Economic Quarterly—Volume 104, Number 4—Fourth Quarter 2018—Pages 153–171

Why Do Platforms Use Ad Valorem Fees? Evaluating Two Alternative Explanations

Zhu Wang

Platforms that intermediate transactions between sellers and buyers have become increasingly important in the economy. People are familiar with, for example, online marketplaces (such as Amazon and eBay), payment platforms (such as Visa, MasterCard, and Paypal), and hotel booking sites (such as Booking.com and Expedia). However, there has been a great pricing puzzle associated with these platforms in that they almost universally rely on ad valorem fees, in which cases platforms charge sellers fees proportional to the transaction value plus sometimes small per-transaction fees. Given that these platforms do not incur significant costs that vary with transaction value, it is puzzling why ad valorem fees are so prevalently used.

In this article, we review two alternative explanations on this pricing puzzle. One theory, provided by Shy and Wang (2011) and others, emphasizes the vertical relation between the platform and the sellers. It is shown that in the case where the platform (i.e., the upstream) and the sellers (i.e., the downstream) both have market power (i.e., so-called "double marginalization")[1], the platform extracts a higher profit

Research Department, Federal Reserve Bank of Richmond. Email: zhu.wang@rich.frb.org. I thank Eric LaRose, John Weinberg, Alexander Wolman, and Russell Wong for helpful comments. The views expressed are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] In the industrial organization literature, double marginalization refers to the phenomenon in which different firms at different vertical levels in the supply chain (e.g., upstream and downstream) have their respective market powers and apply their own markups in prices. For example, consider that a firm with market power buys an input from another firm that also has market power. The producer of the input will

DOI: https://doi.org/10.21144/eq1040401

by using a proportional fee than using a per-transaction fee. Another explanation, offered by Wang and Wright (2017), instead focuses on the price discrimination angle. The key idea is that for a platform dealing with transactions of many different goods that vary widely in their costs and values, ad valorem fees serve as an efficient form of price discrimination that increases the platform's profit. While these two explanations provide alternative views, we will show that they indeed complement each other in explaining the ad valorem fee puzzle.

Our article contributes to a growing literature on platforms and their fee structures. In fact, besides the two theories analyzed in this article, there are additional (competing or complementary) views on ad valorem platform fees. For example, Loertscher and Niedermayer (2012) consider a mechanism design approach in an independent private values setup with privately informed buyers and sellers, in which an intermediary's optimal fees converge to linear fees as markets become increasingly thin. Muthers and Wismer (2013) show that if a platform can commit to proportional fees, this can reduce a hold-up problem that arises from the platform wanting to compete with sellers after they have incurred costs to enter the platform. Hagiu and Wright (forthcoming) provide a theory that ad valorem contracts align the incentives between upstream firms (principals) and downstream firms (agents), which allows the principal to achieve the same profits as if it could observe the demand shocks and control price.

The article is organized as follows. In Section 1, we first lay out two simple models that each justify one of the two explanations: double marginalization versus price discrimination. In Section 2, we then study a generalized model that accommodates both explanations. Our findings suggest that, in reality, platforms may choose a simple ad valorem fee schedule that addresses both double marginalization and price discrimination considerations. In Section 3, we apply the generalized model to a calibration exercise using data on DVD sales on Amazon and quantify the relative importance of the two explanations. Finally, Section 4 offers concluding remarks.

---

price above marginal cost when it sells the input to the other firm, who will then price above marginal cost again when they sell the final product that uses the input. This means the input is being marked up above marginal cost twice, which is called double marginalization.

1. TWO ALTERNATIVE EXPLANATIONS

In this section, we lay out two simple models that each highlight one of the two alternative explanations: double marginalization versus price discrimination.

Double Marginalization

We first study a model environment similar to Shy and Wang (2011), where double marginalization motivates the use of ad valorem fees.[2] Consider that a monopoly seller sells a good on a monopoly platform. The good is indexed by c, the per-unit cost of the good to the seller, which is known to everyone in the market. There is a unit mass of buyers, each of whom wants to purchase one unit of the good. The value of the good to a buyer is $c\,(1+b)$, where $b \geq 0$ is a parameter that the buyer draws.[3] We assume that $1+b$ is randomly distributed according to a cumulative distribution function F. Only buyers know their own b, while F is public information.

For illustrative purposes, we assume that F takes on a simple Pareto distribution

$$F\,(x) = 1 - x^{-\lambda}. \tag{1}$$

Accordingly, the number of transactions Q_c for the good c is the measure of buyers who obtain a nonnegative surplus from buying the good at price p_c, $\Pr\left(c\,(1+b) - p_c \geq 0\right)$. Therefore, the demand function for good c is

$$Q_c(p_c) = 1 - F\left(\frac{p_c}{c}\right) = \left(\frac{p_c}{c}\right)^{-\lambda}, \tag{2}$$

which has the constant elasticity λ. For the monopoly pricing problem to be well-defined, we require that $\lambda > 1$.

The platform incurs a cost of $d \geq 0$ per transaction, and it can potentially charge fees to either the buyer side or the seller side or

---

[2] In a similar vein, several studies (e.g., Foros et al. 2013; Gaudin and White 2014; and Johnson 2017) have explored the advantages of the so-called agency model used by mass retailers such as Amazon, where the retailer lets suppliers (i.e., sellers) set final prices and receive a share of the revenue, which is equivalent to using a percentage fee. Like Shy and Wang (2011), they also show that the revenue sharing used in the agency model has the advantage of mitigating double marginalization.

[3] A higher c (i.e., higher cost) implies in the model that the gains from trade are higher in expectation (due to the multiplicative connection between c and b). One interpretation for this specification, as shown in Wang and Wright (2017), is that such a platform reduces trading frictions, and as a result the value to buyers of using the platform (so that they can avoid the loss of using a less-efficient trade intermediary) is proportional to the cost or price of the goods traded. Note that the assumption $b \geq 0$ is an innocuous normalization because consumers whose valuation for a product is less than its cost can be ignored.

both. Regardless of which side is charged, the final price faced by buyers will reflect any fees, and the buyer treats these the same whether she faces them directly or through sellers. Due to this standard result on the irrelevance of the incidence of taxes across the two sides, we can assume without loss of generality that only the seller side is charged.

In terms of timing, the platform moves first and announces the fee schedule it would charge the seller. Taking the fee schedule as given, the seller then decides the price of the good. Finally, buyers make purchase decisions.

Given the model setup, we are interested in the following question: If the platform can choose among a per-transaction fee, a proportional fee, or a mix of both fees, what type of fee schedule would the platform prefer?

To answer the question, we consider that the platform decides on an affine fee schedule, $T(p_c) = t_0 + t_1 p_c$, which covers all the possibilities listed above. We assume that the platform cannot subsidize the seller to operate by setting $t_0 < 0$. Doing so is likely to create an adverse incentive for which the seller could just collect t_0 but not sell anything real. This imposes the requirement that $t_0 \geq 0$.

The model can be solved backward. Because the platform would make its fee decision by incorporating the seller's response, we solve the seller's problem first. The seller, taking the affine fee schedule (t_0, t_1) charged by the platform as given, would choose p_c to maximize her profit:

$$\max_{p_c} (p_c - c - t_0 - t_1 p_c) \left(\frac{p_c}{c}\right)^{-\lambda},$$

which implies

$$p_c^* = \frac{\lambda (c + t_0)}{(\lambda - 1)(1 - t_1)}. \quad (3)$$

Anticipating the seller's pricing decision p_c^*, the platform would then choose t_0 and t_1 to solve

$$\pi = \max_{t_0, t_1} \; (t_0 + t_1 p_c^* - d) \left(\frac{p_c^*}{c}\right)^{-\lambda}$$

subject to the constraint $t_0 \geq 0$**.** We can verify that the constraint $t_0 \geq 0$ is binding at the maximum, so the optimal affine fee schedule is just a proportional fee:

$$t_0 = 0, \quad t_1 = \frac{c + d(\lambda - 1)}{\lambda c + d(\lambda - 1)}. \quad (4)$$

Given that $\lambda > 1$, we know $1 > t_1 > 0$.

This simple model yields several useful findings. First, in the presence of double marginalization (i.e., when both the platform and the seller have market power), the platform strictly prefers a proportional fee to a per-transaction fee. Note that the use of a proportional fee allows the platform to mitigate, but not eliminate, double marginalization. In fact, if the seller side has no market power (or the platform owns the seller), the platform, being the single monopoly in the market, would earn an even higher profit and would be indifferent with a proportional fee or a per-transaction fee, as we will show in the analysis coming next. Second, to implement the optimal proportional fee, the platform needs to know c unless the marginal cost d of the platform is zero, in which case the platform has a simple formula $t_1 = 1/\lambda$. Considering that d is typically small in reality, a platform may use $t_1 = 1/\lambda$ as a good proxy even if it has no knowledge of c.

The model above serves as a simple illustrative example. As shown in Shy and Wang (2011) and others, the result holds in more general settings, including the cases where sellers engage in Cournot competition with or without free entry.[4]

Price Discrimination

In contrast to the double marginalization explanation, we now study an alternative model proposed by Wang and Wright (2017) where price discrimination motivates the use of ad valorem fees. In doing so, we consider the same model setup as above except for two things: (i) a variety of goods is being sold on the platform, with the costs c differing widely across goods; and (ii) for each good c, there are multiple sellers who engage in Bertrand competition, so sellers have no market power.[5] The rest of the model specification remains unchanged—for each good c, there is a unit mass of buyers each of whom wants to purchase one unit of the good. Buyers draw their benefit $1 + b$ from a simple Pareto distribution, and as a result sellers face constant-elasticity demand. The platform considers charging sellers an affine fee schedule, $T(p_c) = t_0 + t_1 p_c$, subject to the constraint $t_0 \geq 0$.

Assume c takes on a finite number of distinct values in the set of C. The probability distribution of c on C is denoted g_c, with $\sum_{c \in C} g_c = 1$. As before, we solve the sellers' problem first. For each good c, taking

[4] Cournot competition refers to an oligopoly market structure in which multiple firms producing a homogeneous product compete by choosing outputs independently and simultaneously. Assuming a fixed number of Cournot sellers, Shy and Wang (2011) show that the platform earns a higher profit by using a proportional fee than a per-transaction fee. Miao (2013) shows that the result continues to hold under free entry of sellers.

[5] Bertrand competition is a model of competition in which multiple firms producing a homogeneous product compete by setting prices simultaneously and consumers want to buy everything from a firm with a lower price.

the affine fee schedule as given, Bertrand sellers compete by setting the lowest possible price just to break even, so that

$$p_c^* = c + t_0 + t_1 p_c^* \Longrightarrow p_c^* = \frac{c + t_0}{1 - t_1}.$$

Anticipating sellers' pricing decisions, the platform would then choose t_0 and t_1 to solve

$$\Pi = \max_{t_0, t_1} \sum\nolimits_{c \in C} g_c \left[(t_0 + t_1 p_c^* - d) \left(\frac{p_c^*}{c} \right)^{-\lambda} \right]. \quad (5)$$

To derive the solution to (5) intuitively, we first consider the hypothetical scenario where the platform could perfectly observe the cost and valuation for each good c and set a different optimal fee (t_0, t_1) for each as follows:

$$\Pi_c = \max_{t_0, t_1} (t_0 + t_1 p_c^* - d) \left(\frac{p_c^*}{c} \right)^{-\lambda},$$

which is equivalent to solving

$$\Pi_c = \max_{t_0, t_1} \left(\frac{t_0 + ct_1}{1 - t_1} - d \right) \left(1 + \frac{t_0 + ct_1}{(1 - t_1)c} \right)^{-\lambda}.$$

The first-order condition implies a unique value of $\frac{t_0^* + ct_1^*}{1 - t_1^*}$ such that

$$\frac{t_0^* + ct_1^*}{1 - t_1^*} = \frac{c + \lambda d}{\lambda - 1}, \quad (6)$$

which could be potentially consistent with different fee schedules (t_0^*, t_1^*). For example, the optimal fee could be a pure per-transaction fee or a pure proportional fee, but those fee schedules have to depend on c. However, one can verify that there is a unique affine fee

$$t_0^* = d; \quad t_1^* = \frac{1}{\lambda} \quad (7)$$

that also satisfies the condition (6), but the fee schedule does not depend on c. This means that the affine fee (7) maximizes the platform's overall profit (5) without requiring the platform to keep track of the goods traded.

This yields several new findings. First, for a given good, when the cost c is known to the platform and sellers have no market power, the platform is indifferent between charging a proportional fee and a per-transaction fee. This contrasts our finding above that a proportional fee is strictly preferred to a per-transaction fee when sellers do have market power. Second, the platform can maximize profit by implementing the affine fee (7) without conditioning on c, which is a great advantage. There are often a large number of goods being traded on a platform, and the platform may not be able to track each good's cost and value. In this case, using the affine fee (7) requires no information of c, so it can be easily used by the platform. This results in optimal price discrimination in the sense that charging ad valorem fees (7) allows the platform to achieve the same level of profit that could be obtained under third-degree price discrimination as if the platform could perfectly observe the cost and valuation for each good traded. Finally, note that the optimal affine fee (7) has a per-transaction term $t_0^* > 0$ only if the platform incurs a positive marginal cost d; otherwise, a proportional fee $t_1 = 1/\lambda$ is optimal. Again, considering that d is typically small in reality, a simple proportional fee $t_1 = 1/\lambda$ can be a good proxy in practice.

The model is a simple illustrative example. Wang and Wright (2017) show the result holds broadly, including the demand takes more general functional forms or involves unobserved random variations.

2. A GENERALIZED ANALYSIS

The two theories noted above provide alternative justifications for the use of ad valorem fees by platforms. However, these two theories are not necessarily exclusive to each other. In this section, we provide a generalized analysis that accommodates both explanations. We show in reality a platform can choose a simple ad valorem fee that addresses both double marginalization and price discrimination considerations. The analysis and results in this section draw heavily from the online appendix of Wang and Wright (2017).

In the generalized analysis, we consider a variety of different goods being traded on a platform. We suppose that for each good there are $n_c \geq 1$ identical quantity-setting sellers on the platform (i.e., Cournot competitors). This covers different intensities of seller competition, including the two special cases discussed in Section 1: when $n_c = 1$, a good is sold by a monopoly seller; when $n_c \to \infty$, sellers are perfectly competitive. As before, each seller obtains the goods at a unit cost c and sells them at a retail price p_c.

On the demand side, we assume as before that the value of good c to a buyer drawing the benefit parameter $b \geq 0$ is $c(1+b)$. To generalize the analysis, we now consider that $1+b$ is distributed according to the broad family of generalized Pareto distributions (GPD), of which the simple Pareto distribution is a special case. Accordingly, the cumulative distribution function F is defined as

$$F(x) = 1 - (1 + \lambda(\sigma - 1)(x - 1))^{\frac{1}{1-\sigma}}, \tag{8}$$

with $\lambda > 0$ being the scale parameter and $\sigma < 2$ being the shape parameter. Only buyers know their own b, while F is public information.

Note that the generalized Pareto distribution implies the demand functions for sellers on the platform are defined by the class of demands with constant curvature of inverse demand[6]

$$Q_c(p_c) = 1 - F\left(\frac{p_c}{c}\right) = \left(1 + \frac{\lambda(\sigma - 1)(p_c - c)}{c}\right)^{\frac{1}{1-\sigma}}. \quad (9)$$

The constant σ is the curvature of inverse demand, defined as the elasticity of the slope of the inverse demand with respect to quantity. When $\sigma < 1$, the support of F is $[1, 1 + 1/\lambda(1 - \sigma)]$ and it has increasing hazard. Accordingly, the implied demand functions $Q_c(p_c)$ are log-concave and include the linear demand function ($\sigma = 0$) as a special case. Alternatively, when $1 < \sigma < 2$, the support of F is $[1, \infty)$, and it has decreasing hazard. The implied demand functions are log-convex and include the constant elasticity demand function ($\sigma = 1 + 1/\lambda$) as a special case. When $\sigma = 1$, F captures the left-truncated exponential distribution $F(x) = 1 - e^{-\lambda(x-1)}$ on the support $[1, \infty)$, with a constant hazard rate λ. This implies the exponential (or log-linear) demand $Q_c(p_c) = e^{-\frac{\lambda(p_c - c)}{c}}$.

Taking as given that demand belongs to the generalized Pareto class, we allow c to take on potentially many different values in $[c_L, c_H]$, with the set of all such values being denoted C. The cumulative distribution of c on C is denoted G, and g_c is the probability corresponding to the realization c.

The platform incurs a cost of $d \geq 0$ per transaction. Without loss of generality, we assume that the platform only charges the seller side to maximize its profit.

Below, in Section 2.1, as a benchmark, we first derive the platform's optimal affine fee in a setting with generalized Pareto demand and Bertrand sellers (or equivalently, sellers engage in Cournot competition, but the number of sellers goes to infinity). This extends the results we derived in Section 1.2, and we name the resulting fee schedule the "Bertrand affine fee." In this general case, as in Section 1.2, the Bertrand affine fee achieves optimal price discrimination given that sellers have no market power. In Section 2.2, we show that in a setting where sellers have market power and engage in Cournot competition, the Bertrand affine fee continues to do well. Particularly, we show that without knowing each good's cost and how many sellers are

[6] This class of demands has been considered by Bulow and Pfleiderer (1983), Aguirre et al. (2010), Bulow and Klemperer (2012), and Weyl and Fabinger (2013), among others.

competing, the platform can continue to use the Bertrand affine fee and earn a higher profit than if it knew everything and set the optimal per-transaction fee for each good. This is because the Bertrand affine fee now achieves more than price discrimination; it also mitigates double marginalization. We then derive analytical results for the case $d = 0$ and show that while the Bertrand affine fee is not necessarily the optimal affine fee when sellers have market power, it can be very close. Therefore, in practice, a platform can implement the Bertrand affine fee as a good proxy.

Bertrand Affine Fee

We start with deriving the Bertrand affine fee. Consider that the platform charges sellers the fee schedule $T(p_c)$. Assuming that sellers engage in Bertrand competition, the price p_c for good c solves

$$p_c = c + T(p_c). \tag{10}$$

Accordingly, the platform's profit is $\Pi_c = (T(p_c) - d)\, Q_c(p_c)$ for good c, where $Q_c(p_c)$ is given by (9). The platform's problem is to choose $T(p_c)$ to maximize

$$\Pi = \sum_{c \in C} g_c \Pi_c. \tag{11}$$

In Wang and Wright (2017), it is shown that the optimal fee schedule is affine, given by

$$T(p_c) = \frac{\lambda d}{1 + \lambda(2 - \sigma)} + \frac{p_c}{1 + \lambda(2 - \sigma)}, \tag{12}$$

which maximizes (11).[7] Similar to our finding in Section 1.2, while the affine fee (12) does not condition on c, it achieves optimal price discrimination. To see this, note that the solution in (12) is equivalent to the platform charging the optimal per-transaction fee

$$T_c^* = \frac{\lambda d + c}{\lambda(2 - \sigma)} \tag{13}$$

for each different good c, which would be possible if the platform could identify each good c and set its optimal per-transaction fee accordingly.

Our result in Section 1.2 is a special case of the Bertrand affine fee given by (12), with $\sigma = 1 + 1/\lambda$. In the general case, the platform's optimal affine fee again has a fixed per-transaction component only if

[7] With this model setting, the optimal platform fee schedule is affine and does not condition on c if and only if the distribution of buyers' benefits F is the generalized Pareto distribution. See Wang and Wright (2017) for a detailed proof.

there is a positive cost to the platform of handling each transaction (i.e., $d > 0$). Given $\lambda > 0$ and $\sigma < 2$, the fee schedule is increasing (higher prices imply higher fees are paid) but with a slope less than unity (this implies (10) has a unique solution for any given $c > 0$). The result in (12) also implies the platform can maximize its profit without tracking each individual good c or knowing the distribution G of goods that are traded.

Seller Market Power and Bertrand Affine Fee

We now study the platform's fee setting when sellers do have market power. We will show in the case of Cournot sellers, the platform can continue to use the Bertrand affine fee, which not only addresses the price discrimination, but also mitigates double marginalization. As a result, it leads to a higher platform profit than using optimal per-transaction fees.

Optimal per-transaction fees

To start, we consider the problem of a platform with full information on c (i.e., each good's cost) and n_c (i.e., the number of Cournot sellers) setting an optimal per-transaction fee for each good.

Suppose the platform charges a per-transaction fee T_c for good c. Let $q_{c,i}$ denote the output sold by seller i for good c. Each seller i sets $q_{c,i}$ taking the output by competing sellers $q_{c,-i} = Q_c - q_{c,i}$ as given and maximizes its profit $(p_c - c - T_c)\, q_{c,i}$. Assuming F follows the GPD distribution (8), the total demand for good c is given by (9), which implies that the inverse demand is

$$p_c = c\left(1 + \frac{Q_c^{1-\sigma} - 1}{\lambda(\sigma - 1)}\right).$$

Therefore, an individual seller's profit maximization problem is

$$\max_{q_{c,i}} c\left(1 + \frac{(q_{c,-i} + q_{c,i})^{1-\sigma} - 1}{\lambda(\sigma - 1)}\right) q_{c,i} - (c + T_c)q_{c,i}.$$

The first-order condition for good c is

$$c\left(1 + \frac{(q_{c,-i} + q_{c,i})^{1-\sigma} - 1}{\lambda(\sigma - 1)}\right) = q_{c,i}\left(\frac{c(q_{c,-i} + q_{c,i})^{-\sigma}}{\lambda}\right) + c + T_c.$$

In a symmetric Cournot equilibrium, $q_{c,i} = q_c$ for every seller, so the total sellers' output is $Q_c = n_c q_c$. We can then rewrite the first-order condition as

$$\frac{c(n_c q_c)^{1-\sigma} - c}{\lambda(\sigma - 1)} = \frac{c(n_c q_c)^{1-\sigma}}{n_c \lambda} + T_c$$

and derive

$$Q_c = n_c q_c = \left(\frac{cn_c + \lambda(\sigma - 1)T_c n_c}{cn_c - (\sigma - 1)c} \right)^{\frac{1}{1-\sigma}}. \tag{14}$$

Accordingly, the price of good c is

$$p_c = c\left(1 + \frac{\lambda T_c n_c + c}{\lambda c(n_c + 1 - \sigma)}\right) = \frac{T_c n_c}{(n_c + 1 - \sigma)} + \frac{1 + (n_c + 1 - \sigma)\lambda}{(n_c + 1 - \sigma)\lambda} c. \tag{15}$$

The platform takes (14) as given and maximizes its profit by setting a per-transaction fee for good c as follows

$$\max_{T_c} \ (T_c - d) \left(\frac{cn_c + \lambda(\sigma - 1)T_c n_c}{cn_c - (\sigma - 1)c} \right)^{\frac{1}{1-\sigma}}.$$

The first-order condition implies the optimal per-transaction fee T_c^f:

$$T_c^f = \frac{\lambda d + c}{\lambda\,(2 - \sigma)}, \tag{16}$$

which is the same optimal per-transaction fee that we derive in the Bertrand seller setting (13). The optimal per-transaction fee does not depend on the number of sellers and so also holds for a monopoly seller. Note that to ensure a meaningful solution (i.e. $T_c^f > d$) , it is required that

$$d\,(\sigma - 1) + \frac{c}{\lambda} > 0. \tag{17}$$

This is satisfied for the GPD demand specification: When demand is log-linear or log-convex, the GPD specification requires that $\sigma \geq 1$ so the condition in (17) holds. When demand is log-concave, the GPD specification requires that $\sigma < 1$ and $d < \frac{c}{\lambda(1-\sigma)}$, so the condition in (17) again holds.

Substituting (16) into (14) and (15), we get

$$p_c = \frac{n_c d}{(2 - \sigma)(n_c + 1 - \sigma)} + \frac{n_c + (2 - \sigma) + (2 - \sigma)(n_c + 1 - \sigma)\lambda}{(2 - \sigma)(n_c + 1 - \sigma)\lambda} c, \tag{18}$$

and

$$Q_c = \left(\frac{\lambda(\sigma - 1)n_c d + cn_c}{(2 - \sigma)(cn_c - (\sigma - 1)c)} \right)^{\frac{1}{1-\sigma}}. \tag{19}$$

As a result, the platform profit from good c is

$$\pi_c = \left(\frac{(\sigma - 1)d}{2 - \sigma} + \frac{c}{(2 - \sigma)\lambda} \right) \left(\frac{\lambda(\sigma - 1)n_c d + cn_c}{(2 - \sigma)(cn_c - (\sigma - 1)c)} \right)^{\frac{1}{1-\sigma}}.$$

Comparing Bertrand affine fee and optimal per-transaction fees

We now compare Bertrand affine fee and optimal per-transaction fees in the Cournot seller setting.

Consider Cournot sellers facing an affine fee schedule $T(p_c) = t_0 + t_1 p_c$ for each transaction. With GPD demand, the sellers' problem is to choose $q_{c,i}$ to maximize

$$((1-t_1)p_c - c - t_0)q_{c,i}, \tag{20}$$

where

$$p_c = c\left(1 + \frac{(q_{c,-i} + q_{c,i})^{1-\sigma} - 1}{\lambda(\sigma-1)}\right). \tag{21}$$

In a symmetric Cournot equilibrium, $q_{c,i} = q_c$ for every seller, so the total sellers' output is $Q_c = n_c q_c$. The first-order condition then requires

$$(1-t_1)c\left(\frac{1}{\lambda(\sigma-1)} - 1\right) + c + t_0 = \frac{(1-t_1)cQ_c^{1-\sigma}}{\lambda}\left(\frac{1}{\sigma-1} - \frac{1}{n_c}\right). \tag{22}$$

Substituting the Bertrand affine fee from equation (12) into (22) gives the same price and output for a given c as we found above in (18) and (19) for the full information case. That is, the price and output for each good are identical to that implied by the optimal per-transaction fee (16). However, the per-transaction fee for good c implied by the Bertrand affine fee is now

$$\begin{aligned} T^*(p_c) &= t_0 + t_1 p_c = \left(\frac{\lambda}{1+(2-\sigma)\lambda} + \frac{n_c}{(1+(2-\sigma)\lambda)(2-\sigma)(n_c+1-\sigma)}\right)d \\ &\quad + \left(\frac{1}{1+(2-\sigma)\lambda}\right)\left(\frac{n_c + (2-\sigma) + (2-\sigma)(n_c+1-\sigma)\lambda}{(2-\sigma)(n_c+1-\sigma)\lambda}\right)c, \end{aligned}$$

which is strictly higher than the fee in (16) if and only if the condition (17) holds. This implies the platform earns a higher profit using the Bertrand affine fee than if it used the optimal per-transaction fee for each different good assuming full information. This result holds for any $n_c \geq 1$ and so also holds for monopoly sellers.

This result shows that the Bertrand affine fee can be used in this setting to solve the price discrimination problem. It delivers the same price and output for each good without using any information on each good's cost. At the same time, the Bertrand affine fee generates a higher profit for the platform because it mitigates the double marginalization problem associated with using the optimal per-transaction fee for each good, allowing the platform to collect a higher fee from each good while achieving the same level of final price and output.

Comparing Bertrand affine fee and optimal affine fee

We have so far shown that Bertrand affine fee profit dominates per-transaction fee when sellers have market power. In this section, assuming $d = 0$, we show that the Bertrand affine fee schedule (12) is indeed very close to the optimal affine fee schedule under Cournot sellers.[8] Note that given $d = 0$, the Bertrand affine fee (12) implies the proportional fee schedule

$$T^* (p_c) = \left(\frac{1}{1 + (2 - \sigma)\lambda} \right) p_c. \tag{23}$$

We can then check whether this is the optimal affine fee schedule under Cournot sellers.

Consider a platform maximizing its profit by using an affine fee schedule $t_0 + t_1 p_c$. As before, we assume that the platform cannot subsidize sellers to operate by setting $t_0 < 0$. This imposes the requirement that $t_0 \geq 0$.

Cournot sellers take the platform affine fee schedule $T(p_c) = t_0 + t_1 p_c$ as given for each transaction. As shown above, with a GPD demand, the sellers' problem is given by (20) and (21), and the first-order condition for seller's profit-maximizing problem is given by (22).

Anticipating sellers' responses, the platform then solves the following problem:

$$\pi = \max_{t_0, t_1} \sum_c g_c \ (t_0 + t_1 p_c) \left(1 - F\left(\frac{p_c}{c} \right) \right)$$

subject to the constraint $t_0 \geq 0$ as well as the conditions

$$p_c = c \left(1 + \frac{Q_c^{1-\sigma} - 1}{\lambda(\sigma - 1)} \right) \tag{24}$$

and

$$(1 - t_1)c \left(\frac{1}{\lambda(\sigma - 1)} - 1 \right) + c + t_0 = \frac{(1 - t_1)cQ_c^{1-\sigma}}{\lambda} \left(\frac{1}{\sigma - 1} - \frac{1}{n_c} \right), \tag{25}$$

where (24) is given by the GPD demand and (25) is the first-order condition (22). We can verify that the constraint $t_0 \geq 0$ is binding at the maximum, so the optimal affine fee schedule is also just a proportional fee schedule. Moreover, given that $t_0 = 0$, p_c/c does not depend on c, so the platform can solve for the optimal t_1 without knowing the

[8] If $d > 0$, the results will depend on the distribution of c. We discuss this case in Section 3.

distribution of c. The first-order condition on t_1 requires

$$\begin{aligned} & (1+\lambda(1-\sigma))(1-t_1-t_1\lambda(1-\sigma))(1-t_1)-t_1\lambda(1+\lambda(1-\sigma)) \\ = \; & n_c\left(\frac{t_1}{1-t_1}\lambda^2(2-\sigma)-\lambda\right). \qquad (26) \end{aligned}$$

The optimal proportional fee implied by (26) is in general not equal to the proportional fee implied by (23), but based on an examination of some common demand functions, it is very close and so are the profits, as discussed below.

Consider first the case of constant elasticity demand, where $\sigma = 1+\frac{1}{\lambda}$ and $\lambda > 1$. In this case, both (23) and (26) yield $t_1 = 1/\lambda$ and so have identical profits. Thus, in this case, the Bertrand affine fee coincides with the optimal affine fee schedule. This result confirms our findings in Sections 1.1 and 1.2 that when $d = 0$, the optimal affine fee under double marginalization (i.e., $t_0 = 0$, $t_1 = 1/\lambda$) coincides with that which achieves optimal price discrimination (which is again $t_0 = 0$, $t_1 = 1/\lambda$).

Next, consider the case of exponential demand where $\sigma = 1$. Then (26) implies the optimal proportional fee satisfies

$$(1-t_1)^3+\lambda(1-t_1)(n_c-t_1) = n_c t_1\lambda^2,$$

which has a unique solution. In contrast, (23) implies the proportional fee

$$t_1 = \frac{1}{1+\lambda}.$$

The two fees are not exactly equal, but they are very close. For the empirically meaningful range where the proportional term t_1 of the Bertrand affine fee satisfies $t_1 \leq 50$ percent (or equivalently, $\lambda \geq 1$), the Bertrand affine fee can recover more than 98.5 percent of the profit under the optimal affine fee schedule when all sellers are monopolists (so $n_c = 1$ for all c). Moreover, the profit gap between using the Bertrand affine fee and using the optimal affine fee schedule decreases monotonically in n_c, and the two converge as the number of Cournot sellers gets large.

Finally, consider the case of linear demand where $\sigma = 0$. Then (26) implies that the optimal proportional fee satisfies

$$\begin{aligned} & (1-t_1)^2(1+\lambda)(1-t_1-t_1\lambda)-t_1(1-t_1)\lambda(1+\lambda) \\ = \; & n_c\left(2t_1\lambda^2-\lambda(1-t_1)\right), \end{aligned}$$

which has a unique solution. In contrast, (23) implies the proportional fee

$$t_1 = \frac{1}{1+2\lambda}.$$

For the empirically meaningful range where the proportional term t_1 of the Bertrand affine fee satisfies $t_1 \leq 50$ percent (or equivalently, $\lambda \geq 0.5$), the Bertrand affine fee can recover more than 97.5 percent of the profit under the optimal affine fee schedule when all sellers are monopolists (so $n_c = 1$ for all c). Again, the profit gap between using the Bertrand affine fee schedule and using the optimal affine fee decreases monotonically in n_c, and the two converge as the number of Cournot sellers gets large.

The findings in Section 2 are summarized below.

Assume that the demand functions for sellers on the platform belong to the generalized Pareto class with $\lambda > 0$ *and* $\sigma < 2$ *and that for each good* c *there are* $n_c \geq 1$ *identical sellers that set quantities. Then we have the following results:*
(i) the platform obtains a higher profit using the Bertrand affine fee than if it sets the optimal per-transaction fee for each good;
(ii) if sellers face constant elasticity demand ($\sigma = 1 + \frac{1}{\lambda}$ *and* $\lambda > 1$*) and* $d = 0$*, the Bertrand affine fee is the optimal affine fee schedule;*
(iii) if sellers face exponential demand ($\sigma = 1$*),* $\lambda > 1$*, and* $d = 0$*, the Bertrand affine fee can recover more than* 98.5 *percent of the profit under the optimal affine fee schedule;*
(iv) if sellers face linear demand ($\sigma = 0$*),* $\lambda > 0.5$*, and* $d = 0$*, the Bertrand affine fee can recover more than* 97.5 *percent of the profit under the optimal affine fee schedule.*

3. A QUANTITATIVE EXERCISE

Finally, we may consider the general case in which $d > 0$ and compare the platform's profit from the Bertrand affine fee (12) with its profit from the optimal fee schedules, including nonlinear ones. This exercise was carried out in detail in Wang and Wright (2017), and we summarize the findings here.

Once we allow for a nonlinear fee schedule, the optimal fee schedule will depend on the distribution of goods $G(c)$. This is also true for the optimal affine fee schedule once we allow $d > 0$. Therefore, to proceed, one needs to assume some realistic distribution for c and calculate the profitability of different fee schedules numerically. Wang and Wright (2017) use the distribution based on fitting a log-normal distribution to the actual distribution of sales obtained from sales ranks of DVDs sold on Amazon.[9] It is assumed that sellers face constant elasticity

[9] Using a web robot, Wang and Wright (2017) collected data on every DVD listed under "Movies & TV" on Amazon's marketplace in January 2014. Given shipping fees are often not included in the listed price, the focus is on the items where the listed

demand, and $d = 1.35$ and $\sigma = 1.15$ so that the calibrated Bertrand fee schedule matches the actual fee schedule used by Amazon for DVDs (which is \$1.35+15 percent). Sellers are assumed to be monopolists (i.e., $n_c = 1$).[10]

With these assumptions, it is found that the platform obtains a profit of 0.383 with a fixed per-transaction fee (i.e., without any price discrimination).[11] If the platform could observe each different good sold by the sellers, it could do better setting the per-transaction fee that is optimal for each good c. This increases its profit by 17.7 percent to 0.457, which represents the gain due to price discrimination. Moreover, the benefits of price discrimination can be obtained by using the Bertrand fee schedule, which does not require any information on the values of c and has the added benefit of mitigating double marginalization. Indeed, the platform can increase its profit to 0.537, or a further 16.3 percent, by using the Bertrand fee schedule. Taking into account that sellers are monopolists and the particular distribution of c, the platform can increase its profit by a further 1.5 percent by moving to the optimal affine fee schedule.

Finally, Wang and Wright (2017) obtain the platform's profit for the optimal nonlinear fee schedule, which comes from solving for the optimal polynomial fee schedule of degree k, starting with $k = 1$ (the affine fee schedule) and considering higher and higher k until the platform's profit no longer increases. Compared with the optimal affine fee schedule, moving to the optimal nonlinear fee schedule only increases the platform's profit by a further 1.3 percent. The results are summarized in Table 1. The table also shows the results from repeating the exercise with linear demand.

Quantitatively, the results show that the platform loses little from restricting fee schedules to affine fee schedules or indeed the Bertrand affine fees. In the constant-elasticity demand case, price discrimination and double marginalization have similar quantitative effects on justifying the platform's use of the Bertrand affine fee: using the Bertrand

price included free shipping, resulting in a sample with 191,280 distinct items. The data collected include the title, unique ASIN number identifying the DVD, the price, and sales rank of each DVD. Given that the sale of each DVD is not directly observable, a power law is used to infer it from the sales rank data, so $Q_c = aR_c^{-\phi}$, where Q_c is the estimated sale of an item c and R_c is the corresponding sales rank. The parameter a does not affect the analysis, so it is normalized as $a = 1$. It is assumed $\phi = 1.7$, which is the number suggested by an experimental study on DVD sales on Amazon.

[10] This quantitative exercise evaluates how well the Bertrand affine fee performs under Cournot sellers. Assuming monopoly sellers is the most extreme alternative to Bertrand competition, so it provides the most conservative results.

[11] Note that because the sales of DVDs are inferred from data on sales ranks with scale normalized, only the relative (but not the absolute) value of the platform profit is meaningful for comparison.

Table 1 Profitability of Different Types of Fees

Fee schedule	Constant-elasticity demand Profit	Constant-elasticity demand Profit gain	Linear demand Profit	Linear demand Profit gain
Fixed per-transaction fee	0.383		0.632	
Per-trans. fee varying by good	0.457	17.7%	0.966	42.4%
Bertrand affine fee	0.537	16.3%	1.039	7.2%
Optimal affine fee	0.545	1.5%	1.041	0.3%
Optimal nonlinear fee	0.553	1.3%	1.043	0.1%
Total profit gain (%)		36.7%		50.0%

affine fee increases platform's profit by 33.8 percent compared with using a fixed per-transaction fee, where 17.7 percent comes from price discrimination and 16.3 percent comes from mitigating double marginalization. In the linear demand case, price discrimination's effect turns out higher than double marginalization: using the Bertrand affine fee increases platform's profit by 49.6 percent compared with using a fixed per-transaction fee, where 42.4 percent comes from price discrimination and 7.2 percent comes from mitigating double marginalization.

4. CONCLUSION

In this article, we review two alternative explanations for why platforms use ad valorem fees: double marginalization versus price discrimination. Using a generalized framework, we show that the two theories complement each other in explaining this pricing puzzle, and their relative importance is quantified in a calibration exercise.

Our findings set the stage for normative analysis. Given that platforms do not incur significant costs that vary with transaction prices, there have been policy concerns regarding their use of ad valorem fees. Using the framework discussed in this article, one could evaluate the welfare consequences of regulating platforms' use of ad valorem fees. In fact, Shy and Wang (2011) and Wang and Wright (forthcoming) have shown that banning platforms' use of ad valorem fees tends to reduce social welfare in the presence of double marginalization or price discrimination. Therefore, caution ought to be taken when policymakers consider intervening in platforms' use of ad valorem pricing.

REFERENCES

Aguirre, Iñaki, Simon Cowan, and John Vickers. 2010. "Monopoly Price Discrimination and Demand Curvature." *American Economic Review* 100 (September): 1601–15.

Bulow, Jeremy, and Paul Pfleiderer. 1983. "A Note on the Effect of Cost Changes on Prices." *Journal of Political Economy* 91 (February): 182–85.

Bulow, Jeremy, and Paul Klemperer. 2012. "Regulated Prices, Rent Seeking, and Consumer Surplus." *Journal of Political Economy* 120 (February): 160–86.

Foros, Øystein, Hans Jarle Kind, and Greg Shaffer. 2013. "Turning the Page on Business Formats for Digital Platforms: Does Apple's Agency Model Soften Competition?" Working Paper.

Gaudin, Germain, and Alexander White. 2014. "On the Antitrust Economics of the Electronic Books Industry." Dusseldorf Institute for Competition Economics Discussion Paper 147 (May).

Hagiu, Andrei, and Julian Wright. Forthcoming. "The Optimality of Ad Valorem Contracts." *Management Science*.

Johnson, Justin. 2017. "The Agency Model and MFN Clauses." *Review of Economic Studies* 84 (July): 1151–85.

Loertscher, Simon, and Andras Niedermayer. 2012. "Fee-setting Mechanisms: On Optimal Pricing by Intermediaries and Indirect Taxation." Governance and the Efficiency of Economic Systems Discussion Paper 434 (October).

Miao, Chun-Hui. 2013. "Do Card Users Benefit from the Use of Proportional Fees?" *Review of Network Economics* 12 (September): 323–41.

Muthers, Johannes, and Sebastian Wismer. 2013. "Why Do Platforms Charge Proportional Fees? Commitment and Seller Participation." Working Paper.

Shy, Oz, and Zhu Wang. 2011. "Why Do Payment Card Networks Charge Proportional Fees?" *American Economic Review* 101 (June): 1575–90.

Wang, Zhu and Julian Wright. 2017. "Ad Valorem Platform Fees, Indirect Taxes, and Efficient Price Discrimination." *RAND Journal of Economics* 48 (Summer): 467–84.

Wang, Zhu, and Julian Wright. Forthcoming. “Should Platforms Be Allowed to Charge Ad Valorem Fees?” *Journal of Industrial Economics*.

Weyl, Glen, and Michal Fabinger. 2013. “Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition.” *Journal of Political Economy* 121 (June): 528–83.

Economic Quarterly—Volume 104, Number 4—Fourth Quarter 2018—Pages 173–189

What Can We Learn from Online Wage Postings? Evidence from Glassdoor

Marios Karabarbounis and Santiago Pinto

Tracking economic activity and interpreting economic phenomena are the most basic functions of economic research. However, obtaining an accurate description of the economy—in the form of economic data—is a challenging endeavor. Basic economic variables such as gross domestic product, consumption expenditures, investment, real wages, and others are available at the aggregate level. They are useful for time-series analysis but not to study issues such as wage or wealth inequality. To study heterogeneity, economists rely on household-level data from sources like the Panel Study of Income Dynamics (PSID), the Survey of Consumer Finances, or the Consumption Expenditure Survey. However, these typically include only a sample of the population and are often subject to measurement errors.

For this reason, economists have recently started to incorporate alternative sources that provide granular or disaggregated data, for example, websites that offer job and recruiting services. A growing number of sites give online information about different jobs around the US and worldwide. The websites collect, at the same time, personal and financial data from users. In light of this recent phenomenon, a question naturally arises: Can economists view these websites as a reliable source of new information?[1]

Contact information: marios.karabarbounis@rich.frb.org and santiago.pinto@rich.frb.org. We thank Andrew Chamberlain and Glassdoor for generously providing us with the data. For useful comments, we thank Andreas Hornstein, John Bailey Jones, and John Weinberg. We also thank Mohamed Abbas Roshanali for outstanding research assistance. Any opinions expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] Kudlyak et al. (2013) employ information from an online posting website to analyze how job seekers direct their applications over the course of job search. Hershbein

In this paper, we take a small step toward addressing this issue. We present information from millions of salaries from Glassdoor.com (henceforth, Glassdoor), a leading job website that helps people find jobs and companies recruit employees. To use the service, registered users are asked, among other things, to report their current occupation title (job position), company, salary (in addition to other payment schemes), location, and level of experience. In return, users can get access to user-generated content including ratings and reviews of companies, interview questions, CEO approval rates, and summary statistics of salaries for job positions within each company.

We compare the salary information in Glassdoor with two other widely used sources. The first is the Quarterly Census for Employment and Wages (QCEW) published by the US Census Bureau. QCEW provides information on salaries and employment at various industry and geographic area levels. The second is the PSID, which includes a long panel of data available at the household level. Both datasets are frequently used as sources of information by researchers.[2]

There are two main concerns with using data from an online posting site such as Glassdoor. First, online data may not be representative of the population. Our first—and not surprising—finding is that user entries in Glassdoor do not accurately represent the national employment distribution across industries. For example, Glassdoor is overrepresented in industries such as information technology, finance, and telecommunications. In contrast, it is underrepresented in industries such as construction, restaurants and food services, and especially health care. We find that the Glassdoor data, however, are well-represented across metropolitan statistical areas (MSAs), with a correlation of the share of user entries by MSA in Glassdoor and QCEW of 0.94. However, we consider the industry misrepresentation more important, as labor income is likely to depend more on industry rather than regional characteristics. Nevertheless, estimating a population mean on the basis of a sample that fails to represent the target population can be addressed by weighting the entries.[3]

The second, and more important, issue is potential measurement error. Online respondents may intentionally or unintentionally misreport their salary. We test for the presence of measurement error by

and Kahn (2017) use online job postings to show that skill requirements differentially increased in MSAs that were hit hard by the Great Recession.

[2] Chamberlain and Nunez (2016) develop a statistical model based on Glassdoor data and compare median weekly earnings of full-time wage and salary workers to the Current Population Survey, which covers about 60,000 households. The authors report a relatively small deviation between the two, around 5 percent.

[3] For more on this topic, the reader can refer to the paper by Solon et al. (2013).

comparing the mean and the standard deviation of the distribution of salaries in Glassdoor, conditional on a group characteristic, with the respective moments in QCEW and PSID. We focus on two characteristics, the worker's industry and region.

When we compare average salaries between Glassdoor and QCEW, we find a reasonably high correlation both across industries and regions. For example, in the real estate sector, the average salaries in QCEW and Glassdoor are $52,509 and $51,805, respectively; in entertainment, they are $36,118 and $39,395, respectively; and in manufacturing, they are $64,999 and $63,964, respectively. The most important discrepancies between Glassdoor and QCEW are observed in industries where workers receive high salaries. These include finance, media, and biotech and pharmaceutical. Overall, the crossindustry correlation between QCEW and Glassdoor is 0.87. When we compute the correlation of average annual salaries across MSAs, we find a correlation of 0.83.

PSID gives an even higher correlation in average wages when it is compared to Glassdoor (equal to 0.9). When we compare the within industry dispersion between salaries in Glassdoor and PSID, we find a correlation of 0.77, which is still high but considerably lower than the correlation in average salaries.

We conclude that the wage distribution (conditional on industry or region) in Glassdoor represents the respective distributions in other datasets, such as QCEW and PSID fairly well. In contrast, the industry employment shares in Glassdoor do not represent the employment distribution across industries in the US well.

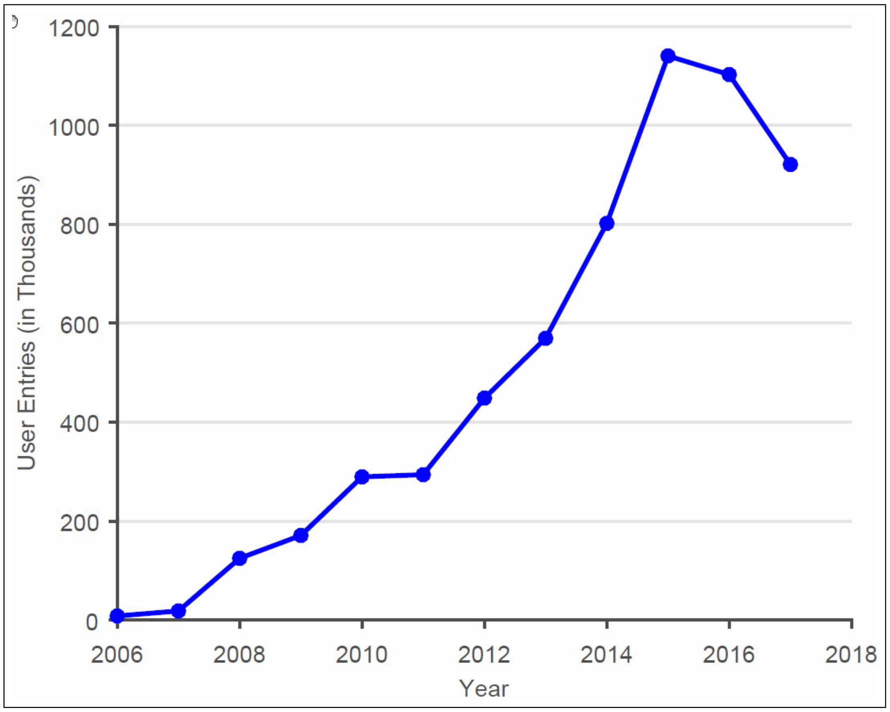
1. DESCRIPTION OF DATASETS

Data from Glassdoor

Glassdoor is one of the leading job sites people use to find jobs and companies use to recruit prospective employees. Users are required to register in order to access user-generated content, which includes company ratings and reviews, typical job interview questions, and CEO approval rates, among other things. Glassdoor requires all registered users to provide some information about their current job, such as their occupation title (job position), the name of the company, and their salary. Users describe their sources of income as well: they distinguish between annual salary (or hourly wage rate) and tips, stock options, or bonuses. They also post information about their experience and the geographical location of the job, described by the city name.

We examine around 6 million salary entries in the Glassdoor database. Figure 1 plots the total number of salary entries by year (flow). As the website became more popular, the number of online users has

Figure 1 User Entries in Glassdoor



Notes: Number of new user entries in website between 2006-17. Entries are reported in thousands.

been expanding. Between 2010 and 2017, the user entries went from around 290,000 to around 1,100,000. We also have 218,462 observations for the first five months of 2018, which we include in the analysis.

Each user has a unique ID number. Since a user may have reported multiple salaries for the same or different jobs, there may be multiple salary entries per user. However, very few users do so. Specifically, 96.4 percent of the users reported one salary, 3.1 percent two salaries, and 0.4 percent three salaries. For each entry we have the exact date of the record, the user's job title, salary, company name, industry, and city name.

Job titles can range from graphic designer, bartender, and nanny to sales associate, project manager, and engineer. There are 190,336 distinct job titles in Glassdoor. Table 1 shows the twenty most common job titles found in the data and their respective shares as a fraction of

Table 1 Most Common Job Titles and Companies in Glassdoor

Job titles		Companies	
Job title	Freq.	Company	Freq.
Manager	4.86%	Amazon.com	1.29%
Software engineer	2.79%	Deloitte	0.78%
Sales associate	2.29%	AT&T	0.76%
Project manager	1.73%	Target	0.68%
Store manager	1.68%	Walmart	0.64%
Cashier	1.46%	Ernst and Young	0.52%
Customer serv. representative	1.42%	Wells Fargo	0.46%
Account manager	1.27%	Microsoft	0.45%
Consultant	1.19%	Bank of America	0.45%
Intern	1.09%	IBM	0.41%
Account executive	1.08%	Best Buy	0.40%
Engineer	1.01%	Home Depot	0.37%
Operations manager	0.96%	Starbucks	0.37%
Administrative assistant	0.94%	Lowe's	0.35%
Registered nurse	0.88%	J.P. Morgan	0.34%
Associate	0.88%	Apple	0.32%
Analyst	0.87%	Walgreens	0.32%
Marketing manager	0.84%	PwC	0.31%
Business analyst	0.82%	Macy's	0.31%
Sales representative	0.82%	US Army	0.31%

Notes: The twenty most common job positions and companies as they appear in Glassdoor. Frequency is the number of user entries in Glassdoor in a specific job position/company as a fraction of total user entries across all years.

the total number of observations. The job with the highest representation is manager followed by software engineer. This makes sense as workers in these job positions are more likely to feel comfortable using job-posting websites. In addition, there are many jobs affiliated with the retail sector, such as retail sales associate, store manager, cashier, and sales representative. Other frequent jobs include analyst, different types of accountants, and project managers.

We perform a similar analysis with respect to companies. There are 222,982 distinct companies in Glassdoor. Companies with the highest representation are most often in the retail sector: Target, Walmart, Amazon.com, Best Buy, Macy's, and others. The others are in software and electronic product development such as Microsoft, IBM, and Apple, or in the financial sector such as Wells Fargo, Bank of America, JPMorgan, and PricewaterhouseCoopers. Although not reported in the table, we also find the cities with the highest representation. There are 17,437 distinct cities. The most-represented city is New York (6.5

percent), followed by Chicago (3.2 percent), San Francisco (2.3 percent), Houston (2.1 percent), Atlanta (2.1 percent), Los Angeles (2.0 percent), Seattle (1.9 percent), Washington (1.8 percent), Boston (1.8 percent), Dallas (1.7 percent), and Austin (1.3 percent).

Users can report their labor income payments at an annual or hourly frequency. When users are asked about their salary, they are asked about their base pay as well as cash bonuses, stock bonuses, profit shares, commissions, and tips. Around 64 percent of observations have annual salary entries, while 34 percent have hourly rates. Around 2 percent report their labor earnings in a monthly frequency. About 23 percent of our sample has information on cash bonuses, 3 percent on stock bonuses, 3 percent on profit sharing, 6 percent on commissions, and 1 percent on tips.

Users also report years of experience. This variable (available for 99.9 percent of the entries) takes values between zero and sixty. In the database, 16 percent report zero years of experience, 9 percent report five years, 6 percent report ten years, 3 percent report fifteen years, and 3 percent report twenty years. Glassdoor also provides some demographic characteristics about the users. Available information includes the users' highest education level, gender, and race. From all Glassdoor responses, 34 percent have nonmissing entries for highest attained education level, 66 percent for gender, and 5 percent for race.

Quarterly Census of Employment and Wages

The Department of Labor's Bureau of Labor Statistics (BLS) runs and maintains three datasets that examine and track the behavior of labor markets at the state and local levels: the Current Employment Statistics, the Local Area Unemployment Statistics, and the QCEW. From all these sources, the most reliable and straightforward counterpart to the Glassdoor data are the data released by the QCEW program.

QCEW provides thorough information on the number of establishments, monthly employment, and quarterly wages in the US. The data are collected from state and federal unemployment insurance records. Since approximately 9 million businesses report this information to state and federal unemployment insurance agencies, the data cover 98 percent of all salary and civilian employment in the country. The information is available at different levels of geographical detail (MSA, county, state, and national levels) and industry detail (down to six-digit NAICS codes). We use data from the period 2010-16, which roughly correspond to the years of data available on Glassdoor.

QCEW data have some limitations, which we briefly describe here. First, for confidentiality reasons, nearly 60 percent of the most detailed level data are suppressed. Second, QCEW does not account for some categories of employment such as self-employed, nonprofit, and military workers, among others. And third, the way the data are collected by states may not be fully consistent, since standards for unemployment insurance coverage vary across states.

Panel Study of Income Dynamics

The PSID includes a long panel of households. The survey was conducted annually until 1997 and biannually from 1999-2015. We use, in the present analysis, data from the period 2003-11. For each year, we use the information associated with the head of the household, including total amount of hours supplied, annual labor income, and industry. The latter is available at the three-digit level. For hours we use the variable "Head Annual Hours of Work." This variable represents the total annual work hours for all jobs including overtime. For labor income, we use the variable "Head Wage," which includes wages and salaries. We deflate salaries using the CPI deflator.

Summary of Available Information: QCEW vs. PSID vs. Glassdoor

Table 2 compares the information available in Glassdoor to the corresponding information in QCEW and PSID. Glassdoor data offer many advantages relative to the other two datasets. In Glassdoor, labor income is available at the worker level. Glassdoor also offers information on the job title, employer, and industry. PSID offers information on the three-digit occupation/industry of the worker, which is broader than the exact job title. Moreover, both Glassdoor and QCEW include detailed geographical information while PSID does not. At the same time, data from Glassdoor have a few shortcomings. As mentioned earlier, Glassdoor is a repeated cross-section of workers and not a panel. Moreover, there is no information on working hours on Glassdoor, although there is some information on part-time versus full-time work.

2. MEASUREMENT ISSUES

We compare Glassdoor with a) QCEW in terms of employment shares and average wages by industry and geographic area and b) PSID in terms of average wages and dispersion in wages by industry. Industries in Glassdoor are not directly comparable to industries in QCEW and PSID. Glassdoor uses an industry descriptor that roughly corresponds to four-digit industry codes. Some examples of industries or

Table 2 QCEW vs. PSID vs. Glassdoor

	QCEW	PSID	Glassdoor
Worker ID	X	✓	✓
Job Title	X	X	✓
Occupation	X	✓	X
Employer	X	X	✓
Industry	✓	✓	✓
Location	✓	X	✓
Panel Data	✓	✓	X
Information on Labor Income	✓	✓	✓
Information on Hours	X	✓	X
Survey	✓	✓	✓

Notes: Comparison between datasets: QCEW, PSID, and Glassdoor.

industry bundles are accounting and legal, consumer services, finance, government, health care, real estate, retail, information technology, manufacturing, and others. Glassdoor does offer a narrower definition of industries (such as car rentals, bars and restaurants, oil and gas exploration, airlines, and other groups of economic activity), but this information is not available for all entries, so we use the broader industry definition.

Our first task is, therefore, to match as closely as possible the industry sectors reported in Glassdoor and QCEW. For some industry categories, there is a direct mapping between the two databases. Some examples are manufacturing; arts, entertainment, and recreation; real estate; business services; telecommunications; and retail. For other sectors, we construct a mapping using a bundle of industries from QCEW. As an example, for biotech and pharmaceuticals, we use industry codes 3254 and 5417, which correspond to pharmaceutical and medicine manufacturing and scientific research and development, respectively. Matching geographical areas between Glassdoor and QCEW is a more straightforward exercise. In particular, to make geographic areas consistent across databases, we merge cities to the appropriate MSA.

Matching industries between Glassdoor and PSID also involves combining different industry codes in PSID and matching them to a corresponding sector in Glassdoor. For example, for accounting/legal we combine industry codes 727 and 728 in the PSID to get the closest possible match, while for government, we combine fifteen different industry codes, ranging from 937 to 987.

A second issue is to transform hourly rates to annual salaries because in Glassdoor, 34 percent of user entries report compensation in

Table 3 Employment Shares By Industry

Sector	QCEW (%)	Glassdoor (%)
Accounting/Legal	1.00	2.99
Aerospace/Defense	0.39	2.21
Agriculture/Forestry	0.30	0.24
Arts/Entertainment/Recreation	1.80	1.41
Biotech/Pharmaceuticals	0.38	1.94
Business services	18.90	11.02
Construction/Repair/Maintenance	5.18	1.58
Consumer Services	3.98	1.11
Education	2.40	6.52
Finance	1.13	7.55
Government	4.96	2.72
Health Care	15.22	7.33
Information Technology	2.37	13.35
Insurance	1.69	2.60
Manufacturing	9.93	8.37
Media	0.11	2.48
Mining/Metals	0.13	0.11
Oil/Gas/Energy/Utilities	0.26	1.85
Real Estate	1.34	1.20
Restaurants/Bars/Food services	9.09	3.89
Retail	14.79	13.02
Telecommunications	0.66	2.71
Transportation/Logistics	3.28	2.04
Travel/Tourism	0.72	1.77
All sectors	100.00	100.00

Notes: Employment shares by industry in QCEW and Glassdoor.

hourly rates. We transform hourly rates into annual salaries by multiplying the hourly rate by 2,000 hours, which is about the average hours worked for a full-time worker per year. We then calculate average salary in industry/area i as follows:

$$\text{Average salary}_i = \begin{cases} \text{fraction salaried workers}_i \times \text{average salary}_i \\ + \\ \text{fraction hourly paid workers}_i \times \text{average hourly rate}_i \times 2000. \end{cases}$$

3. RESULTS

In this section, we compare employment shares and average wages across industries and areas between Glassdoor and QCEW. We also compare average and standard deviation in wages across industries between Glassdoor and PSID. For Glassdoor, we use the cumulative data between 2010-17; for QCEW, we use the averages for the period

Table 4 Employment Shares for Selected MSAs

MSA	QCEW (%)	Glassdoor (%)
Atlanta	2.30	3.36
Boston	2.42	3.82
Chicago	3.45	5.52
Detroit	1.76	1.57
Houston	2.21	2.88
Los Angeles	3.26	5.57
Miami	2.28	1.71
New York	7.24	8.92
Philadelphia	2.18	0.28
Seattle	1.92	3.51
10 Large MSAs	29.02	37.14

Notes: Employment shares by selected geographical area (MSAs) in QCEW and Glassdoor.

2010-16; and for the PSID, we use averages for the period 2003-11. It is possible that some of the differences between Glassdoor and PSID arise due to the different time periods analyzed.

Employment Shares: Glassdoor vs. QCEW

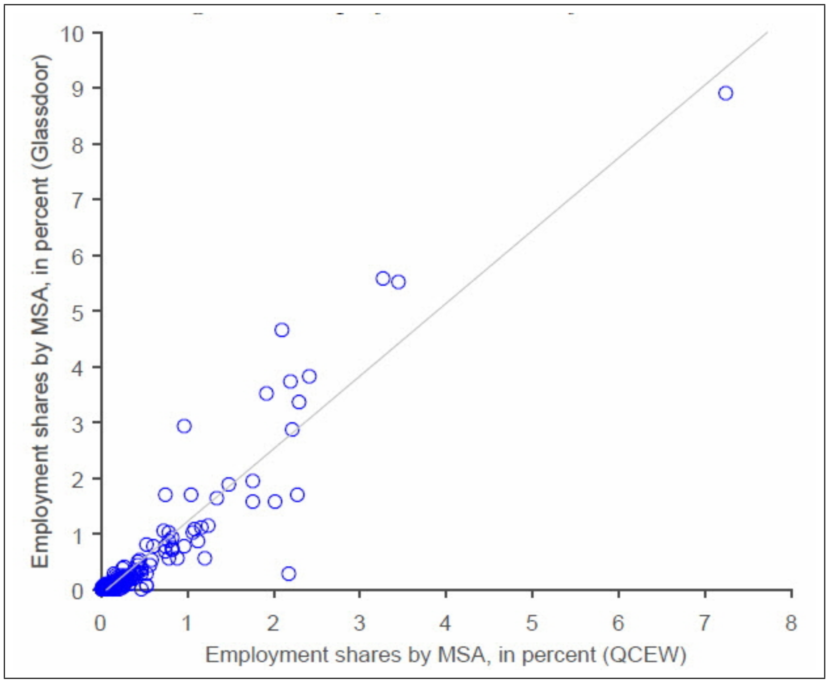
We compare employment shares in a given industry or region in Glassdoor with the respective shares in QCEW. Employment share in Glassdoor is the share of entries in a given industry or region relative to the total number of respondents. Employment share in QCEW is the total number of employed workers in an industry or region as a fraction of total employment.

Table 3 shows employment shares by industry for all years. The observations from Glassdoor are significantly underrepresented in a number of industries including business services, construction, restaurants, food services, and, more importantly, health care. In contrast, Glassdoor is overrepresented in information technology and finance, among others. The correlation between the variables from the two databases is 0.65.

Table 4 describes employment shares obtained from the two databases for ten large US MSAs. From the table, it is clear that large MSAs tend to be overrepresented in Glassdoor. Specifically, employment shares for the ten large MSAs reported in the table is about 37 percent in Glassdoor, and 29 percent in QCEW.

Figure 2 compares employment shares by MSA between QCEW and Glassdoor for all MSAs. We also include a linear fit. The correlation

Figure 2 Employment Shares by MSA



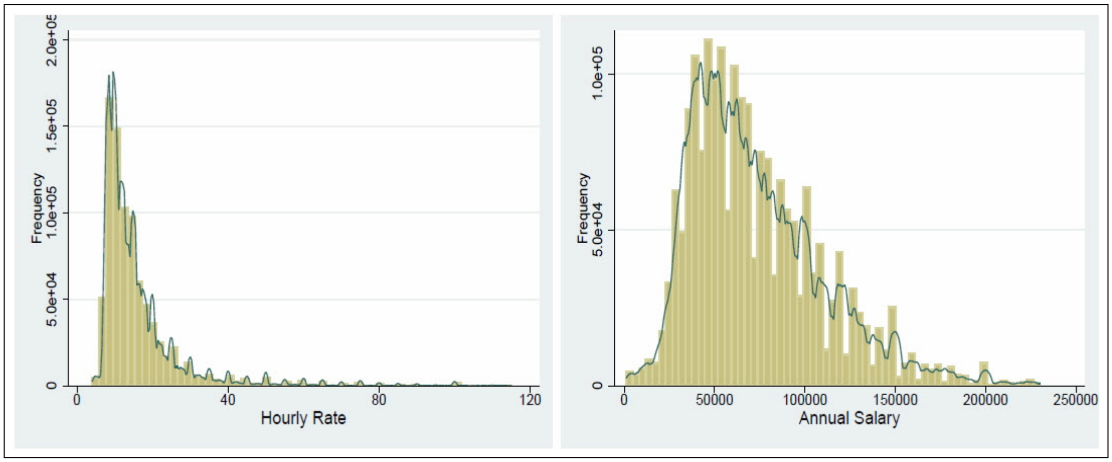
Notes: Employment shares for all geographical area (MSAs) in QCEW and Glassdoor.

is very high, equal to 0.94, which suggests that Glassdoor data are substantially more representative at the MSA level than at the industry level. MSAs with low employment shares (less than 2 percent) seem to be equally represented in both databases. The largest discrepancies are observed for MSAs with relatively large employment shares. As stated earlier, Glassdoor tends to attract respondents disproportionately from those large MSAs.

Average Salaries: Glassdoor vs. QCEW

In this section, we compare average salaries betwen Glassdoor and QCEW. We start by analyzing some salary statistics from Glassdoor. In Figure 3, we plot the distribution of reported salaries and hourly rates, respectively, as they appear in Glassdoor data for all years. The panel on the left shows the distribution of hourly rates. We have

Figure 3 Distribution of Hourly Rates and Salaries in Glassdoor



Notes: Left panel: Distribution of hourly rates in Glassdoor. Right panel: Distribution of annual salaries in Glassdoor.

dropped observations reporting less than $4, which roughly corresponds to half the minimum wage, and also trimmed the top 1 percent of the distribution. The panel on the right shows the distribution of annual salaries. For salaried workers, we dropped observations with less than $1,000 annually and again trimmed the top 1 percent of the distribution.

As mentioned, around 34 percent of user entries report jobs paid in hourly rates. The median hourly rate is $13. The bottom 10 percent in the distribution receives $8.41, while the top 10 percent receives $25. Salaried workers account for approximately 64 percent of user entries in Glassdoor.[4] The median annual salary is $65,000. The bottom 10 percent in the distribution receives $35,000, while the top 10 percent receives $125,000.

So how do the average salaries reported in Glassdoor compare to those in QCEW? Table 5 shows average salaries by industry.

The average wages line up reasonably well for transportation ($48,106 in QCEW vs. $46,966 in Glassdoor), construction ($54,826 vs. $57,534),

[4] As mentioned before, the rest of the workers, around 2 percent, report their labor earnings in a monthly frequency. For simplicity, we will abstract from this group in our analysis.

Table 5 Average Annual Salaries by Industry

Industry	QCEW ($)	Glassdoor ($)
Accounting/Legal	79,087	69,065
Aerospace/Defense	94,501	74,965
Agriculture/Forestry	27,458	53,896
Arts/Entertainment/Recreation	36,118	39,395
Biotech/Pharmaceuticals	116,956	76,298
Business Services	67,175	58,775
Construction/Repair/Maintenance	54,826	57,534
Consumer Services	32,905	40,171
Education	47,096	43,732
Finance	113,685	64,126
Government	52,966	61,991
Health Care	47,061	53,940
Information Technology	89,989	81,908
Insurance	76,132	59,937
Manufacturing	64,999	63,964
Media	88,090	62,987
Mining/Metals	86,408	66,943
Oil/Gas/Energy/Utilities	122,102	72,498
Real Estate	52,509	51,805
Restaurants/Bars/Food Services	17,309	28,341
Retail	28,770	36,906
Telecommunications	78,223	62,448
Transportation/Logistics	48,106	46,966
Travel/Tourism	32,776	42,081

Notes: Average annual salaries in QCEW and Glassdoor by industry.

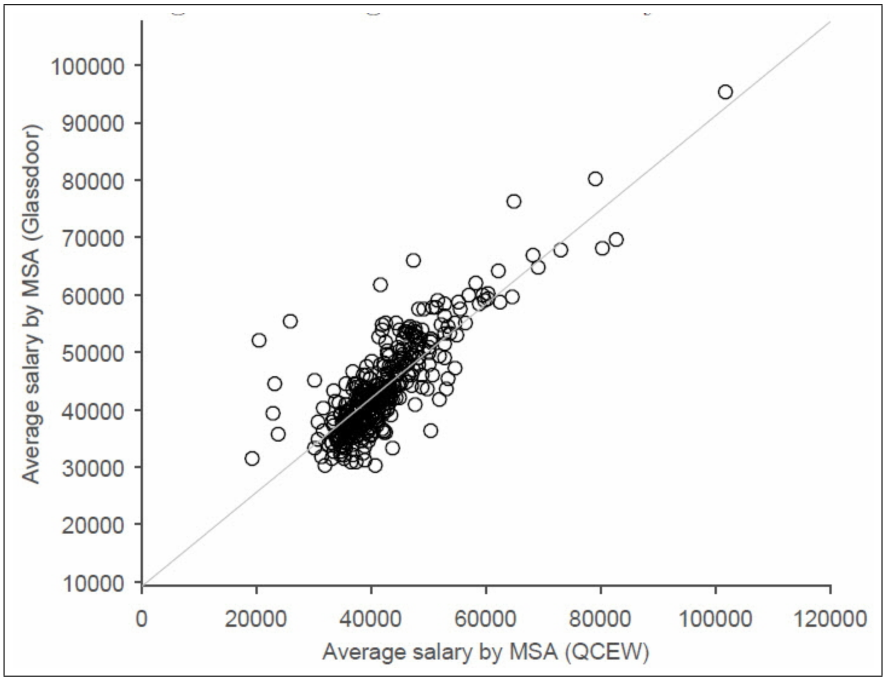
education ($47,096 vs. $43,732), arts and entertainment ($36,118 vs. $39,395), real estate ($52,509 vs. $51,805), and manufacturing ($64,999 vs. $63,964). Overall, the correlation between QCEW and Glassdoor is 0.87.

In Figure 4, we perform a similar comparison across MSAs. In particular, we compare the average salary in a location, as it appears in QCEW, with the average salary in the area from Glassdoor. The correlation between the two is 0.83.

Average Salaries: Glassdoor vs. PSID

In this section, we compare data between Glassdoor and PSID. Both datasets are available at the worker level. We focus on the average salary and the dispersion of the wage distribution (standard deviation). As mentioned, we perform only crossindustry comparisons as detailed geographical information are not available in the PSID. The median industry in the PSID includes 659 observations. The largest number of

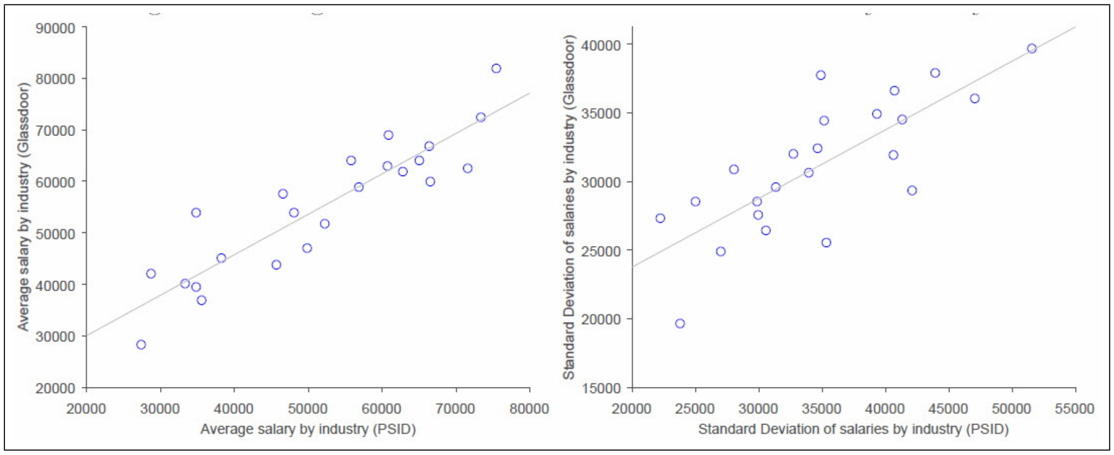
Figure 4 Average Annual Salaries by MSA



Notes: Average annual salaries in QCEW and Glassdoor by MSA.

observations is in manufacturing (4,665), and the smallest is in mining (136). The left panel in Figure 5 plots average salary by industry in PSID and Glassdoor, respectively. The right panel in Figure 5 plots the standard deviation of annual salaries across industries in PSID and Glassdoor. Table 6 gives the numbers used to construct the right panel in Figure 5. The correlation in average salaries between PSID and Glassdoor is even higher than the one with QCEW, equal to 0.9. However, the within-industry dispersion in salaries in Glassdoor is not as close to the PSID as the correlation in average salary. The correlation is 0.77.

Figure 5 Average and Standard Deviation of Annual Salaries by Industry



Notes: Left panel plots average annual salaries by industry for PSID and Glassdoor. Right panel plots standard deviation of annual salaries by industry in PSID and Glassdoor.

4. CONCLUSION AND SUMMARY OF FINDINGS

Glassdoor collects and records millions of observations on salaries by job titles, companies, and cities. The purpose of our paper is to evaluate the extent to which the salary data reported by Glassdoor replicates more traditional datasets, namely QCEW and the PSID. Our findings are summarized in Table 7. The correlation between industry employment shares in Glassdoor and QCEW is relatively low, equal to 0.65. The correlation between MSA employment shares in Glassdoor and in QCEW is higher though, equal to 0.94. Regarding average annual wages, the correlation is fairly high, namely 0.87 across industries and 0.83 across MSAs. Finally, the correlation in average salaries between Glassdoor and PSID is 0.90, and in industry-wide dispersion in salaries it is 0.77.

Table 6 Standard Deviation in Annual Salaries

Sector	Glassdoor ($)	PSID ($)
Accounting/Legal	36,018	47,053
Agriculture/Forestry	30,870	28,019
Arts/Entertainment/Recreation	28,528	29,815
Business services	34,478	41,293
Construction/Repair/Maintenance	29,624	31,317
Consumer Services	28,542	25,007
Education	24,872	26,994
Finance	37,913	43,916
Government	32,010	32,696
Health Care	30,659	33,966
Information Technology	39,702	51,549
Insurance	31,917	40,614
Manufacturing	34,425	35,171
Media	36,654	40,699
Mining/Metals	32,377	34,625
Nonprofit	25,517	35,291
Oil/Gas/Energy/Utilities	34,889	39,344
Real Estate	29,345	42,060
Restaurants/Bars/Food services	19,611	23,740
Retail	27,554	29,957
Telecommunications	37,715	34,937
Transportation/Logistics	26,401	30,528
Travel/Tourism	27,363	22,244

Notes: Standard deviation in annual salaries in PSID and Glassdoor by industry.

Table 7 Summary of Findings

Correlations	Industries	Areas	Data
Employment share	0.65	0.94	Glassdoor/QCEW
Avg. annual salaries	0.87	0.83	Glassdoor/QCEW
Avg. annual salaries	0.90	N/A	Glassdoor/PSID
St. dev. annual salaries	0.77	N/A	Glassdoor/PSID

REFERENCES

Chamberlain, Andrew, and Mario Nunez. 2016. "Glassdoor Local Pay Reports Methodology." Glassdoor Research Studies (December).

Hershbein, Brad, and Lisa B. Kahn. 2017. "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings." Working Paper 22762. Cambridge, Mass.: National Bureau of Economic Research. (September).

Kudlyak, Marianna, Damba Lkhagvasuren, and Roman Sysuyev. 2013. "Systematic Job Search: New Evidence from Individual Job Application Data." Federal Reserve Bank of Richmond Working Paper 12-03R (September).

Solon, Gary, Steven Haider, and Jeffrey Wooldridge. 2013. "What Are We Weighting For?" Working Paper 18859. Cambridge, Mass.: National Bureau of Economic Research. (February).