022). Employers are required to check driving data for workers who regularly drive commercial vehicles on the job (U.S. Department of Transportation 2023), and many employers also check driving data for other workers who drive on the job (Rogers 2023). In addition, some employers use drunk-driving records, together with non-driving criminal records, in their hiring decisions (Agan et al. 2022). Because driving data are used for fewer purposes than the financial data studied in a previous paper (Soloveichik 2023), the

revisions to overall GDP shown in this paper are smaller than the revisions to overall GDP shown in that paper. Nevertheless, driving data are an important production input, and tracking driving data changes measured productivity for the insurance industry significantly and changes measured productivity for other industries noticeably.

This paper assumes businesses that use driving data implicitly purchase those data from their owner. These implicit data purchases are part of a bundled transaction. For example, a motor vehicle insurer might offer special discounts to customers who share their insurance claims history, an employer might offer all job applicants a regular position but also offer those job applicants who consent to have their driving record checked an opportunity to apply for highly paid traveling sales positions, or an equipment supplier might offer discounts on truck leases if a small-business owner shares their personal vehicle insurance claims history. This paper is focused on tracking data, and so it adjusts all bundled transactions to explicitly track both the value of data given to firms and the value that firms give in return for those data. For customer data, this paper adjusts the value of purchased products upward and business intermediate inputs upward to include the value of data provided by customers. In the example given earlier, the insurance company's sales would be adjusted upward to include value of the special discounts given to customers who share their insurance claims history and the insurance company's intermediate inputs would be adjusted upward to include the value of data provided by customers who share their claims history. These two adjustments add precisely the same amount to output and intermediate inputs so that tracking customer data does not change measured business value added. Similarly, this paper shifts the value of data provided by workers from labor compensation to business intermediate inputs without changing measured gross operating surplus. In the example given earlier, the value of the opportunity to apply for traveling sales positions is considered an implicit payment for the data provided by job applicants. This implicit payment is shifted from labor compensation to a payment for data provided by job applicants who consent to have their driving history checked. Finally, this paper shifts the value of data provided by self-employed individuals from general proprietors' income to capital services from data without changing either measured business value added or measured gross operating surplus. In the example given earlier, the discount on a truck lease is considered an implicit payment for the data provided by the small-business owner. This implicit payment is shifted from general proprietors' income to a payment for the value of data provided by the small-business owner. To be clear, all of these treatments only apply to data usage. The implicit payments for data creation by insurers and other businesses were described in an earlier section.

This paper does not study vehicle ownership, driver risk taking, or other indirect behavior changes that may be impacted by driving data. It is very common for individuals to consider their future insurance costs when buying a vehicle. Forward-looking individuals may also consider the expected damage to

their driving record from speeding or driving drunk. Conversely, governments and insurers may structure driving-data creation to influence current and future behavior. The combined impact of these efforts is theoretically ambiguous and difficult to solve without a complete model of the entire economy. For now, this paper only focuses on the direct impact of driving data.

Unshared Data: Driving Behavior and Vehicle Insurance Claims Before CLUE

The driving-behavior data system is much simpler than the system for shared data. Insurers currently keep the driving-behavior data they collect in-house and use those data for their own purposes (Jin and Vasserman 2021). Similarly, employers keep the driving-behavior data they collect in-house and use those data for their own purposes (Overman 2014). Before CLUE started, vehicle claims were handled similarly to driving-behavior data. The insurance company kept claims data in-house and used those data for their own purposes. Unlike shared data, which are owned by the entity described, in-house data are owned by the entity which holds them. In other words, the four main entities in the shared data system are all combined in the unshared data system.

Because neither current driving-behavior data nor historical vehicle claims data are shared, they are not implicitly bundled with other insurance services. Accordingly, including those data in measured GDP increases measured output for the insurance industry and employers without any adjustment to their primary output. Similarly, including data services in the production accounts increases measured capital services for the insurance industry and employers without any adjustment to either output or other inputs. This paper's treatment of these in-house data is similar to other national accounting papers' treatment of other types of in-house data (Coyle and Li 2021) (Mitchell et al. 2022).

Why Did Companies Start Sharing Claims Data?

One might think that data sharing benefits all insurance companies because each insurance company benefits from access to more claims data. In other words, data are more valuable when they are combined with other data (Coyle 2022). But the CLUE database is regulated by the 1970 Fair Credit Reporting Act and individuals can get a copy of their CLUE report without charge (Girardin 2023). So, insurance companies need not send claims data for their existing customers to access claims data for new customers. Another possible explanation might be collusion. It is true that an antitrust literature has argued that data sharing can facilitate collusion when market conditions are uncertain (Vives 1984) (Raith 1996). However, this particular mechanism does not apply to the insurance industry because regulators generally allow cooperative insurance rate setting even without data sharing (Eisenach 1985). The remainder of this subsection will suggest alternative explanations for the existence of CLUE.

Data sharing may have benefited insurance companies by reducing customer churn. Expected accident risk and therefore expected payouts are higher for individuals who have had accidents in the past (Rubinstein 2023). When claims data are kept in-house , individuals with a recent accident often switch insurers to avoid a premium increase (Cohen 2005). But when claims data are shared, competing insurance companies quote a higher premium to individuals who have had an at-fault accident and therefore decrease the incentive for those individuals to switch insurers. Of course, competing insurance companies also quote a lower premium to individuals who have not had an at-fault accident and therefore increase the incentive for those individuals to switch insurers. If individuals are loss averse, then they will be more interested in switching carriers after a premium increase (Hoffman 2013)—and therefore the net impact of data sharing on customer churn is negative.

It is also possible that the CLUE database benefited insurers by deterring claims. It may be true that the CLUE database is not widely known (Marte 2014), and therefore it is unlikely that ordinary customers changed their claim filing behavior when their insurer joined the CLUE database. However, an unknown but noticeable share of insurance claims are fraudulent (Brenton 1964) and some are even filed by professionals who specialize in fraudulent claims (Emerson 1992). Those professionals may have quickly learned about the CLUE and quickly changed their behavior in response.

In practice, the exact reason for the existence of CLUE does not matter to this paper. Instead, this paper focuses on how CLUE impacts measured GDP and measured productivity. These empirical results are the same regardless of why insurance companies chose to join CLUE in the early 1990s.

2. Valuing Driving-Record Creation

This paper focuses on data creation costs alone. Accordingly, it measures the cost of creating driving records rather than the cost of actually learning how to drive. For example, time spent filling out paperwork at the department of motor vehicles is included but not time spent practicing driving. Measured investment would be much larger if human capital investment was included together with data creation, but the qualitative trends shown in this paper would be similar. This paper also does not study fines, fees, and other payments to the government associated with driving data. Measured investment would be slightly larger if those payments were tracked as a tax on driving data, but the qualitative trends shown in this paper would once again be similar.

This paper focuses on driving data, and so it only studies records related to driving. But the general ideas developed in this section could apply to records related to other activities. For example, a business

license could be seen as a data asset owned by a small business. Or a civil citation for loud parties could be seen as a data asset owned by the rental housing unit. The national accounting impact of those other records is a topic for future research.

Licenses

Government spending on licenses is assumed to equal the total budgets of all the departments of motor vehicles in the United States. Each state has its own department of motor vehicles and so government spending is calculated by summing spending in each state. Information on government spending was taken from publicly reported state budgets and other online sources. In total, the paper calculates that governments spent $16 billion on licenses in 2017. This $16 billion includes driver's licenses, vehicle licenses, and associated services.

Consumer spending on licenses is based on product line detail in the 2017 Economic Census. This paper assumes that consumers attend drivers' education courses to get certificates that are required to obtain or keep a driver's license rather than to learn actual driving skilled. The Economic Census reports that commercial and vocational schools earned $0.8 billion from drivers' education courses. This $0.8 billion includes classes for prospective drivers and also current drivers who are improving their skills. Information on the revenue earned by public schools from sold drivers' education courses was not located. Similarly, information on the revenue earned by driving tutors or other businesses who do not answer the Economic Census was not located. For now, this paper assumes that those two suppliers are about equal in size to the businesses tracked in the Economic Census and consumers spent a total of $1.6 billion on licenses in 2017.

Household production of licenses is based on the time spent obtaining a license. To start out, the initial learners permit requires a visit to the department of motor vehicles and filling out paperwork. Including travel time and waiting, this process requires about four hours of time from the prospective driver. Most states also require a prospective driver to attend drivers' education courses that involves more paperwork. Finally, the prospective driver must visit the department of motor vehicles again and fill out even more paperwork before they are granted a new license. Prospective drivers often fail either the exam for their initial learner's permit or the exam for their license, and so may be required to visit the department of motor vehicles more than twice before they can obtain a new driver license (Greg's Driving School 2023). In addition, drivers under eighteen are required to bring a parent to fill out paperwork at each department of motor vehicles visit and even drivers over eighteen drivers may need a companion to transport them and the vehicle they will be using for their driving test. For now, this paper assumes that each new driver's license requires 80 hours of unpaid time. This paper also assumes

that individuals with a current driver's licenses spend an average of 4 hours per year renewing either their driver's license or their vehicle licenses. This paper then multiplies that time by the hourly cost of tax preparation (Soloveichik 2023) to calculate that household production was $57 billion in 2017.

This paper calculates government spending, consumer spending, and household time costs for every year from 1929 onwards. Annual data on the number of licensed drivers is available from 1963 onwards (Federal Highway Administration 2020). Between 1950 and 1963, this paper uses estimates of drivers for select years and interpolates between those years (Federal Highway Administration 2022). Before 1950, this paper uses information on the number of registered vehicles (Federal Highway Administration 2000) and on state laws associated with driving (Federal Highway Administration 1997) to extrapolate the number of licensed drivers back to 1929.[1] Finally, the paper multiplies that index of licensed drivers by the cost index that will be developed in a later section to calculate annual data creation for each year.

Criminal Driving Citations

Government spending on criminal driving citations is calculated using a three-step procedure. First, information on total spending on police, courts, and prisons is taken from NIPA table 3.15.5. Second, the paper splits total spending between criminal data creation and other government functions. The share of police time devoted to criminal law enforcement is used as a proxy to split both police and court costs. In particular, information on the share of police time devoted to criminal law enforcement is taken from recent studies of select police departments (Asher and Horowitz 2020). Similarly, the share of prison time devoted to pre-trial detention is used as a proxy to split prison costs. Information on the share of prison time devoted to pre-trial detention is calculated from 2017 U.S. Bureau of Justice Statistics information on the share of incarcerated people in jails and the share of jailed individuals who are in pre-trial detention (Maruschak and Minton 2020) (Zheng 2019). This paper then multiplies police and court spending by the share of police time devoted to criminal law enforcement and multiplies prison spending by the share of prison time devoted to pre-trial detention to get an estimate of government spending on all criminal records. This paper assumes that all criminal records are long-lived data—and therefore assigns 100 percent of government spending on criminal records to government data investment. Police and court spending is then multiplied by the ratio of drunk-driving arrests to total arrests,[2] as reported in the Uniform Crime Reports (U.S. Department of Justice 1933–2022), and prison spending is then multiplied by the share of unconvicted jail inmates who are held for drunk-

1. States that did not require licenses are assumed to require neither government spending, consumer spending, nor household production to become a legal driver. States that did not require a driving test are assumed to require only half the resources to become a licensed driver.

2. Arrests for unspecified crimes that may or may not be driving related are dropped from the sample.

driving arrests (Greenfield 1988) (U.S. Department of Justice 1985). In 2017, this paper calculates that governments created $12 billion of criminal driving citation data.

Consumer spending on criminal driving citations is calculated using a similar three-step process. First, information on total legal output is taken from BEA's estimates of gross output by industry. Second, information on the share of lawyer revenue associated with criminal law services is taken from the product line detail in the 2017 Economic Census. This criminal law share is then multiplied by legal output to get an estimate of consumer spending on all criminal records. This estimate is then multiplied by the same ratio of drunk-driving arrests to total arrests to estimate consumer spending of $2 billion on legal services associated with criminal driving citations in 2017.

Household production of criminal driving citations is calculated based on time spent. To start out, the initial arrest process requires an individual to accompany police officers to a station, answer questions, and cooperate with relevant evidence collection. This process takes at least 2 hours (Devendorf 2021) and may take longer. Some individuals are released immediately after arrest, and some individuals are held until they can be bailed out or even until trial. Based on a one-time study of the jail population (Greenfield 1988), this paper calculates that the average drunk-driving arrest required 16.5 hours in jail. Finally, individuals are required to appear in court. Court appearances often require careful preparation ahead of time and long waits on the day of the court appearance. Information on the time associated with court appearances could not be located. For now, this paper assumes that each individual arrested for drunk driving spends a total of 80 hours of effort dealing with their arrest. This paper then multiplies that time by the hourly cost of tax preparation (Soloveichik 2023) to calculate that households spent $4 billion worth of effort managing criminal driving citations. This estimate of household production is somewhat imprecise. On the one hand, costs would be lower if some of the time associated with an arrest was considered short-term conflict resolution rather than long-term data creation. On the other hand, costs would be higher if the paper accounted for the very small but nonzero number of deaths associated with arrests (Edwards et al. 2019). For simplicity, this paper assumes these two adjustments cancel out so that the total household creation of criminal driving data is $4 billion.

Noncriminal Driving Records

Government spending on noncriminal driving records is calculated similarly to criminal driving records. The only difference is that accused individuals are almost never placed in pre-trial detention and therefore prison costs are irrelevant. This paper calculates police and court spending by multiplying total police and court spending with an estimate of the share of police time devoted to traffic (National Highway Traffic Safety Administration 1975) and then subtracting the police time related to criminal

driving records that was estimated in an earlier subsection. The time share reported in that early study is reasonably similar to share of police time devoted to noncriminal driving records in a recent study (Asher and Horowitz 2020).[3] Between those two years with data, this paper uses the ratio of drunk-driving arrests to total arrests to interpolate and extrapolate the share of police time devoted to noncriminal driving records from 1929 to 2020. Based on those assumptions, the paper calculates that government expenditures on noncriminal driving records were $41 billion in 2017.

Consumer spending on noncriminal driving records is calculated using a similar two-step process. First, information on total legal output is taken from BEA's estimates of gross output by industry. Vehicle accidents account for three-quarters of personal injury legal fees (Anderson et al. 2010). This paper assumes that lawyer fees are split evenly between plaintiffs and defendants, so that three-eighths of civil law negligence services are associated with individuals suing insurers for damages associated with vehicle accidents.[4] Information on the share of lawyer revenue associated with civil negligence law services is taken from the 2017 Economic Census. This number is then multiplied by total legal output to calculate that consumer expenditures on noncriminal driving records were $19 billion in 2017.

Household production of noncriminal driving records is calculated based on time spent. Based on a Bureau of Justice Statistics survey, this paper calculates that Americans in 2017 had approximately 38 million encounters with police associated with noncriminal driving records (Langan et al. 2001).[5] This paper assumes that each police encounter is associated with 80 hours of driver time. To be clear, the 80 hours of driver time includes much more time than just time spent talking to police offices during a traffic stop. Drivers spend a lot of time watching for police vehicles and communicating with others on how to avoid police vehicles (Yawman 1977) (Dube 2021). And drivers who receive documents from the police must decide how to respond to those documents. This paper then multiplies that time by the hourly cost of tax preparation to calculate an hourly time cost of $63 billion in 2017.

---

3. The category "responding to noncriminal calls" includes both driving accident responses and non-driving responses. For now, this category is assumed to be distributed similarly to other police time.

4. It may be true that non-accident related traffic citations are common. But individuals rarely hire lawyers to fight those citations. This paper assumes that expenditures on non-accident related traffic citations are only 1 percent of expenditures on criminal citations.

5. This paper assumes that 10 percent of reports made by crime victims are for vehicle theft or other driving-related crimes. Individuals who report two or more encounters with a police officer are assumed to have 2.5 encounters.

Driving-Behavior Data

Insurers, employers, and other entities started collecting data on driving behavior in the 1990s, and data collection has grown steadily since then (Karapiperis et al. 2015). For example, rideshare companies automatically monitor speed (Smialowski 2016) for the self-employed workers associated with their platforms. Furthermore, modern programmers are using sophisticated techniques to summarize raw telematics data into risk scores (Gao et al. 2022) that are much more predictive than the mileage based scores suggested by early theorists (Vickrey 1968). Accordingly, including driving-behavior data has a noticeable impact on total data creation in recent years but a minimal impact before the 2000s.

Expenditures per monitored driver are estimated from industry literature. Consumer monitoring costs are taken from a recent academic paper which reported that data collection costs $70 per driver per year (Jin and Vasserman 2019). On-the-job monitoring costs are taken from a Federal Motor Carrier Safety Administration publication which estimated that data collection for commercial truck drivers costs $419 per truck per year (Federal Register 2015). The higher annual cost for on-the-job data monitoring is likely due to both more hours of data collection and more intensive data collection. To be clear, the data collection costs shown here only apply to sophisticated electronic monitoring. Non-electronic monitoring devices like paper logbooks or bumper stickers that ask other drivers to report driver behavior are assumed to neither cost much nor produce much useful data—and therefore are ignored. This paper then extrapolates the costs of electronic monitoring forward and backward using BEA's pre-existing price index for cellular services (NIPA table 2.4.5U, line 283).

Monitored drivers also suffer a disutility cost from the loss of privacy. The same academic paper that reported consumer data collection costs of $70 per year also reported a consumer disutility cost of $186 per monitored driver per year (Jin and Vasserman 2019). Precise estimates of the on-the-job disutility cost of monitoring were not located, but the industry literature suggests that workers who are monitored on-the-job have a high disutility cost from the monitoring (Premack 2022). For now, this paper assumes that the ratio of disutility costs to equipment costs is the same for workers as it is for consumers. This paper also assumes that workers target a fixed level of utility during their work day, and therefore respond to the disutility of monitored driving by reducing the quantity or quality of other output (Thiel et al. 2021). In other words, the disutility of worker monitoring is ultimately borne by businesses. This paper then uses the hourly cost of tax preparation (Soloveichik 2023) to extrapolate those calculated disutility costs forward and backward.

Total data creation costs are calculated by multiplying monitoring costs with an estimate of the number of drivers monitored. Information on the number of monitored consumer drivers is taken from a survey

which reported that the share of consumers purchasing usage-based insurance doubled from 8 percent in 2016 to 16 percent in 2020 (JD Power 2022), an industry report which estimated that 1 million U.S. drivers used Progressive Insurance's Snapshot policies in 2013 (ABIresearch 2013),[6] and a survey which reported that awareness of usage-based insurance grew from 10 percent of drivers in 2010 to 36 percent in 2013 (Yao 2018). Accordingly, this paper calculates that 21 million consumer drivers were monitored in 2017. Information on the number of monitored working drivers was not located, but the industry literature suggests that driver monitoring has been growing at 10 percent per year (Allied Market Research 2022). The industry literature also suggests that the government mandate for truckers to collect data increased monitoring in 2018 (Premack 2022). For now, this paper assumes that the share of employees monitored on the job was 40 percentage points in 2017, grew to 50 percentage points in 2018, and grew 10 percent per year in all other years. BEA's published statistics show that the total number of workers was 160 million in 2017 (NIPA tables 6.5A-D and 6.7A-D) and 11.3 percent of jobs involve at least some driving (Beede et al. 2017). Accordingly, this paper calculates that a total of 7 million workers were monitored on-the-job in 2017. This paper multiplies these two estimates of monitored drivers by the cost of consumer driver monitoring and worker monitoring to calculate total data creation costs of $15 billion in 2017.

3. Valuing Vehicle Insurance Claims

To the best of my knowledge, no academic paper has yet measured either data creation costs or service flows from vehicle insurance claims alone. But the literature suggests that individuals who have both driving records and vehicle insurance claims pay lower insurance premiums, sell their used cars quicker, and earn higher incomes. This paper uses the standard capital formulas (Jorgenson 1963) to calculate service flows from driving-record data. This paper then assumes that all service flows that are not accounted as associated with driving-record data are from vehicle insurance claims and uses the standard capital formulas (Jorgenson 1963) to calculate investment from those service flows.

Lower Insurance Premiums

Newly licensed drivers pay insurance premiums about 40 percent higher than otherwise similar drivers with at least 5 years of experience (Rosanes 2023). These higher premiums cannot be explained by driving experience, because recent immigrants are charged higher rates even if they have driving experience in their prior country of residency (Rosanes 2023). Furthermore, previous research has

6. Progressive was an early leader in usage-based insurance, so they represented a large share of total driver data.

shown that new drivers do not have any hidden information about their accident probability (Chiappori and Salanie 2000) (Cohen 2005) (Spindler et al. 2014). So, the higher premiums cannot be explained by adverse selection. Instead, this paper assumes that the 40 percent premium discount received by experienced drivers is an implicit payment for their many years of data. The 2017 Economic Census reports that drivers paid a total of $263 billion in vehicle insurance premiums. This paper assumes that the 82 percent of drivers licensed for more than 5 years all received 40 percent discounts for their 5 or more years of driving data and the remaining 18 percent of drivers licensed received average discounts of 15 percent for their 0 to 4 years of driving data. Hence, this paper calculates that driving data reduced premiums by a total of $101 billion in 2017.

Insurers also offer other benefits to individuals with long driving-data histories. Vehicle insurers sometimes allow parents with a good claims history to get discounts on their children's insurance (Chiappori and Salanie 2000) (Quinn 2009). Life insurers also offer discounts to individuals who have good driving data (Cornelius 2022). Information on the capital services associated with other insurance data usage was not located. For now, this paper assumes the services associated with parents sharing their good claims history with their children are 10 percent of the services associated with individual usage and the services associated with life insurance usage are 10 percent of the services associated with individual vehicle insurance usage. Based on those assumptions, this paper assumes that driving records and vehicle insurance claims combined to save drivers a total of $122 billion of insurance premiums in 2017.

Faster Used-Car Sales

Used-vehicle sellers often use vehicle-history reports, which include both ownership history and any vehicle insurance claims, to inform potential buyers about a vehicle's condition (Routhier and Lanctot 2023). A recent academic paper demonstrated that used-vehicle sellers who posted free car history reports sold their cars 6 days faster (Huang et al. 2021). That paper calculated that those faster sales increased seller profits by $250 per vehicle.[7] This paper multiplies that $250 estimate by the 39.2 million used-vehicle transactions that occurred in 2017 (James and Paul 2018) to calculate a total seller profit increase of $10 billion in 2017.

7. This methodology could underestimate the value of vehicle history reports if some potential car buyers were able to access vehicle history reports on their own.

Higher Worker Earnings

More than 10 percent of jobs involve at least some driving (Beede et al. 2017). Those jobs are often restricted by law or employer choice to individuals with good driving records. For example, trucking companies are required to check employee driving records (U.S. Department of Transportation 2023), and major rideshare platforms require drivers to have held a U.S. license for at least 1 year and have a good safety record (HyreCar 2020). Previous academic research has shown that undocumented immigrants who become eligible for driver's licenses due to changes in state law earn around 2 to 3 percent more per year (Amuedo-Dorantes et al. 2020) [8] (Cho 2022).[9] On the one hand, this annual increase is much smaller than the 9 percent increase in average hours worked for teenagers who become eligible for driver's licenses due change in the state laws associated with drivers under 18 (Argys et al. 2019),[10] the 37 percent average earnings increase associated with a taxi license applicant passing their licensing exam (Kardja and Sundberg 2023), the $12,000 hypothetical earnings loss associated with suspended licenses (Chien et al. 2022), or the 53 percent hiring rate reported by participants in a program that helped individuals get their suspended license reinstated (Seidman Research Institute 2017). On the other hand, this annual increase is much larger than either the annual earnings decrease associated with a suspended license (Pager et al. 2022) (Finlay et al 2023) or the annual cost of restoring a suspended license (Sweeney et al. 2022).[11] For now, this paper will use the papers focusing on immigrants because that population seems more representative of the general driving population than either teenagers, potential taxi drivers, or individuals with suspended licenses. BEA's published statistics show that total worker earnings (labor compensation plus proprietors' income) was $12 trillion in 2017. Furthermore, U.S. Federal Highway Administration data show that approximately 85 percent of the total adult population was licensed in 2017 (Federal Highway Administration 2020).[12] Hence, this paper calculates that the total earnings increase associated with driving data ranges from $203 to $304 billion [(2 to 3 percent) * ($11,940 billion) * (85 percent)]. This paper then takes the midpoint of that range, $253 billion, as the earnings increase from driving data.

---

8. That paper studies the impact on employment, hours worked, and wages per hour separately for both men and women. This paper combines the coefficients in table 2 to calculate earnings increase for each gender and then multiplies by the earnings for each gender in table 1 to calculate an average impact.

9. Based on the results after 3 years shown in table A.1

10. Based on the employment decrease shown in table 3 and the hours increase shown in table 4.

11. Based on the $920 million debt and 1.133 million drivers suspended reported on page 5, it would cost $811 to restore the typical suspended license. That paper does not report a mean length for suspension, but it says that 75 percent last more than 1 year and 33 percent last more than 5 years. Hence the annual cost is well under $800.

12. This 85 percent includes the small but nontrivial portion of licensed drivers who are uninsured (Kluender 2022).

At first glance, the $253 billion of higher earnings associated with driving data seem very large. This paper calculates that reduced business costs from lower on-the-job vehicle insurance premiums associated with driving data were only $11 billion in 2017. Even when workers' compensation and other costs are included, total business costs associated with on-the-job accidents were only $39 billion in 2019 (Miller and McKnight 2021). However, accidents are only one cost associated with bad driving. Fuel efficiency decreases when drivers speed (Friedman et al. 2009) and bad on-the-job driving may damage a business's brand (Rossi 2016). In addition, bad driving data can sometime indicate the presence of conditions like attention deficit hyperactivity disorder (Chang et al. 2015) or alcohol dependence (Hingson and Winter 2003) that can be associated with non-driving problems at work. Hence, it is quite plausible that the total costs correlated with bad driving records are sufficiently high that firms are willing to pay workers $253 billion extra for their driving data.

To be clear, the $253 billion of increased earnings is not a precise measure of the benefits received by newly licensed drivers. On the one hand, individuals with driving licenses work more hours (Amuedo-Dorantes et al. 2020) and are more likely to commute to work by car (Cho 2022). These changes all impose a cost on households that partially offset the increased earnings. On the other hand, some types of consumption are associated with driving data. For example, people who live far away from a grocery store may need a driver's license to obtain fresh fruit (Economic Research Service 2009), and people who want to rent a vehicle may need good driving data (Murphy 1994). Similarly, some types of recreation are associated with driving data. For example, older adults who stop driving report higher levels of social isolation (Qin et al. 2020). For simplicity, this paper assumes that the household costs associated with higher earnings are equal to the consumption benefits and recreation benefits. Accordingly, this paper will use the $253 billion figure without adjustment.

Calculating Vehicle Insurance Claim Creation

In total, this paper calculates that driving data yielded $374 billion (122+10+253-11) of benefits to individuals in 2017. This paper then calculates[13] that driving records yielded $65 billion of shared data services to governments, $206 billion of shared data services to individuals, $7 billion of in-house data services to insurers, and $4 billion of in-house data services to businesses. Hence, this paper calculates that the vehicle insurance claims yielded $168 billion (374–206) of shared data services in 2017.

---

13. Government capital services are assumed to equal to consumption of fixed capital (BEA 2022), and private capital services are assumed to consumption of fixed capital plus a return on the capital stock (Jorgenson 1963). Both government capital and private capital are calculated using the perpetual inventory method and the depreciation rates that will be discussed in section 5. The rate of return on private capital is also discussed in that section.

This paper then uses the standard capital service formula and a standard 7 percent real rate of return (Jorda et al. 2019) to calculate that $153 billion of claims data were created in 2017.

At first glance, $153 billion of claims data creation appears implausibly large. After all, NIPA table 2.4.5U shows that consumers only received $73 billion of vehicle insurance services in 2017. However, those $73 billion of consumer vehicle insurance services are calculated based on net premiums (gross premiums less expected insurance claims) (BEA 2022). In other words, the same vehicle insurance claims that generate the claims data studied in this section are not even included in BEA's current measures of insurance industry output. The 2017 Economic Census reports that vehicle insurers collected a total of $263 billion of gross premiums. This paper tracks those gross premiums as total industry output and also tracks costs associated with insurance claims (repairs, medical care, and so on) as intermediate inputs. The increased value of insurance industry output associated with using gross premiums rather than net is approximately equal to the value of the newly tracked claims data. As a result, this paper does not change non-data insurance industry output much.

Information on claims data creations costs for years other than 2017 could not be located. For now, this paper extrapolates costs using incurred losses as a proxy. To be clear, only a portion of incurred losses create useful data and some data are created without any associated losses. Nevertheless, incurred losses are generally associated with claims data, and so are a reasonable proxy in the absence of better information. First the paper collected data on aggregate incurred losses from the Insurance Industry Factbook (Insurance Information Institute 2008, 2012, 2013, 2015, 2017 and 2021). Consistent information on aggregate incurred losses before 1997 could not be located. Accordingly, this paper uses a periodic Illinois government report comparing the state loss ratio with the national loss ratio (Clark 2003) (Shapiro 1999), an academic paper on automobile liability losses (Derrig and Tennyson 2008), a GAO report giving the average loss ratio for a selected sample of insurers between 1978 and 1987 (Posner 1989), and the Historical Statistics of the United States: Colonial Times to 1970 (U.S. Department of Commerce 1975). Gross premiums are reported in the Statistical Abstract of the United States (U.S. Census Bureau 2011, 2005, 2001, 1993, 1986, 1984, 1980, 1977, and 1970) and the Historical Statistics of the United States: Colonial Times to 1970 (U.S. Department of Commerce 1975). Between years with information on average losses, the ratio of PCE net vehicle insurance premiums to private passenger premiums is used as an interpolator.

4. Nominal GDP and Nominal Household Production

This section tracks five separate intangible assets: licenses, criminal driving citations, noncriminal driving citations, driving behavior, and vehicle insurance claims. All five intangible assets combine digital and non-digital data together. Digital and non-digital data are combined for both practical and theoretical reasons. Practically, reliable information splitting digital data and non-digital data was not located. Theoretically, the storage method has little impact on data production, data depreciation, or data usage. Accordingly, splitting digital driving data and non-digital driving data would add measurement complexity without changing anything theoretically important. Readers might worry that combining digital and non-digital data conflicts with the official guidelines' focus on computerized data (United Nations 2008, sec. 10.109). However, national accountants studying the digital economy routinely include similar non-digital products in their analysis. For example, BEA's digital economy satellite account includes television programs throughout the entire time period studied (Highfill and Surfield 2022) even though television in the United States remained analog for several years after the digital economy satellite account started (Cole 2009). Similarly, previous BEA research on specific data-related activities does not distinguish between activities that create digital data and activities that create non-digital data (Calderon and Rassier 2022) (Soloveichik 2023). Hence, this paper's treatment of non-digital driving records and non-digital claims data does not conflict with national accounting practices even if it is inconsistent with the official guidelines.

Figure 1. Data Creation Relative to Nominal GDP

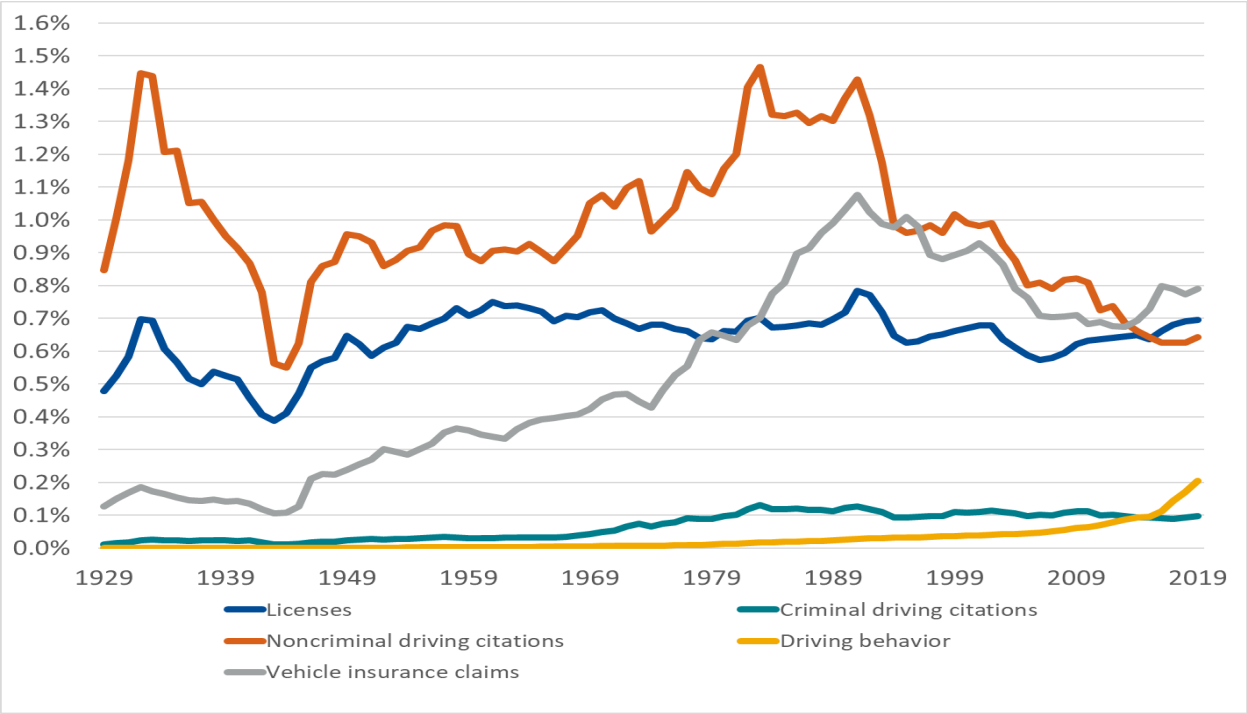


Figure 1 shows that the total ratio of driving-data creation to overall GDP has hovered around 2 percent from 1980 onward. This ratio is smaller than the 5 percent ratio of data creation to overall GDP found in an earlier paper that studied financial data (Soloveichik 2023). Nevertheless, it is still a noticeable amount of data that has the potential to change the economy. The remainder of the paper will discuss how measured GDP growth, measured household, and measured total factor productivity (TFP) growth change when driving data are included.

Current Accounting for Data Creation by Sector

Private business expenditures on driving data do not currently impact measures of GDP that are based on industry value added, because driving data are not considered final output. If data creation services are purchased from a specialist firm, they are included as output of the producing company and intermediate input for the purchasing company. This increase to value added for the producing company is precisely canceled out by the decrease to value added for the purchasing company so that sales of private business data creation services to businesses have no net impact on total value added (just as any other intermediate input). In currently published statistics, own-account data creation by businesses is not included as either output or intermediate input and therefore does not directly impact any industry's value added in official BEA statistics. Similarly, private business data do not directly impact measures of GDP that are based on final demand.

Government and nonprofit expenditures on data creation do impact measures of GDP. For those sectors, BEA measures output (and implicitly value added) based on costs rather than market revenue. Wages for workers who create driving data are included in total labor costs and therefore implicitly included in measured value added of the government or non-profit entity producing data. Furthermore, a government purchase of driving-data creation services from a private business raises measured value-added for the private business without lowering measured value-added for the government.

Consumer purchases of insurance data creation services are already included in measured GDP. Driving schools explicitly sell certificate creation services to consumers, and law firms explicitly sell driving-record creation services to consumers. In addition, insurance firms implicitly bundle claims data creation services together with their primary insurance services. As a result, consumer purchases of insurance data creation services are currently included in PCE. Household data production is not included in measured GDP.

Proposed Accounting for Data Creation and Data Ownership by Sector

This paper proposes to track all data as intangible capital assets. This treatment raises measured business output by the newly included data investment and raises measured government output by the newly included depreciation on data capital. Consumer purchases of data are already included in final output as a component of PCE. To avoid double-counting, this paper shifts those purchases from PCE services to PCE consumer durables without changing nominal GDP at all. Household data production and the services from consumer data are included in household production but not in GDP.

The paper assigns initial ownership of data to the sector that funded their creation. Governments account for a noticeable share of driving record creation, so they are treated as owning a portion of driving-record data. Consumers own both the portions of driving data created by businesses on behalf of consumers and the portions of driving data created by consumers themselves. Similarly, businesses own both the portions of driving data that are created by businesses on behalf of other businesses or by the businesses themselves. Consumer driving-behavior data are an insurance industry asset, and worker driving-behavior data are an asset of the firm that hires the workers. Finally, in-house insurance claims data are treated as an insurance industry asset.

Data are often transferred from consumers to businesses. For example, a young adult might build a good driving history by using their personal vehicle responsibly and later use that good driving record to get insurance on their business vehicle. This is treated as a negative purchase of consumer durables and a positive business investment. The converse situation, where businesses fund the creation of driving data that are used by consumers, is very rare. Consistent with the asymmetric transfers from consumers to businesses, insurance industry statistics show the CLUE database was associated with a reduction in the commercial share of the vehicle insurance industry, from 16 percent in 1990 to 11 percent in 1999 (U.S. Census Bureau 2002). Therefore, neither labor compensation nor proprietors' income are adjusted for on-the-job data creation.

Revisions to Nominal GDP Associated With Driving Data

Figure 2. Revisions to Nominal GDP From Capitalizing In-House Data

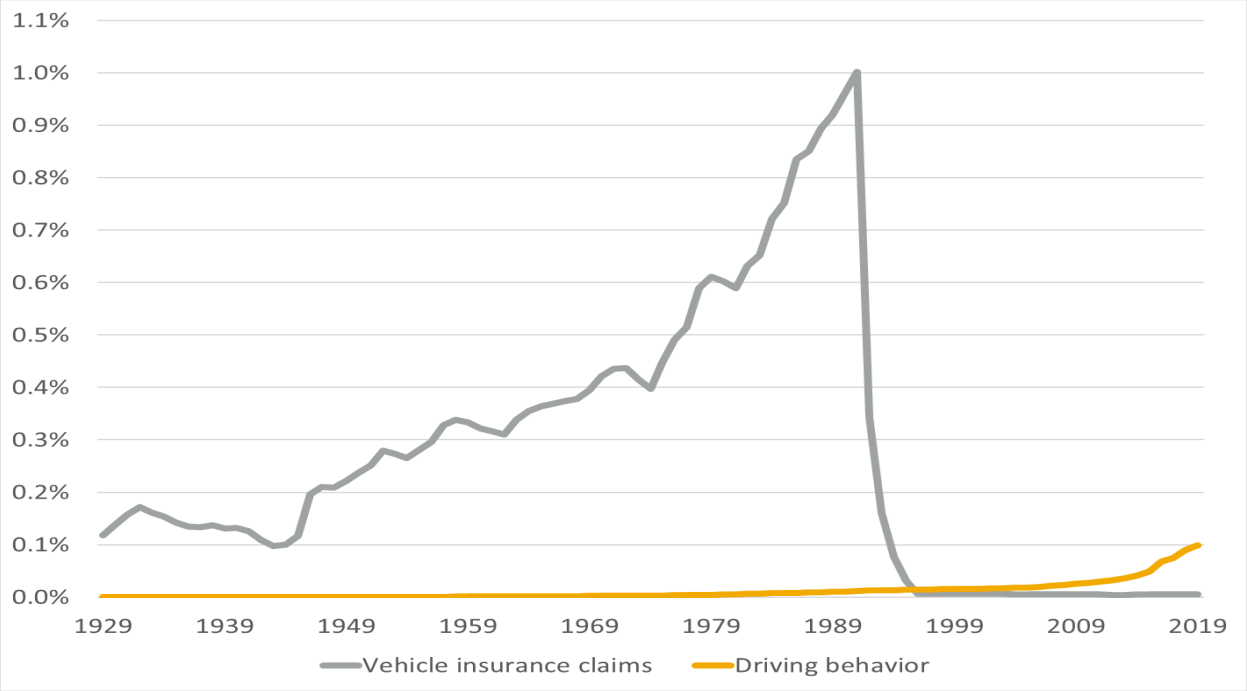
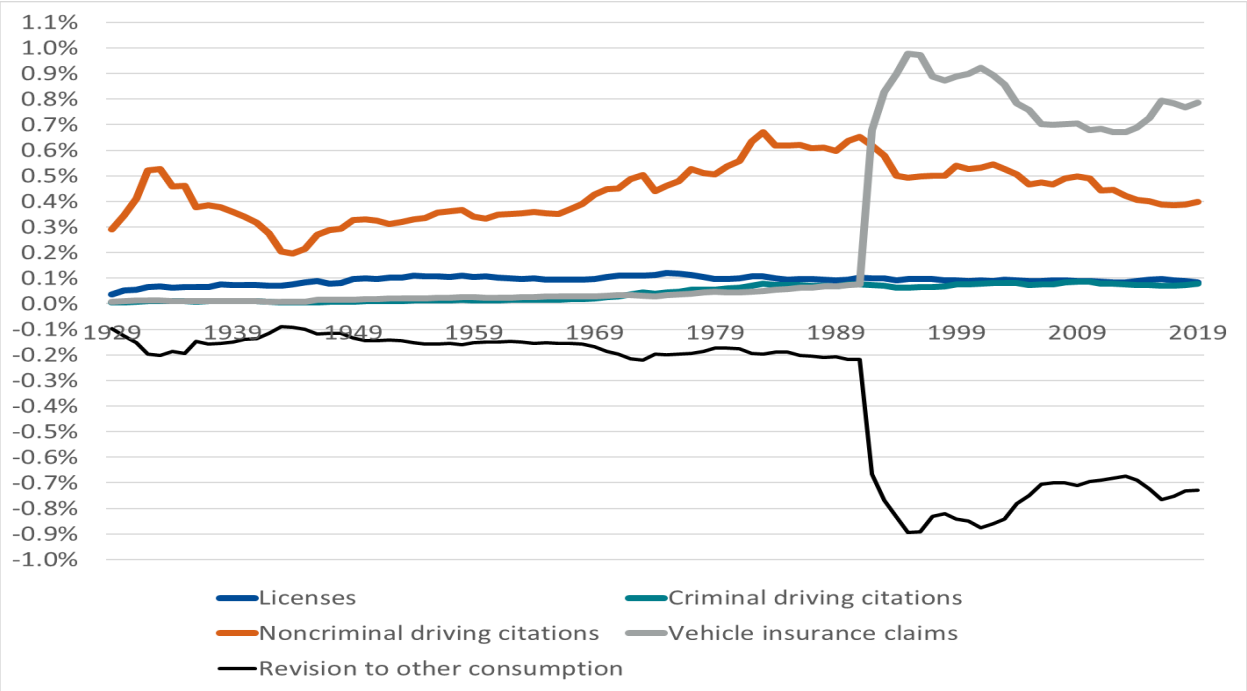


Figure 3. Revisions to Nominal GDP From Capitalizing Shared Data



Figures 2 and 3 show that vehicle insurance claims suddenly switched from being in-house data to being shared data. To remind readers, this switch is due to the introduction of the CLUE system in the early 1990s. Interestingly, in-house claims data have a larger GDP impact than shared claims data. The difference can be explained by the treatment of claims data associated with consumer vehicle policies. When claims data are owned by businesses, data creation is included in GDP without any compensating adjustments. But when claims data are owned by consumers, data creation is exactly offset by a downward revision to PCE services.

Figures 2 and 3 also show that capitalizing data raises the level of nominal GDP in every year studied. It may be true that some shared data were already counted in measured personal consumption expenditures or measured government consumption and were therefore already part of GDP. However, neither the in-house data shown in figure 2 nor the portion of the shared data shown in figure 3 that are business assets were ever included in measured consumption. Accordingly, capitalizing driving data increases private-sector output by the value of business investment in driving data. In addition, government output increases by the consumption of fixed capital associated with government-owned driving records.

Depreciation Rates, Capital Stock, and Capital Services Associated With Driving Data

Each type of driving data has its own depreciation rate. Licenses require regular renewal (Insurance Institute for Highway Safety 2023) but the renewal process is typically straightforward and generally requires only one visit to the department of motor vehicles (Hawley 2023). This paper assumes that licenses depreciate at 5 percent per year. Criminal driving citations can impact insurance rates for up to 10 years (Gusner and Tilford 2023), so this paper assumes that they depreciate at 10 percent per year. The 1970 Fair Credit Reporting Act mandated that noncriminal driving citations (Spencer 2020) and shared claims data be deleted after 7 years (Masterson and Metz 2022). This paper uses an adjusted version of the one-hoss shay (Diewert 2004) to calculate an annual depreciation rate of 31 percent for noncriminal driving citations and shared claims data. To be clear, academic research has shown that claims data predict accident risk for many years (Rubinstein 2023), and so the 31 percent depreciation rate used in this paper might decrease noticeably if a longer period before data deletion was allowed.[14] Neither current driving-behavior data nor historical vehicle claims data are regulated by the 1970 Fair Credit Reporting Act—and therefore they have the potential to last longer than 7 years. However, both

14. Shared claims data were almost nonexistent before 1970, and so the reduced lifespan associated with the Fair Credit Reporting Act had no impact on measured GDP or productivity. Noncriminal citations did exist before 1970, but auto insurers then had just started using them to price policies and their initial formulas only looked at the past three calendar years (Kiplinger 1968). Hence, these data were also unaffected by legal changes in 1970.

are kept in-house and therefore lose value whenever a consumer switches insurance carriers or a worker switches jobs. The industry literature (Bendig 2023) and the academic literature (Cohen 2005) suggest that consumers switch insurance carriers frequently and therefore data kept in-house depreciates rapidly. For simplicity, this paper assumes that the reduced depreciation associated with potential uses after 7 years cancels out the increased depreciation associated with switching so that in-house data depreciates at the same 31 percent per year as shared claims data.

Government services from insurance data are assumed to equal depreciation (BEA 2022). Based on the depreciation rates reported above and the assumed rate of return, this paper estimates that governments receive services equal to 5 percent of their license data stock, 10 percent of their criminal citation data stock, 31 percent of their noncriminal driving citation stock and 31 percent of their claims data each year. Both business and consumer services from driving data are assumed to equal government services plus a 7 percent rate of return based on the average real return for housing and stocks over the past 150 years (Jorda et al. 2019).[15] However, there is one crucial difference between business services and household services. When driving data are owned by insurers, data services are implicitly included in GDP as part of insurance services. But when driving data are owned by consumers, data services are explicitly included in household production as part of consumer durable services. All else equal, a switch from in-house claims data to shared claimed data lowers nominal GDP growth. Conceptually, this decrease in measured GDP is similar to the GDP drop that occurs when cooking or cleaning time switches from the market sector to the household sector (Coyle 2015) (Bridgman 2023).

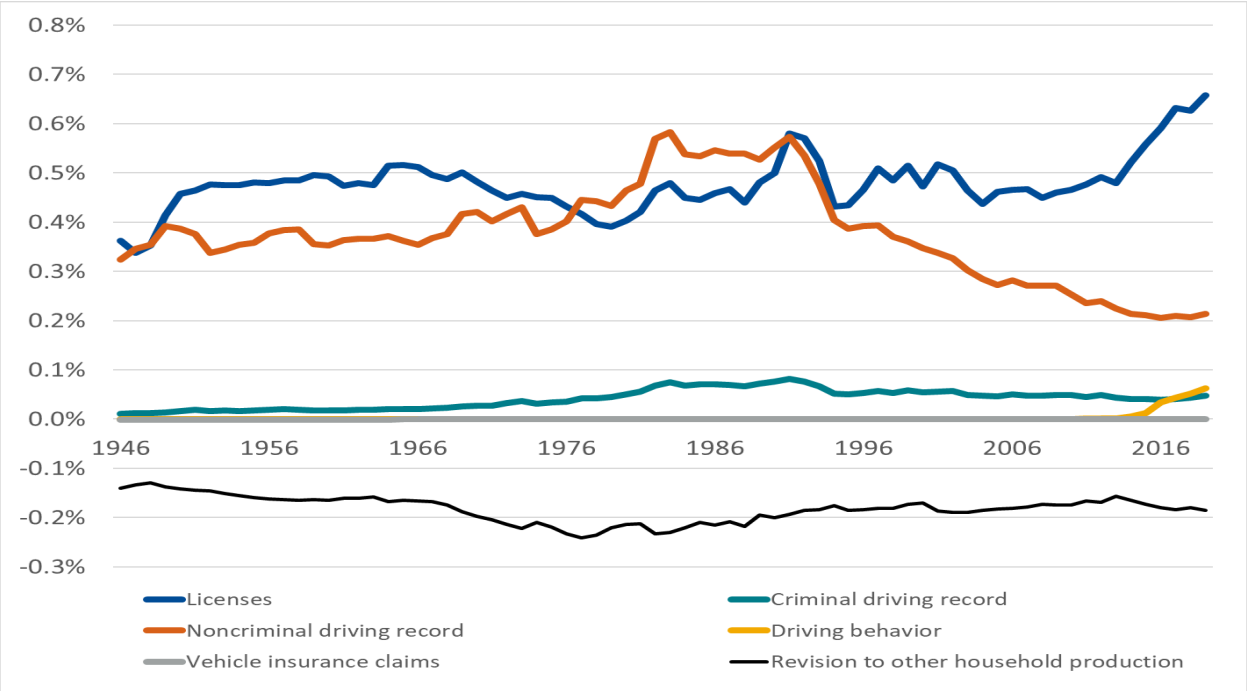
Revisions to Nominal Household Production Associated With Driving Data

Total consumer time devoted to driving-data creation was 2.3 billion hours in 2017. It is not obvious how these 2.3 billion hours are recorded in the American Time Use Survey (U.S. Bureau of Labor Statistics 2023). Time spent helping a teenager obtain a license might be included in the 71 billion hours that Americans spend on child or adult care. Time spent looking for police while driving or speaking to a police officer during a traffic stop might be included in the 105 billion hours that American adults spend in their private vehicles. Time spent paying a ticket might be included in the 2.6 billion hours that Americans spend on financial management. Time spent obtaining a license for yourself or resolving a traffic ticket might be included in the 0.1 billion hours that Americans spend obtaining licenses and paying fines, fees, and taxes. Finally, time spent in police custody or in court might be included in the 0.2 billion hours that Americans spent on civic obligations and government services. Information on the

15. The calculations add general inflation and subtract data inflation to get a specific rate of return for driving data.

share of driving-record creation time that is currently included in household production could not be located. This paper assumes that only 25 percent of driving-data creation time is currently included in household production. In addition, this paper uses a specialist wage to value data creation time that is higher than the generalist wage used to value other household production time. As a result, the value of newly tracked data creation is much larger than the downward revision to other household production.

Figure 4. Revisions to Household Production From Data Creation



Figures 4 shows that driving data creation accounts for a growing share of household production. A portion of this relative growth can be explained by growth in the average amount of time spent on driving-data creation. For example, the number of speeding tickets issued per capita increased between 1946 and 2019. Another portion can be explained by the vehicle insurance claims data switching from a business asset to a consumer durable in the 1990s. However, the lion's share of this relative growth can be explained by a decline in the share of time spent on other household production (Aguiar and Hurst 2007). As a result, a steady amount of time spent at the department of motor vehicles accounts for an ever-growing share of household production.

Figure 5. Revisions to Nominal Household Production From Data Services

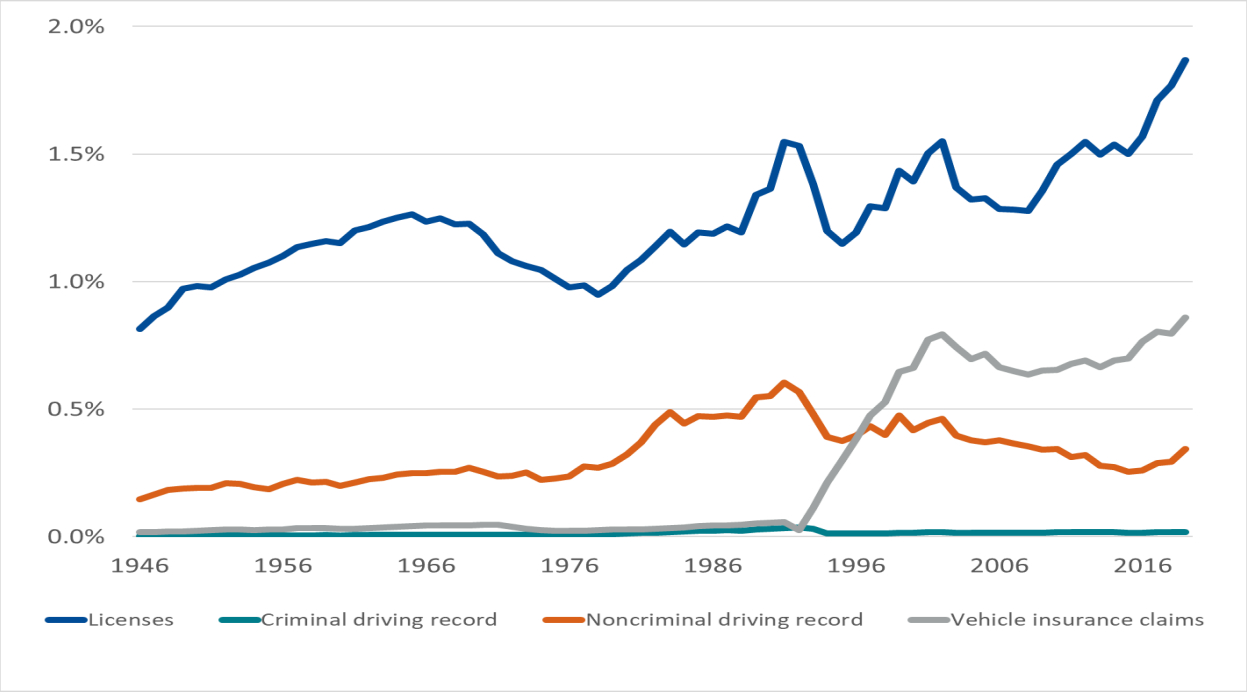


Figure 5 also shows that driving data services have grown faster than overall household production. To save space, the remainder of this paper combines household data creation with household data services and discusses the two together.

5. Data Prices, Real GDP, and Real Household Production

This paper studies five types of driving data: licenses, criminal driving citations, noncriminal driving citations, driving behavior, and vehicle insurance claims. All of these types except driving behavior can be created without computers and stored on simple text files. As a result, digitization lowers their cost without transforming them entirely. In contrast, driving-behavior data relies on sophisticated information and communication technology that did not exist before the 1990s. Driving-behavior data are still a small category, but they have the potential to transform the automobile insurance industry and prevent many deaths (Ben-Shahar 2023). Hence, the combined price for all types of driving data may fall when sufficiently high-quality driving-behavior data make the old data types obsolete.

Driver's License, Criminal Record, and Noncriminal Record Prices

Production costs are calculated as the weighted average of costs in scope for GDP and costs out of scope for GDP. For licenses and noncriminal driving citations, costs in scope for GDP are assumed to track

government costs from BEA's pre-existing price index for the government function of "public order and safety." For criminal driving citations, costs in scope for GDP are calculated by simply dividing the nominal costs estimated in section 3 with the number of drunk-driving arrests in each year. The costs out of scope for GDP are assumed to equal the hourly cost of tax paperwork (Soloveichik 2023). To be clear, those hourly costs do not include that paper's adjustment for labor productivity growth associated with tax software and other technology improvements. Each of the three data types has a separate cost share for hourly costs. Hence, the weighted average costs calculated for each data type are different.

Production costs are then adjusted for quality growth associated with the Driver License Compact. Since its formation in 1960, the Driver License Compact has facilitated data sharing between the participating states. Based on the share of accidents suffered by out-of-state drivers (Solomon 1964) and the share of the population who move across state lines (Palarino et al. 2023), this paper estimates that about 10 percent of driving data were impacted by lack of communication across states. Accordingly, this paper assumes that the full data sharing associated with every state participating in the Driver License Compact or an equivalent organization would represent a 10 percent increase in data quality. It then calculates data quality in each year based on the weighted fraction of states that are currently participating in the Driver License Compact.

Finally, production costs for criminal driving citations are adjusted for the quality growth associated with the National Instant Criminal Background Check System (NICS). Before the 1990s, it was difficult for businesses to access criminal data (Senate Hearings 1986). In the early 1990s, the FBI created the NICS, which made criminal records easy to access and used standardized formats. The NICS is now used regularly for background checks by employers (Jacobs and Crepet 2008). In addition, modern search engines now allow both businesses and individuals to view criminal records directly (Finlay 2009) without going through the NICS or background check firms. This paper treats the easier access and standardized format associated with the NICS as a sudden doubling in the quality of criminal data and therefore a halving of the quality-adjusted price. Driver's licenses and noncriminal driving citations were never added to the NICS, and are therefore their quality did not increase during the 1990s.

Driving-Behavior Prices

Production costs are calculated as the weighted average of data collection costs and the disutility cost for drivers of having their behavior monitored. Equipment costs are assumed to track BEA's pre-existing price index for cellular services (NIPA table 2.4.5U, line 283) and the disutility cost is assumed to track the hourly cost of tax paperwork used earlier (Soloveichik 2023). The weights for the two components are taken from the weights reported in previous (Jin and Vasserman 2019).

Going forward, production costs may need to be adjusted for data sharing. Early research has shown that driving-behavior data can be used to diagnose Alzheimer's disease (Bayat et al. 2021), solve crimes (Koops et al. 2019), improve bus service quality (Coyle and Diepeveen 2021), and so on. Future researchers may discover additional ways to use shared data. Hence, data quality might improve dramatically if the insurance industry switched from in-house data to shared data. But all driving-behavior data are currently used in-house, and therefore quality has not yet improved noticeably.

Vehicle Claims Prices

Production costs for claims data are calculated as nominal losses per insured driver. Production costs are then adjusted for the quality growth associated with data sharing. In-house vehicle claims data are less useful because they cannot be used to screen workers, sell used cars, or set life insurance premiums. This paper calculates that these uses of driving data account for more than two-thirds of the total value associated with claims data. In addition, the net present value of newly created data are lower when they have the geometric depreciation schedule associated with in-house data rather than the one-hoss shay schedule associated with data regulated by the Fair Credit Reporting Act. Considering both the additional uses of shared data and the different depreciation schedule, this paper calculates that the CLUE database increased claims data quality more than fivefold.

Revision to Overall GDP Prices and Household Production Prices

Capitalizing data impacts overall GDP prices through three separate channels. Most importantly, average price growth changes when a new output category with its own price index is added to the national accounts. Similarly, household production prices also change when a new output category with its own price index is added to the household satellite accounts. In addition, BEA currently uses input prices to impute government output and household production prices. As a result of this imputation, measured government output prices change when driving-record data services are included as an input to government services, and measured household output prices change when both driving-record data services and insurance claims data services are included as inputs to household production. Finally, measured PCE prices change when the consumer purchases related to insurance data are reclassified from consumer services to the consumer durable of insurance data if the price indexes differ.[16]

16. If all data-related services had the same price, this shift would not change overall GDP prices. However, BEA's PCE deflators for insurance services and legal services have both risen faster than the price of data.

Figure 6. Revision to GDP Prices From In-House and Shared Data

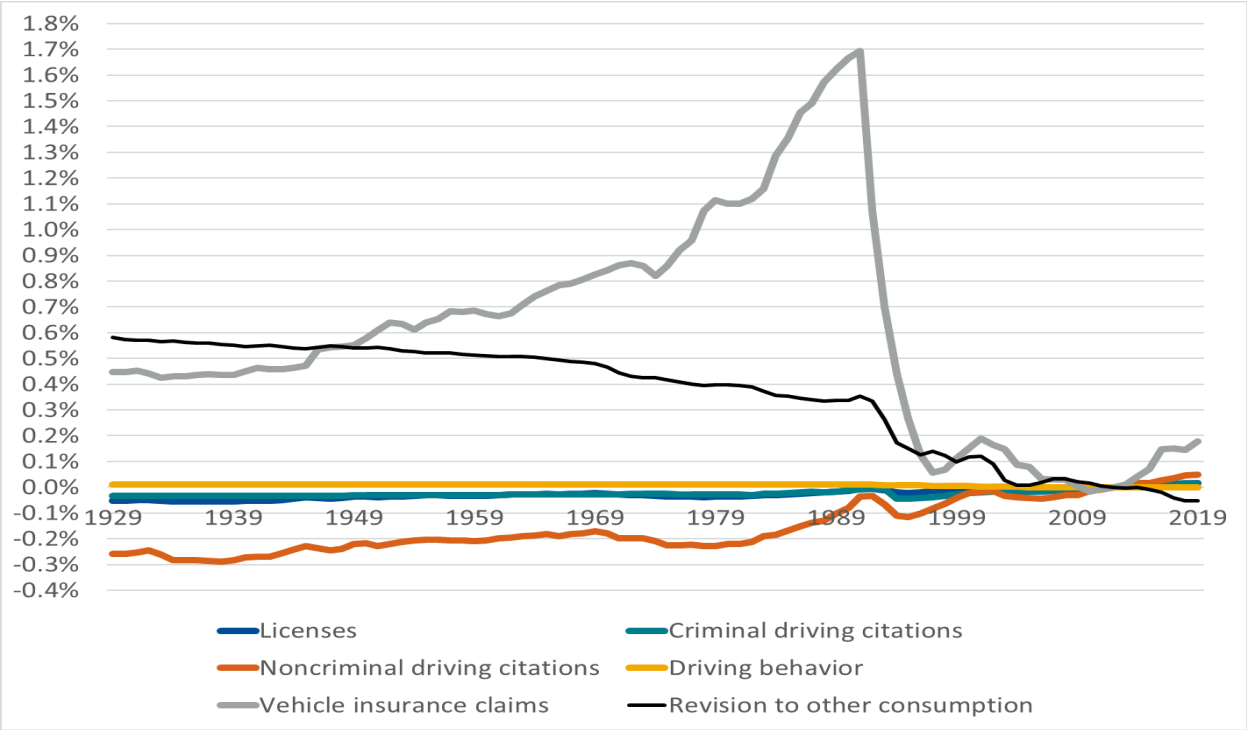
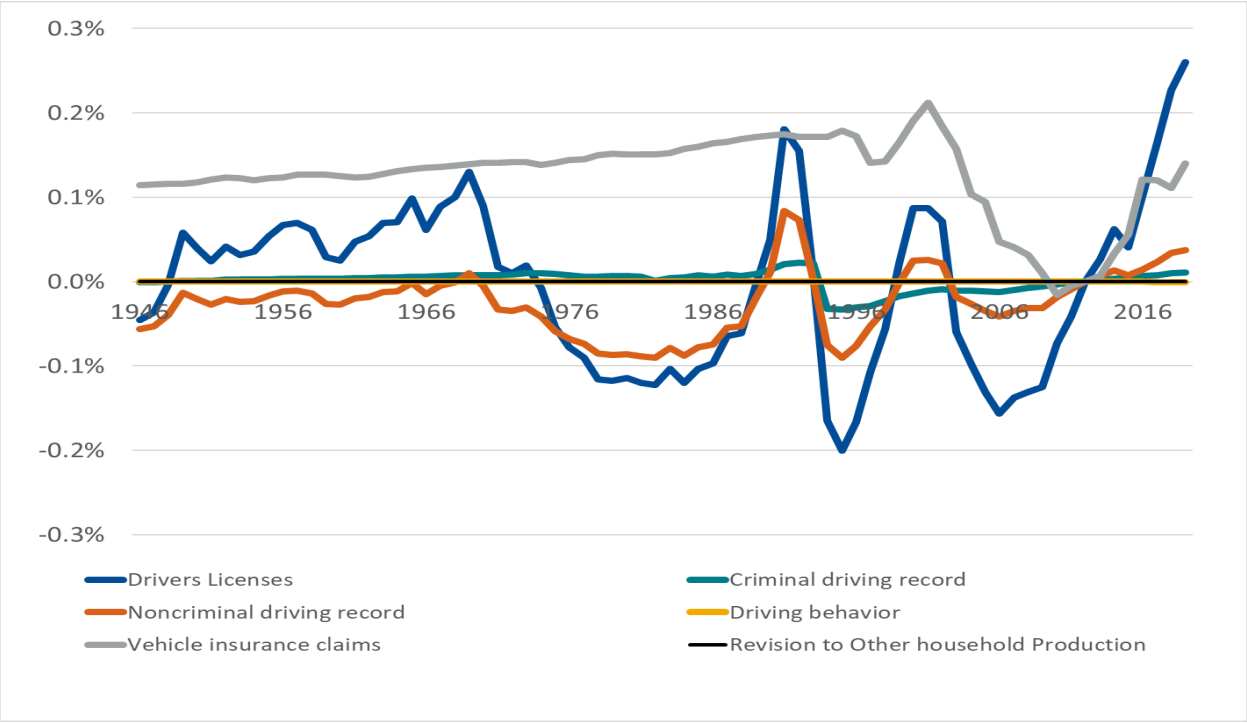


Figure 7. Revision to Household Production Prices From Tracking Data Creation and Services



Revision to Overall GDP Quantities and Household Production Quantities

Figure 8. Revision to GDP Quantities From In-House and Shared Data

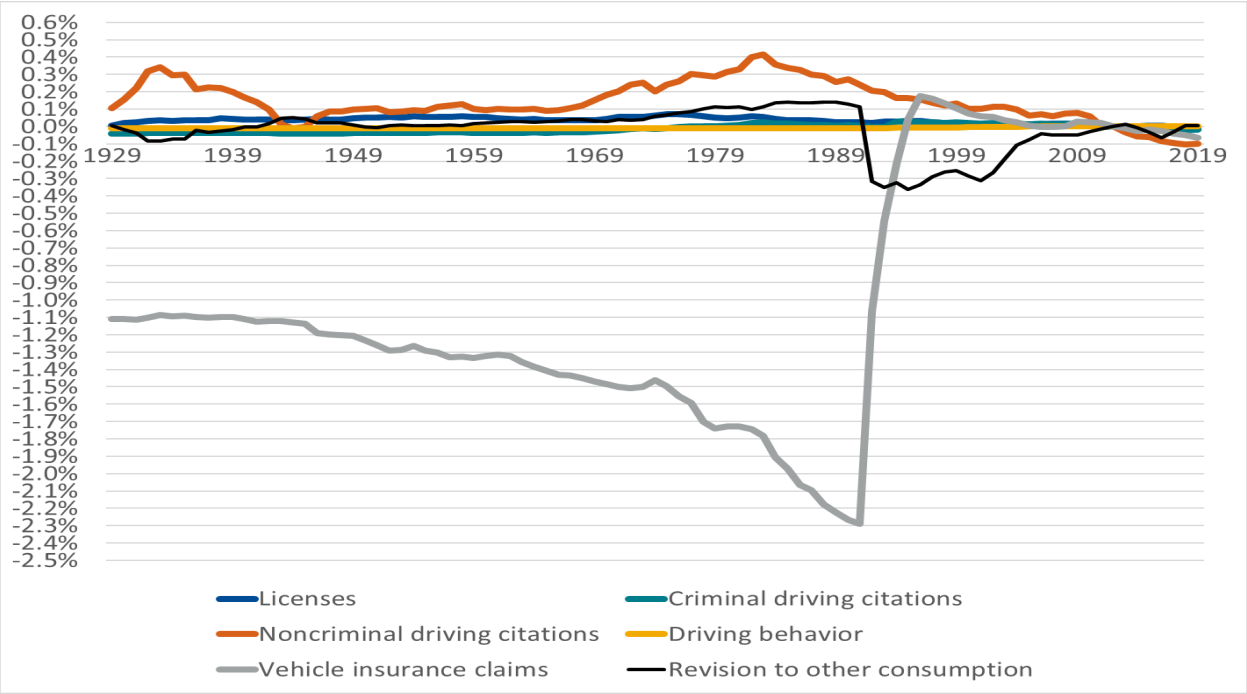


Figure 8 shows that including driving data raises real GDP growth in the early 1990s by a total of 2.3 percentage points. This large growth increase is almost entirely due to the assumption that shared data are much higher quality than in-house data. The exact growth increase in each year is somewhat sensitive to precise timing of the switch from in-house data to shared data associated with the CLUE database—but the general result is relatively robust. To be clear, the GDP increase associated with the CLUE system and the digitization of driving records may have costs that offset the increased real output. Individuals value their privacy (Acquisti et al. 2016) and might have enjoyed significant nonmonetary benefits from negative data being kept secret. This ambiguity reinforces the fact that GDP is not necessarily equivalent to welfare.

The growth increase shown in figure 8 may be large enough to change our understanding of macroeconomic history. The late 1990s are well-known for enjoying rapid economic growth, and many previous authors have argued that the growth acceleration should be attributed to information and communication technology (Oliner and Sichel 2000) (Anderson and Kliesen 2006) (Jorgenson et al. 2008). Figure 8 shows including driving data in GDP decreases the measured acceleration to real GDP growth for the period 1994–1999, relative to the period 1989–1994, from 1.4 percentage point per year

to 1.2 percentage point per year. In other words, shared driving data are capable of explaining 0.2 percentage point of the 1.4 percentage point growth acceleration.

It is possible that other shared data may explain additional portions of the growth acceleration. The CLUE database started sharing claims for home and other property insurance around the same time it started sharing claims for vehicle insurance and the National Practicioners Data Bank started sharing medical malpractice claims data in 1990 (Waters and Budetti 2004). In addition, Walmart created a system to automatically share sales data with suppliers in 1990 (Basker 2007) and the NICS shared data on non-driving related crimes. If those types of data sharing had anywhere near the impact of vehicle data sharing, then they might be able to explain much if not all of the measured growth acceleration during the late 1990s.[17]

To be clear, this paper does not study either price dispersion or consumption inequality. When risk data are available, insurance companies charge high premiums to high-risk customers and low premiums to low-risk customers. High-risk customers often suffer from other problems like low grades in school (Shuldiner 1968), low credit scores (Federal Trade Commission 2007), and substance abuse disorder (Recovery Centers of America 2019). These high premiums may seem particularly unfair because insurance is often not a genuinely voluntary purchase: most states require drivers to maintain liability insurance (Kluender 2022), and almost all lenders require the owner of a financed car to maintain comprehensive coverage as well (Deventer and Buckley 2023). Researchers studying the automobile loan market have calculated that deletion of negative credit data can sometimes be a low-cost method of reducing inequality between individuals with good credit and bad credit (Jansen et al. 2022). It is possible the deletion of negative driving data would also be a low-cost method of reducing inequality between individuals with good driving data and bad driving data.[18] BEA's current measures of inequality focus on income inequality (Fixler et al. 2020) and are therefore not directly impacted by consumer price dispersion. The impact of risk data on measured consumption inequality is a topic for future research.

---

17. The specific data sharing systems introduced in the early 1990's relied on information and communication technology (ICT) . Hence, the argument that data sharing explains some of the productivity acceleration does not necessarily conflict with the argument that ICT explains the growth acceleration.

18. Risk data may create other social problems. For example, individuals who are quoted higher insurance prices due to negative data may choose to drive without insurance and therefore raise insurance premiums for other drives (Sun and Yannelis 2016). These externalities are beyond the scope of GDP.

Figure 9. Revision to Household Production Quantities From Tracking Data Creation and Services

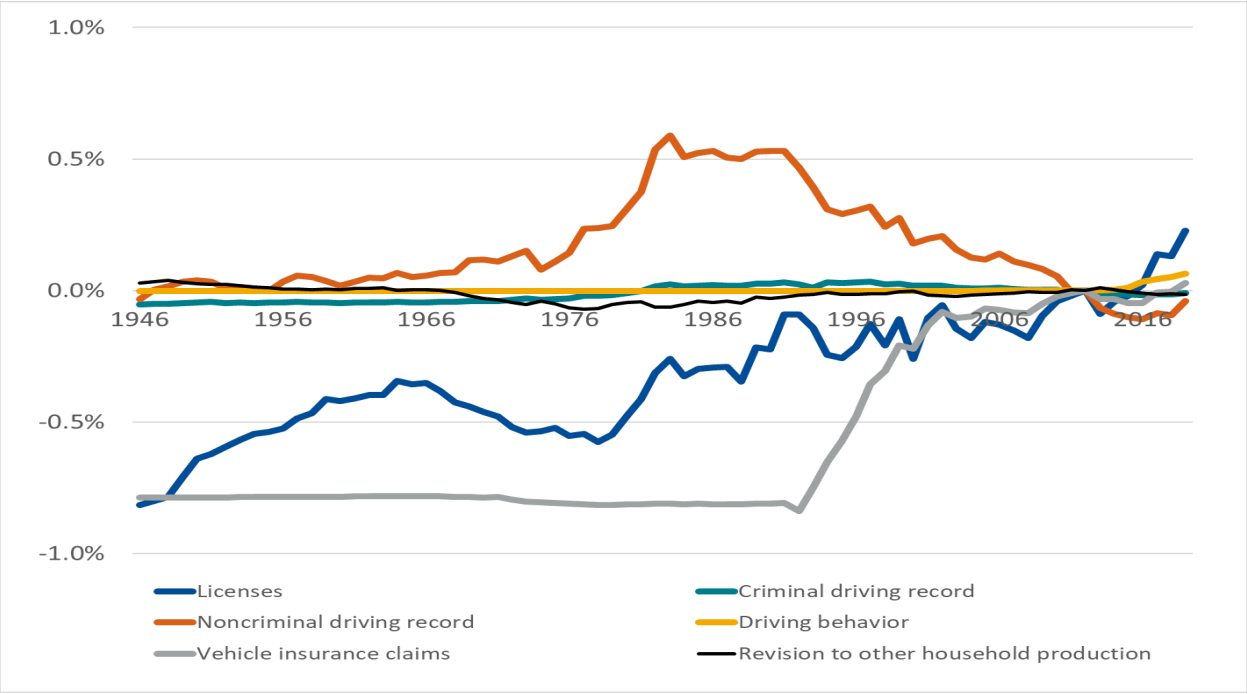


Figure 9 shows that including insurance data also raises real household production growth. By itself, the upward revision to real growth shown in figure 9 is not enough to change the qualitative trends in household production reported in BEA's satellite account (Bridgman et al. 2022). However, a previous paper on financial data found that including tax forms and individual credit reports also raised real growth in some periods (Soloveichik 2023). If including other types of long-lived data also raises real growth, then the overall revision associated with including all types of household data may be enough to change post–1975 qualitative trends. This is a topic for future research.

6. Impact of Driving Data on Business Productivity

The productivity calculations presented in this paper are based on existing industry-level production accounts that track labor, capital services, and intermediate inputs for 61 separate private business sector industries (Garner et al. 2020). These production accounts calculate productivity for each industry separately and then aggregate across the entire economy with Domar weights. This paper calculates revisions to productivity for each industry separately. Each industry has its own outputs and inputs, and so the relative impact of tracking insurance data can be very different across industries. This paper then calculates how those revisions change TFP in each industry and aggregates each industry's TFP revision using revised Domar weights to calculate a total TFP revision for the overall private business sector.

This paper makes five revisions to each of the existing “before redefinitions” industry-level production accounts. First, data created by businesses and owned by businesses are considered part of value added and therefore increase measured business output. For example, commercial driver's license data created on-the-job by truckers are now tracked as output of the trucking industry. Second, data that are given to a customer as part of a service bundle are deflated with a newly developed price index for data creation services rather than deflated with the industry’s primary output price index. For example, claims data reported by automobile insurance companies are deflated with a claims data price index rather than the production account’s existing price index for insurance services. Third, newly included insurance data services received from customers, business owners, and employees increase the measured intermediate inputs of data users. For example, the data services associated with vehicle claims data that are shared with an insurance company by an insurance applicant are now tracked as an intermediate input of the insurance company. Fourth, measured output for data users increases by discounts on the list price that they use to implicitly pay customers for access to customer insurance data. For example, insurance services are increased by the implicit payment for the data services associated with an applicant’s claims data. Fifth, capital services for data users decreases by the portion of capital compensation that is classified as payment for insurance data of business owners and labor services for data users decreases by the portion of labor compensation that is classified as payment for insurance data of workers. For example, the proprietors’ income from a small business is split between general proprietors’ income and services associated with the small-business owner’s driving record. Businesses use both data created by businesses and data created by households, and so both types of data services are included identically in the productivity account revisions described.

Figure 10. Impact on Business Productivity from Including Driving Data

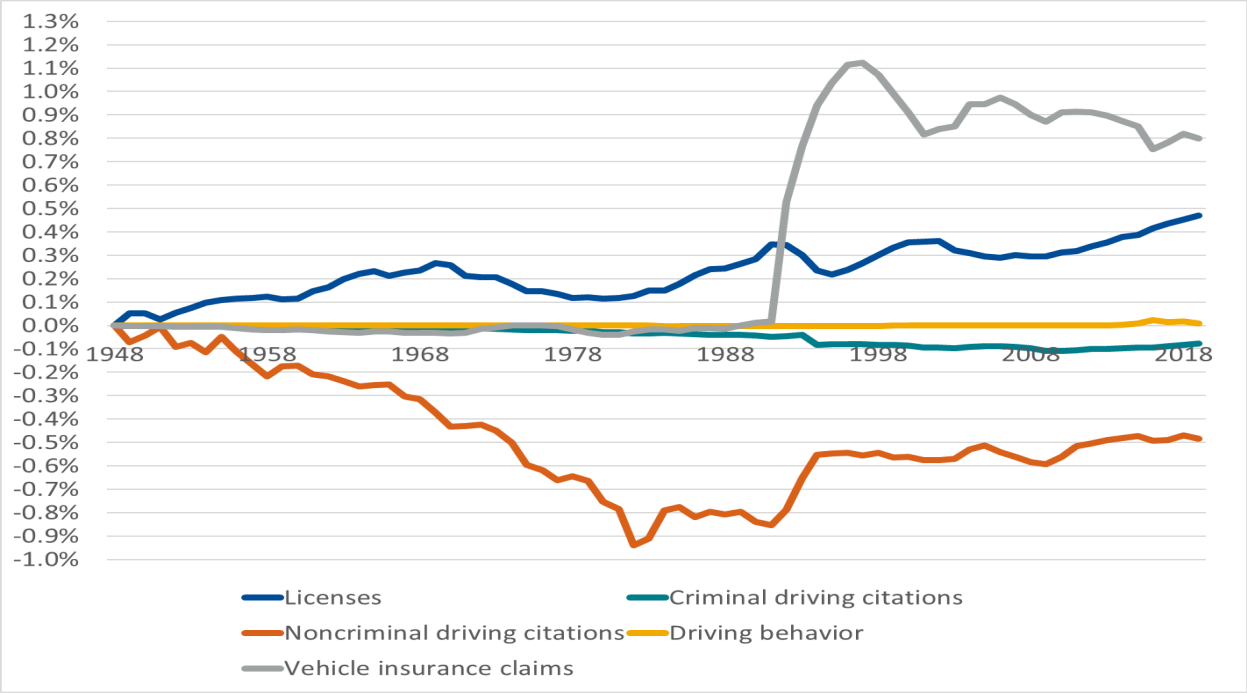


Figure 10 shows that productivity growth spikes noticeably when vehicle insurance claims data are included. The increase to measured productivity shown in figure 10 can be decomposed between outputs and inputs. On the one hand, real output jumped immediately after the CLUE database increased claims data quality. On the other hand, real data services grew more slowly because it took a few years for the higher quality claims data to provide additional capital services. The net impact is a spike in measured productivity during the early 1990s and then a leveling off. The measured acceleration to productivity growth for the period 1994–1999, relative to the period 1989–1994, decreases from 0.43 percentage point per year to 0.23 percentage point per year when driving data are included in the productivity accounts.

Figure 10 also shows that neither criminal driving citations nor driving behavior have much impact on measured productivity growth. The null result for criminal driving citations is due to the fact that data of this type are mostly created and used by governments rather than businesses. So, they have little impact on measured business productivity. A broader measure of productivity that includes the government sector might show a very different results for this data type. Driving-behaviour data have little impact on measured productivity growth because they are still a very small data type. But these data are growing rapidly and may impact measured growth in the future.

Conclusion

This paper presented case studies of two types of driving data: driving records and insurance claims. These two types of data are simple enough that they can be stored on paper (U.S. Women's Bureau 1944) and can be managed by relatively unskilled workers (BLS 2022), but so valuable that they had a creation cost of $382 billion in 2017. This paper calculated that tracking driving data as an intangible asset increases measured GDP in 2017 by $128 billion and measured household production in 2017 by $160 billion. This paper also calculated that including driving data increases measured GDP growth in the early 1990s noticeably and explains half of the late 1990s acceleration to measured productivity growth.

Bibliography

ABIresearch (2013) "Global Insurance Telematics Subscriptions to Exceed 100 Million by 2018, but Auto Insurance Faces Dramatic Changes," posted at *ABIresearch.com* on June 6, 2023, and accessed September 27, 2023.

Acquisti, A., C. Taylor, and L. Wagman (2016) "The Economics of Privacy," *Journal of Economic Literature* 54(2), pages 442–92.

Agan, A., A. Garin, D. Koustas, A. Mas, and C. Yang (2022) "The Impact of Criminal Records on Employment, Earnings, and Tax Filing," unpublished manuscript available at https://www.irs.gov/pub/irs-soi/07-2022-impact-incarceration-employment-earnings-tax-filing.pdf.

Aggarwal, P., A. Brandon, A. Goldszmidt, J. Holz, J. List, I. Muir, G. Sun, and T. Yu (2022) "High-frequency Location Data Shows that Race Affects the Likelihood of Being Stopped and Fined for Speeding," working paper posted at https://bfi.uchicago.edu/wp-content/uploads/2022/12/BFI_WP_2022-160.pdf and accessed August 8, 2023.

Aguiar, M., and E. Hurst (2007) "Measuring Leisure: The Allocation of Time over Five Decades," *Quarterly Journal of Economics* 122(3), pages 969-1006.

Allied Market Research (2022) "Driver Monitoring Systems Market," Report Code A04525.

American Association of Motor Vehicle Administrators (2023) "D20 Traffic Records Systems, Data Dictionary (AMIE)," posted at https://www.aamva.org/getmedia/d4c16fd8-2193-490c-a5ea-21607a3bd51a/D20-Traffic-Records-Systems-Data-Dictionary-(AMIE).pdf and accessed August 8, 2023.

Amuedo-Dorantes, C., E. Arenas-Arroyo, and A. Sevilla (2020) "Labor Market Impacts of States Issuing of Driver's Licenses to Undocumented Immigrants," *Labor Economics* 63(April), 101805.

Anderson, J., P. Heaton, and S. Carroll (2010) *The U.S. Experience with No-Fault Automobile Insurance*, Rand Corporation, posted at https://www.rand.org/pubs/monographs/MG860.html.

Anderson, R., and K. Kliesen (2006) "The 1990s Acceleration in Labor Productivity: Causes and Measurement," *Federal Reserve Bank of St. Louis Review* 88(3), pages 181–202.

Argys, L., T. Mroz, and M. M. Pitts (2019) “Driven from Work: Graduated Driver License Programs and Teen Labor Market Outcomes,” *Federal Reserve Bank of Atlanta Working Paper Series* 2019–16.

Asher, J. and B. Horwitz (2020) “How Do the Police Actually Spend Their Time?,” *New York Times*, published on June 19, 2020, and accessed August 9, 2023.

Barraclough, P., A. Wahlberg, J. Freeman, B. Watson, and A. Watson (2016) “Predicting Crashes Using Traffic Offenses. A Meta-Analysis that Examines Potential Bias Between Self-Report and Archival Data,” *PLoS One* 11(4), e0153390.

Basker, E. (2007) “The Causes and Consequences of Wal-Mart’s Growth,” *Journal of Economic Perspectives* 21(3), pages 177–198.

Bayat, S., G. Babulal, S. Schindler, A. Fagan, J. Morris, A. Mihailidis, and C. Roe (2021) “GPS Driving: A Digital Biomarker for Preclinical Alzheimer’s Disease,” *Alzheimer’s Research and Therapy* 13, article 115.

Becker, G. (1968) “Crime and Punishment: An Economic Approach,” *Journal of Political Economy* 76(2), pages 169–217.

Beede, D., R. Powers, and C. Ingram (2017) “The Employment Effect of Autonomous Vehicles,” *Economics and Statistics Administration Issue Brief* #05–17.

Ben-Shahar O. (2023) "Exploring the Resistance to Data Technology, The Debate over Tracking in Usage-Based Insurance,” working paper posted at https://ilc.law.uconn.edu/wp-content/uploads/-sites/3281/2023/02/BenShahar_Usage_Based_Insurance_UConn.pdf, and accessed August 8, 2023.

Bendig, E. (2023) “Inflation Causes Drivers to Switch Car Insurance Carriers,” posted at *Kiplinger.com* on June 5, 2023, and accessed September 27, 2023.

Boyle, J. (2019) “Answer Man: DMV Accident Reporting System Caused Loss of Uber, Lyft Gigs?” *The Citizen Times*, posted April 4, 2019, and accessed July 25, 2023.

Brady, R. and L. Glover (2023) “How to Get Car Insurance Without a License,” posted at *Nerdwallet.com* on May 15, 2023, and accessed on July 26, 2023.

Brenton, M. (1964) *The Privacy Invaders*, Fawcett World Library, New York City, New York.

Bridgman, B., A. Craig, and D. Kanal, (2022) "Accounting for Household Production in the National Accounts: An Update 1965–2020," *Survey of Current Business* 2022 (2).

Bridgman, B. (2023) "A Disaggregated View of Household Production Trends," *AEA Papers and Proceedings* 113(May), pages 619–622.

Bureau of Economic Analysis (2022) "Concepts and Methods of the U.S. National Income and Production Accounts," https://www.bea.gov/resources/methodologies/nipa-handbook/pdf/all-chapters.pdf, accessed January 17, 2023.

Calderon, J.B., and D. Rassier (2022) "Valuing Stocks and Flows of Data Assets for the U.S. Business Sector," presentation to the BEA Advisory Committee Meeting on May 13.

Carey, M., E. Anderson, R. Mansour, J. Sloan, and A. Curry (2020) "Missed Opportunities to Advance Knowledge on Traffic Safety: Accessibility of Driver Licensing and Crash Data for Scientific Research," *Accident Analysis & Prevention* 139: 105500.

Chang, Z., P. Lichtenstein, B. D'Onofrio, A. Sjolander, and H. Larsson (2014) "Serious Transport Accidents in Adults with ADHD, and the Effect of Medication: A Population Based Study," *Journal of the American Medical Association Psychiatry* 71(3), pages 319–325.

Chapman, R. (2022) "Inherent Diminished Value in Motor Vehicle Torts," *Boston Law Review* 102, pages 1649–1683.

Chiappori, P. and B. Salanie (2000) "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy* 108(1), pages 56–78.

Chien, C., A. George, S. Shekhar, and R. Apel (2022) "Earnings Loss Associated with a Criminal Record and Suspended Driver's License," Arizona Law Review 64, pages 675–732.

Cho, H. (2022) "Driver's License Reforms and Job Accessibility Among Undocumented Immigrants," *Labor Economics* 76 (102174).

ChoicePoint (2001) "C.L.U.E. Personal Auto, How to Read the Report," posted online at www2.epic.org/privacy/-choicepoint/fcraltr/CLUEAutoUnderwriter.pdf and accessed August, 2023.

Clark, J. (2003) “Annual Report to the Illinois General Assembly on Insurance Cost Containment,” posted April 15, 2003, and accessed November 8, 2023.

Cohen, A. (2005) “Asymmetric Information and Learning: Evidence from the Automobile Insurance Market,” *The Review of Income and Statistics* 87(2), pages 197–207.

Cole, T. (2009) “Congress Backs Delay of Digital TV Switch to June,” *National Public Radio* article, posted February 4 2009 and accessed February 2023.

Congressional Research Service (2011) “Child Support Enforcement and Driver’s License Suspension Policies,” https://crsreports.congress.gov/product/pdf/R/R41762/5.

Cornelius, N. (2022) “How Your Driving Record Affects Your Life Insurance,” posted at *Quotacy.com* on September 30, 2022, and accessed November 7, 2023.

Coyle, D. (2015) “GDP: A Brief But Affectionate History—Revised and Expanded Edition,” *Princeton University Press*.

Coyle, D. (2022) “Socializing Data,” *Daedalus* 151 (2), pages 348–359.

Coyle, D., and S. Diepeveen (2021) “Creating and Governing Social Value from Data,” presented at the IARIW-ESCoE Conference in November 2021 and accessed in October 2022.

Coyle, D. and W. Li (2021) “The Data Economy: Market Size and Global Trade,” *SSRN Working Paper* 3973028.

Crocker, K., and S. Tennyson (2002) “Insurance Fraud and Optimal Claims Settlement Strategies,” *The Journal of Law and Economics* 45(2), pages 469–507.

Derrig, R., and S. Tennyson (2008) “The Impact of Rate Regulation on Claims Evidence from the Massachusetts Automobile Insurance,” *Casualty Actuarial Society*, 2008 Discussion Paper Program.

Desind, S. (2015) “Winning in Traffic Court Because the Cop Doesn’t Show Up,” posted at *TheTrafficTicketAttorneys.com* on May 1, 2015, and accessed July 26, 2023.

Devendorf, J. (2021) “DUI Booking and Bail,” posted at Lawinfor.com on May 27, 2021, and accessed September 27, 2023.

Deventer, C., and A. Buckley (2023) “Car Insurance for Financed Vehicles,” posted at *Bankrate.com* on April 19, 2023, and accessed September 28, 2023.

Diewert, W.E. (2004) “Issues in the Measurement of Capital Services, Depreciation, Asset Price Changes, and Interest Rates,” Discussion Paper 04-11 http://papers.economics.ubc.ca/legacypapers/dp0411.pdf.

Dube, R. (2021) “The Best Apps for Your Phone to Avoid Speed Traps,” posted at Groovypost.com on June 25, 2021, and accessed August 8, 2023.

Economic Research Service (2009) “Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences,” Report to Congress, posted online and accessed on July 26.’ 2023.

Edwards, F., H. Lee, and M. Esposito (2019) “Risk of Being Killed by Police Use of the United States by Age, Race-Ethnicity, and Sex,” *Proceedings of the National Academy of Sciences* 116(34), pages 16793–16798.

Eisenach, J. (1985) “The Role of Collective Pricing In Auto Insurance,” FTC Staff Report.

Eurostat (2020) “Recording and Valuation of Data in National Accounts,” working paper 3.4 https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.20/2020/mtg1/3.4_Recording_of_Data_in_NA_Eurostat_June_2020_after_SG_comments.pdf.

Emerson, R. (1992) “Insurance Claims Fraud Problems and Remedies,” University of Miami Law Review 46(4), pages 907–973.

Farboodi, M., and L. Veldkamp (2022) “A Model of the Data Economy,” *NBER Working Paper* 28427.

Federal Highway Administration (1997) “Year of First State Driver License Law and First Driver Examination,” posted at https://www.fhwa.dot.gov/ohim/summary95/dl230.pdf and accessed August 9, 2023.

Federal Highway Administration (2000) "State Motor Vehicle Registrations, By Years, 1900–1995," posted at https://www.fhwa.dot.gov/ohim/summary95/mv200.pdf and accessed August 8, 2023.

Federal Highway Administration (2020) "Table DL-220 Licensed Drivers, by Sex and Age Group, 1963–2019," posted at https://www.fhwa.dot.gov/policyinformation/statistics/2019/pdf/dl220.pdf and accessed August 8, 2023.

Federal Highway Administration (2022) "Licensed Drivers, Population, and Motor Vehicles," posted at https://www.fhwa.dot.gov/ohim/onh00/line5.htm and accessed August 8, 2023.

Federal Register (2015) "Electronic Logging Devices and Hours of Service Supporting Documents," Rules and Regulations 80 (241), pages 78292–78416.

Federal Trade Commission (2007) "Credit-Based Insurance Scores: Impact on Consumers of Automobile Insurance," Report to Congress.

Feinman, J. (2015) "The Regulation of Insurance Claim Practices," UC Irvine Law Review 5(5), pages 1319–1353.

Finlay, K. (2009) "Effect of Employer Access to Criminal History Data on the Labor Market Outcomes of Ex-Offenders and Non-Offenders," *Studies of Labor Market Intermediation*, University of Chicago Press.

Finlay, K., M. Gross, E. Luh, and M. Mueller-Smith (2023) "The Impact of Financial Sanctions: Regression Discontinuity Evidence from Driver Responsibility Fee Programs in Michigan and Texas," working paper available at https://sites.lsa.umich.edu/mgms/wp-content/uploads/sites/283/2023/01/CJARS_FinSanc_DRF_20230119.pdf.

Fixler, D., M. Gindelsky, and D. Johnson (2020) "Measuring Inequality in the National Accounts," *BEA Working Paper* 2020–3.

Friedman, L., D. Hedeker, and E. Richter (2009) "Long-Term Effects of Repealing the National Maximum Speed Limit in the United States," *American Journal of Public Health* 99(9), pages 1626–1631.

Gao, G., S. Meng, and M. Wuthrich (2022) "What Can We Learn from Telematics Car Driving Data: A Survey," *Insurance: Mathematics and Economics* 104, pages 185–189.

Garner, C., J. Harper, M. Russell, and J. Samuels. (2020) "Integrated Industry-Level Production Account for the United States and the Sources of U.S. Economic Growth Between 1987 and 2018," *Survey of Current Business* 100.

Gershowitz, A. (2011) "12 Unnecessary Men: The Case for Eliminating Jury Trials in Drunk Driving Cases," William & Mary Law School *Faculty Publications* 1256.

Girardin, M. (2023) "CLUE Report Guide," posted at *Wallethub.com* on March 10, 2023, and accessed September 18, 2023.

Goncalves, F. and S. Mello (2023) "Police Discretion and Public Safety," *NBER Working Paper* 31678.

Greenfield, L. (1988) "Drunk Driving," *Bureau of Justice Statistics Special Report* 109945.
Greg's Driving School (2023) "6 Steps to Your License," posted at *GregsDrivingSchool.net* and accessed on August 8, 2023.

Gusner, P., and A. Tilford (2023) "How Much do Car Insurance Rates Go Up After a DUI," posted at Forbes.com on September 11, 2023, and accessed September 27, 2023.

Harrington, E., W. Murdock, and H. Shaffer (2023) "Prediction Mistakes in the Criminal Justice System: Evidence from Linking Prosecutor Surveys to Court Records," working paper posted at https://conference.nber.org/conf_papers/f189643.pdf and accessed August 8, 2023.

Hawley, D. (2023) "How Long Does it Take to Renew Your Driving License?," posted at *JDpower.com* on April 23, 2023, and accessed September 27, 2023.

Helland, E., and F. Kipperman (2011) "Recovery Under the Medicare Secondary Payer Act," *RAND Occasional Papers Series* 332-ICJ.

Highfill, T., and C. Surfield (2022) "New and Revised Statistics of the U.S. Digital Economy, 2005–2021," BEA report, posted November 2022 and accessed February 2023.

Hingson, R., and M. Winter (2003) "Epidemiology and Consequences of Drinking and Driving," *Alcohol Research and Health* 27(1), pages 63–78.

Hoffman, E. (2013) “Reducing Churn from Price Increases: An Experimental Intervention,” unpublished paper available at https://repository.up.ac.za/bitstream/handle/-2263/40653/Hoffman_Intervention_2013.pdf?sequence=1&isAllowed=y.

Homer, T. (2022) “The Long Strange History of License Plates in the U.S.," posted at *Auto.HowStuffWorks.com* on October 6, 2022, and accessed July 26, 2023.

Huang, G., H. Liu, and H. Luo (2021) “Can You Judge a Car by Its History? Vehicle History Reports and the Efficiency of the Used-Car Retail Market,” *SSRN Working Paper* 3976378.

HyreCar (2020) “Uber Background Checks: Everything You Need to Know,” posted at *HyreCar.com* on February 14, 2020, and accessed September 18, 2023.

Insurance Information Institute (2015, 2017, and 2012) *Insurance Industry Factbook*, posted online.
Insurance Institute for Highway Safety (2023) “License Renewal Procedures,” posted at *IIhs.org* in September 2023 and accessed September 27, 2023.

Jacobs, J., and T. Crepet (2008) “The Expanding Scope, Use, and Availability of Criminal Records,” *New York University Journal of Legislation and Public Policy* 11(2), pages 177–214.

James, T. and M. Paul (2018) “Used Vehicle Sales Hit Record High in 2017, According to Latest Edmunds Used Car Report,” posted on Edmunds.com on March 21, 2018, and accessed July 18, 2023.

Jansen, M., F. Nagel, C. Yannelis, and A. Zhang (2022) “Data and Welfare in Credit Markets,” *NBER Working Paper* 30235.

JD Power (2022) “Usage-Based Auto Insurance Takes Center Stage as Satisfaction Flatlines, J.D. Power Finds,” posted at *JDpower.com* on June 13, 2022, and Access September 27, 2023.

Johnson, K. (2004) “Driver’s Licenses and Undocumented Immigrants: The Future of Civil Rights Law?,” *Nevada Law Journal* 5(1), article 11.

Johnston, P. (1951) “The Administrative Hearing for the Suspension of a Driver’s License,” North Carolina Law Review 30(1), pages 27–47.

Jones, C., and C. Tonetti, (2020) "Nonrivalry and the Economics of Data," *American Economic Review* 110(9), pages 2819–58.

Jorda, O., K. Knoll, D. Kuvshinov, M. Schularick, and A. Taylor (2019) "The Rate of Return on Everything, 1870–2015," *The Quarterly Journal of Economics* 134(3), pages 1225–1298.

Jorgenson, D. (1963) "Capital Theory and Investment Behavior," *American Economic Review* 53(2), pages 247–259.

Jorgenson, D., M. Ho, and K. Stiroh (2008) "A Retrospective Look at the U.S Productivity Growth Resurgence," *Journal of Economic Perspectives* 22(1), pages 3–24.

Karadja, M., and A. Sundberg (2023) "The Labor Market Impact of a Taxi Driver's License," *Institute for Evaluation of Labour Market and Education Policy Working Paper* 2023–6.

Kasperowicz, L., and J. Ohman (2022) "What if My C.L.U.E. and MVR Reports Are Wrong?," posted at *TheTruthAboutInsurance.com* on June 3, 2022, and accessed August 8, 2022.

Kiplinger (1968) "Auto Insurance: The Lowdown on 'Safe-Driver' Policies," published in June 1968.

Kluender, R. (2022) "Pay-As-You-Go Insurance: Experimental Evidence on Consumer Demand and Behavior," *Harvard Business School Working Paper* 23–030 available at https://www.hbs.edu/ris/Publication%20Files/23-030_73d0c3d1-fb78-40f2-b151-741ba02340e4.pdf.

Koops, B. B. Newell, and I. Skorvanek (2019) "Location Tracking by Police: The Regulation of Tireless and Absolute Surveillance," *UC Irvine Law Review* 9(3), pages 635–698.

LaFave, W. (2004) "The "Routine Traffic Stop" from Start to Finish: Too Much "Routine," Not Enough Fourth Amendment," Michigan Law Review 102(8), pages 1843–1905.

Langan, P., L. Greenfeld, S. Smith, M. Durose, and D. Levin (2001) "Contacts between Police and the Public, Findings from the 1999 National Survey," posted at https://bjs.ojp.gov/content/pub/pdf/cpp99.pdf and accessed August 9, 2023.

Lofstrom, M., J. Hayes, B. Martin, D. Premkumar, and A. Gumbs (2022) "Racial Disparities in Law Enforcement Stops," Public Policy Institution of California Report, accessed July 26, 2023.

Mancusi, C. (2018) “Should You File a Claim if the Damage is Less than Your Deductible?,” posted at *RogersGray.com* on November 1, 2018, and accessed on July 26, 2023.

Marte, J. (2014) “Just Asking About an Insurance Claim Can Make Your Rate Go Up,” published in the Washington Post on December 15, 2014, and accessed on July 26, 2023.

Maruschak, L., and T. Minton (2020) “Correctional Populations in the United States, 2017–2018,” *Bureau of Justice Statistics Bulletin* 252157.

Masterson, L., and J. Metz (2022) “What is A C.L.U.E. Report For?” posted at *Forbes.com* on November 7, 2022, and accessed September 27, 2023.

Megna, M., and L. Kasperowicz (2023) “16 Questions to Expect when Getting Car Insurance Quotes,” posted at *Insurance.com* on April 17, 2023, and accessed on July 26, 2023.

Miller, T., and A. S. McKnight (2021) “Cost of Motor Vehicle Crashes to Employers—2019,” Report prepared by the Pacific Institute for Research and Evaluation and posted online at *TrafficSafety.org*.

Mitchell, J., M. Lesher, and M. Barberis, (2022) “Going Digital Toolkit Note: Measuring the Economic Value of Data,” OECD Directorate for Science, Technology and Innovation.

Murphy, M. (1994) “Consumer News,” *Loyola Consumer Law Review* 6(2), pages 34–36.

National Highway Traffic Safety Administration (1975) “Statewide Highway Safety Program Assessment,” posted at https://www.google.com/books/edition/Statewide_Highway_Safety_Program_Assessm/aIPH-LOnp3IC?hl=en&gbpv=0 and accessed August 9, 2023.

Oliner, S., and D. Sichel (2000) “The Resurgence of Growth in the Late 1990s: Is Information Technology the Story,” *Journal of Economic Perspectives* 14(4), pages 3–22.

Overman, S. (2014) “Can Monitoring Company Vehicles Drive Safety,” posted at *SHRM.org* on July 8, 2014, and accessed September 18, 2023.

Pager, D., R. Goldstein, H. Ho, and B. Western (2022) “Criminalizing Poverty: The Consequences of Court Fees in a Randomized Experiment,” *American Sociological Review* 00(0), pages 1–25.

Palarino, J., J. Pugh, and B. McKenzie (2023) "Share of Americans Who Said They Moved Declined Between 2019 and 2021," posted at *Census.gov/library/stories* on April 11, 2023, and accessed September 28, 2023.

Posner, P. (1989) "Profitability of the Automobile Lines of the Insurance Industry, Government Accountability Office 90–04FS.

Premack, R. (2022) "The Quest to Get Truck Drivers to Stop Worrying and Love the ELD," posted at *Freightwaves.com* on November 24, 2022, and accessed on August 28, 2023.

Qin, W., X. Xiaoling, and H. Taylor (202) "Driving Cessation and Social Isolation in Older Adults," *Journal of Aging Health* 32(9), pages 962–971.

Quinn, J. (2009) "Making the Most of Your Money Now: The Classic Bestseller Completely Revised," Simon and Schuster.

Raith, M. (1996) "A General Model of Information Sharing in Oligopoly," *Journal of Economic Theory* 71(1), pages 260–288.

Rakoff, J. (2021) "Why the Innocent Plead Guilty and the Guilty Go Free: And Other Paradoxes of Our Broken Legal System," *Farrar, Straus and Giroux*, United States.

Rassier, D., R. Kornfeld, and E. Strassner (2019) "Treatment of Data in National Accounts," *Paper prepared for BEA Advisory Committee*, posted in May 2019 and accessed in February 2023.

Recovery Centers of America (2019) "Economic Cost of Substance Abuse Disorder in the United States, 2019," posted at *RecoveryCentersofAmerica.com* and accessed July 26, 2023.

RelX (2017) "Annual Reports and Financial Statements 2017," posted at https://www.relx.com/~/media/Files/R/RELX-Group/documents/reports/annual-reports/relx2017-annual-report.pdf and accessed August 8, 2023.

Richmond, D. (1994) "An Overview of Insurance Bad Faith Law and Litigation," *Seton Hall Law Review* 25(74), pages 74–140.

Rogers, M. (2023) "MVR Check for Employers: What You Need to Know," posted at *IProspectCheck.com* on June 7, 2023, and accessed on July 26, 2023.

Rosanes, M. (2023) "A Guide to Cheap Car Insurance for New Drivers," posted at Insurancebusinessmag.com on September 13, 2023, and accessed September 27, 2023.

Rossi, T. (2016) "How Does the Traffic Behavior of Company Vehicles' Drivers Affect Corporate Reputation and Consumers' Purchasing Behavior in Finland?" unpublished master's thesis available at https://www.theseus.fi/bitstream/handle/10024/114038/Rossi_Tomi.pdf?sequence=1&isAllowed=y.

Routhier, S., and C. Lanctot (2023) "Cheap Auto Insurance Companies that Don't Count Accidents Reported on CARFAX in 2023," posted at *Buyautoinsurance.com* on July 24, 2023, and accessed September 18, 2023.

Rubinstein, A. (2023) "Behavioral Professional: Evidence from the Commercial Auto Insurance Industry," unpublished working paper available at https://conference.nber.org/conf_papers/f181404.pdf.

Samuelson, P. (1954) "The Pure Theory of Public Expenditure," The Review of Economics and Statistics 36(4), pages 387–389.

Seidman Research Institute (2017) "The City of Phoenix Municipal Court's Compliance Assistance Program, 2016: An Economic Assessment," report dated June 2, 2017, posted online in 2018, and accessed August 8, 2023.

Senate Hearings (1986) "A Bill to Amend Section 534 of Title 28, United States Code, to Allow Railroad Police and Private University or College Police Access to Federal Government Criminal Identification Records," *Hearings before the Committee on the Judiciary*, testimony given January 22.

Shapiro, N. (1999) "Annual Report to the Illinois General Assembly on Insurance Cost Containment," posted April 15, 1999 and accessed November 8, 2023.

Sheldon, A. (2021) "The History of the Driver's License," posted at *Magazine.Northeast.AAA.com* on July 12, 2021, and accessed on July 26, 2023.

Shuldiner, H. (1968) "The Truth About Auto Insurance," *Popular Science* 193(2), pages 99–115.

Slikboer, R., S. Muir, S. Silva, and D. Meyer (2020) “A Systematic Review of Statistical Models and Outcomes of Predicting Fatal and Serious Injury Crashes from Driver Crash and Offense History Data,” *Systematic Reviews* 9, 220.

Smialowski, B. (2016) “Uber is Tracking Drivers’ Phones to Watch for Speeding,” posted at *Time.com* on June 29, 2016, and accessed December 12, 2023.

Solomon, D. (1964) “Accidents on Main Rural Highways Related to Speed, Driver, and Vehicle,” *Department of Commerce Report* 92.

Soloveichik, R. (2023) “Capitalizing Data: Case Studies of Tax Forms and Individual Credit Reports,” *BEA Working Paper* 2023–7.

Soloveichik, R. (2024) “Private Funding of 'Free' Data: A Theoretical Framework,” unpublished manuscript.

Spencer, D. (2020) “Will My Speeding Ticket Show Up on a Background Check?,” posted at *Addictivetips.com* on December 7, 2020, and accessed July 26, 2023.

Spindler, M., J. Winter, and S. Hagmayer (2014) “Asymmetric Information in the Market for Automobile Insurance: Evidence from Germany,” *The Journal of Risk and Insurance* 81(4), pages 781–801.

Statistics Canada (2019) “The Value of Data in Canada: Experimental Estimates,” posted on July 10, 2019, at https://www150.statcan.gc.ca/n1/pub/13-605-x/2019001/article/00009-eng.htm and accessed in October 2022.

Sun, S., and C. Yannelis (2016) “Quantifying the Premium Externality of the Uninsured,” *Journal of the European Economic Association* 14(2), pages 405–437.

Sweeney, A., M. Russell, J. Robie, and B. Mikelbank (2022) “Road to Nowhere: Debt-Related Driver’s License Suspensions in Ohio,” posted online at *LasClev.org* and accessed on July 26, 2023.

Thiel, C., J. Bonner, J. Bush, D. Welsh, and N. Garud (2021) “Stripped of Agency: The Paradoxical Effect of Employee Monitoring on Deviance,” *Journal of Management* 49(2), pages 709–740.

United Nations Statistics Division. (2008). *Updated System of National Accounts 2008*. Accessed October 2022. http://unstats.un.org/unsd/nationalaccount/sna2008.asp.

U.S. Bureau of Labor Statistics (2022) *Occupational Outlook Handbook*, Insurance Underwriters; Claims Adjusters, Appraisers, Examiners, and Investigators; Information Clerks, at https://bls.gov/ooh, accessed in July 2023.

U.S. Bureau of Labor Statistics (2023) *Time Spent in Detailed Primary Activities and Percent of the Civilian Population Engaging in Each Activity, Averages Per Day Be Sex, Annual Averages 2003–2021* at https://www.bls.gov/tus/tables.htm, accessed in January 2023.

U.S. Census Bureau (2011, 2005, 2002, 2001, 1993, 1986, 1984, 1980, 1977, and 1970) "Statistical Abstract of the United States," posted at https://www.census.gov/library/publications/time-series/statistical_abstracts.html.

U.S. Customs and Border Protection (2023) "Registering a Foreign Registered Vehicle, Car, or Motorcycle in the United States with Department of Motor Vehicles," posted at help.cpb.gov and accessed July 26, 2023.

U.S. Department of Commerce (1975) *Historical Statistics of the United States: Colonial Times to 1970*, Washington, D.C.: Department of Commerce, September 1975, accessed October 2022.

U.S. Department of Justice (1933–2022) "Federal Bureau of Investigation Uniform Crime Report," posted at https://ucr.fbi.gov/crime-in-the-u.s and other places.

U.S. Department of Justice (1985) "Jail Inmates 1983," *Bureau of Justice Statistics Bulletin*.

U.S. Department of Transportation (2023) "Driver's Motor Vehicle Record," posted at https://ai.fmcsa.dot.gov/NewEntrant/MC/Content.aspx?nav=MVR and accessed July 25, 2023.

U.S. Women's Bureau (1944) "Your Questions as to Women in Work Industries: Types of Jobs," available at https://www.google.com/books/-edition/Your_Questions_as_to_Women_in_War_Indust/SiWJ5oATgjQC?hl=en&gbpv=0.

Vallet, M. and J. McCormick (2023) "How to Settle a Car Accident Without Filing a Claim," posted at *Insurance.com* on June 27, 2023, and accessed August 8, 2023.

Vayda, A., and I. Crespi (1981) *Public Acceptability of Highway Safety Countermeasures: Safe Driving Conformance Research*, U.S. Department of Transportation Publisher.

Vickrey, W. (1968) "Automobile Accidents, Tort Law, Externalities, and Insurance: An Economist's Critique," *Law and Contemporary Problems* 33(1968), pages 464–487.

Vives, X. (1984) "Duopoly Information Equilibrium: Cournot and Bertrand," *Journal of Economic Theory* 34(1), pages 71–94.

Waters, T., and P. Budetti (2004) "The National Practitioner Data Bank: Promoting Safety and Quality," *AMA Journal of Ethics Virtual Mentor* 6(3), pages 136–138.

Yao, Y. (2018) "Evolution of Insurance: A Telematics-Based Personal Auto Insurance Study," Honors Scholars Thesis posted at https://opencommons.uconn.edu/cgi/viewcontent.cgi?article=1563&context=srhonors_theses and accessed August 8, 2023.

Yawman, J.G. (1977) "Criminal Law—Obstructing Government Administration—The Physical Interference Requirement and a CB Radar Warning," *Buffalo Law Review* 26(4), pages 729–741.

Zheng, Z. (2019) "Jail Inmates in 2017," *Bureau of Justice Statistics Bulletin* NCJ 251774.