Identification in Symmetric English Auctions with Additively Separable Unobserved Heterogeneity

Daniel Quint
University of Wisconsin

October 2015

Abstract. I consider identification of a symmetric, independent private values model with additively separable unobserved heterogeneity from observation of winning bids in English auctions. If the number of bidders N is observable, the model is identified given exogenous variation in N, and N need only take two values. If N is not observable, the model is identified if observations are available from auctions under two different known probability distributions of N – for example, if N follows a Poisson distribution, with a mean that varies in a known way as a function of an observable "participation shifter" that is independent of valuations.

1 Introduction

It's known (Athey and Haile 2002) that in English auctions, a symmetric independent private values model is identified by bid data, while a general symmetric private values model is not. Athey and Haile (2007, section 6.1) speculate that given exogenous (observed) variation in the number of bidders, an IPV model with additively separable unobserved heterogeneity might be identified; such a model has proved quite tractable in first-price auctions, and is appealing for empirical work.[1]

Aradillas-López, Gandhi and Quint (2013) consider English auctions under a more general correlated private values model, and show that exogenous variation in N leads to identification of bounds on the objects of most interest (expected revenue and bidder surplus as a function of reserve price), but without identifying the entire model. There, identification requires "a lot" of variation in N: point identification occurs only in the limit as N varies unboundedly, and the identified bounds are wide when N takes only a few values.

Here, I show that the model with additively-separable unobserved heterogeneity is indeed identified, and that this requires only the minimal possible variation in N: observation of just two sizes

[1] When bidder valuations are independent conditional on auction-specific information observed by the bidders but not the analyst, different bidders' bids in the same auction can be interpreted as independent noisy signals of that underlying information. This "measurement error" approach introduced by Li and Vuong (1998) has been applied, or extended, by Li, Perrigne and Vuong (2000), Krasnokutskaya (2011), and Hu, McAdams and Shum (2013), and Athey, Levin, and Seira (2011), among others. Much of this work has assumed that the "common" and "idiosyncratic" value components are either additively or multiplicatively separable. This technique does not naturally extend to English auctions, however, since multiple "independent" bids are not available from each auction.

of auctions suffices to fully identify the model. Of course, this means that if N varies more than that, the model is overidentified, and the assumption of exogenous N (or the joint assumptions of exogenous N and additive separability) is therefore testable.

I also consider the case where the number of bidders in each auction cannot be observed. When N is unobservable, I model it as a random variable with a known distribution. I assume the researcher has access to a "participation shifter" – a variable that affects the distribution of N, but not the distribution of valuations. If the researcher has access to transaction price data from auctions with two distinct distributions of N, then the model is again identified. For example, if auctions were held on two different days of the week, with Saturday auctions having a different (known) distribution of N than Wednesday auctions, this would suffice.

I will show the result for observable N first, followed by the result for unobservable N. I prove the results for two different versions of the model: one with distributions with continuous support under a strong smoothness condition, and one with distributions with discrete support (and no additional conditions). For the continuous model, observation of auctions with any two known distributions of N suffices for identification; for the discrete model, I assume that N follows a Poisson distribution, and that the researcher observes auctions with two different (known) Poisson parameters. Proofs omitted from the text are contained in the appendix.

2 General Model

The model is symmetric independent private values with additively separable unobserved heterogeneity. Let N denote the number of bidders in an auction. Bidder i's private value is

$$v_i \quad = \quad \theta + \epsilon_i$$

where θ is a common term (observed by all bidders but not the econometrician) and $\{\epsilon_i\}$ are *i.i.d.* and independent of θ. I assume that θ and $\{\epsilon_i\}$ take nonnegative values; I assume that 0 is the lowest point in the support of both, but this is just a normalization. I will deal separately with the cases where θ and ϵ_i are discrete-valued and continuous-valued.

I assume that any variation in N is exogenous – i.e., that it is independent of θ and $\{\epsilon_i\}$, whose distributions are the same regardless of N.

Finally, I assume that data on past auctions includes the number of bidders N and the winning bid (or transaction price) T, and that the winning bid in each auction is equal to the second-highest valuation.[2] Thus, I assume that the distribution of the second-highest valuation, conditional on N, is identified for certain values of N. Importantly, I assume the past auctions did not have binding reserve prices, so that the distribution of T is observed all the way down to the bottom of the support of $\theta + \epsilon_i$.

[2]Like Aradillas-López, Gandhi and Quint (2013), I ignore losing bids, to avoid worrying about how to interpret them. Unlike in the "incomplete" model of Haile and Tamer (2003), however, I do assume that the transaction price perfectly matches the second-highest valuation, which would not be exactly true in auctions with a discrete bid interval or jump bidding. See Aradillas-López, Gandhi and Quint (2013), p. 493, for a discussion of this assumption.

3 Observable N

3.1 Continuous Valuations

First, as is common in auction theory, I consider the case where θ and ϵ_i are continuous-valued. Let F_θ and F_ϵ denote the distributions of θ and ϵ_i, with f_θ and f_ϵ the corresponding density functions. Suppose the support of f_θ is either $\Re^+$ or $[0, \overline{\theta}]$ for some finite $\overline{\theta}$, and likewise the support of f_ϵ is either $\Re^+$ or $[0, \overline{\epsilon}]$ for some finite $\overline{\epsilon}$.

When N is observable, *any* exogenous variation in N is sufficient to identify the model, provided both f_ϵ and f_θ satisfy a strong smoothness condition:

Theorem 1. *If f_θ and f_ϵ are known to be analytic on their supports, and if N varies exogenously and takes at least two values, then observation of N and T nonparametrically identifies the model.*

That is, as long as N varies exogenously, observation of the distribution of T for two values of N suffices to recover the distributions of θ and ϵ_i.

Let $F_{T|N}$ denote the distribution of transaction prices conditional on N, and $f_{T|N}$ its density function. For any density function f, let $f^{(m)}$ denote its m^{th} derivative, with $f^{(0)} = f$. If f_ϵ and f_θ are analytic, then they are infinitely differentiable and locally equal to their Taylor series. This means that in an open neighborhood around 0, they are uniquely determined by the infinite series of derivatives $\{f_\epsilon^{(k)}(0), f_\theta^{(k)}(0)\}_{k=0,1,2,\ldots}$. The proof of Theorem 1 will therefore follow from the following lemma:

Lemma 1. *Fix n and n', with $n' > n \geq 2$, and let f_θ and f_ϵ be analytic.*

1. *$f_\theta(0)$ and $f_\epsilon(0)$ can be recovered from the derivatives $f_{T|n}^{(n-1)}(0)$ and $f_{T|n'}^{(n'-1)}(0)$.*

2. *For any $k > 0$, the derivatives $f_\theta^{(k)}(0)$ and $f_\epsilon^{(k)}(0)$ can be recovered from the derivatives $f_{T|n}^{(n-1+k)}(0)$ and $f_{T|n'}^{(n'-1+k)}(0)$ and the lower derivatives $\{f_\theta^{(j)}(0), f_\epsilon^{(j)}(0)\}_{j<k}$.*

Given Lemma 1, we can recursively calculate all derivatives of f_ϵ and f_θ at 0, uniquely pinning down f_ϵ and f_θ in a neighborhood of 0. This means if two different sets of primitives $(\widehat{f}_\epsilon, \widehat{f}_\theta)$ and $(\widetilde{f}_\epsilon, \widetilde{f}_\theta)$ could both explain the observed distributions $f_{T|n}$ and $f_{T|n'}$, they would need to be equal on a neighborhood of 0; the Identity Theorem for real-analytic functions would then imply that $(\widehat{f}_\epsilon, \widehat{f}_\theta) = (\widetilde{f}_\epsilon, \widetilde{f}_\theta)$ everywhere, ensuring that the observables could only have been generated by a single set of primitives.

The proof of Lemma 1 is in the appendix, but the following example should give clear intuition for how it works.

Intuition for Lemma 1

Preliminaries. Let $\epsilon^{(2)}$ denote the second-highest of $\{\epsilon_i\}_{i=1,2,\ldots,N}$. Let $F_{y|n}(\cdot)$ denote the distribution of $\epsilon^{(2)}$ given $N = n$, and $f_{y|n}$ its density. Since $T = \theta + \epsilon^{(2)}$,

$$F_{T|n}(t) \quad = \quad \int_0^t F_{y|n}(t-s) f_\theta(s) ds$$

Calculating successive derivatives and evaluating them at $t = 0$,

$$\begin{aligned}
f'_{T|n}(0) &= f_{y|n}(0) f_\theta(0) \\
f''_{T|n}(0) &= f_{y|n}(0) f'_\theta(0) + f'_{y|n}(0) f_\theta(0) \\
f'''_{T|n}(0) &= f_{y|n}(0) f''_\theta(0) + f'_{y|n}(0) f'_\theta(0) + f''_{y|n}(0) f_\theta(0) \\
f''''_{T|n}(0) &= f_{y|n}(0) f'''_\theta(0) + f'_{y|n}(0) f''_\theta(0) + f''_{y|n}(0) f'_\theta(0) + f'''_{y|n}(0) f_\theta(0) \\
f'''''_{T|n}(0) &= f_{y|n}(0) f''''_\theta(0) + f'_{y|n}(0) f'''_\theta(0) + f''_{y|n}(0) f''_\theta(0) + f'''_{y|n}(0) f'_\theta(0) + f''''_{y|n}(0) f_\theta(0)
\end{aligned}$$

Next, note that by properties of order statistics,

$$F_{y|3}(t) \quad = \quad 3F_\epsilon^2(t) - 2F_\epsilon^3(t)$$

Taking successive derivatives and evaluating them at 0,

$$\begin{aligned}
f_{y|3}(0) &= 0 \\
f'_{y|3}(0) &= 6f_\epsilon^2(0) \\
f''_{y|3}(0) &= 18 f_\epsilon(0) f'_\epsilon(0) - 12 f_\epsilon^3(0) \\
f'''_{y|3}(0) &= 24 f_\epsilon(0) f''_\epsilon(0) + 18 (f'_\epsilon(0))^2 - 72 f_\epsilon^2(0) f'_\epsilon(0)
\end{aligned}$$

Likewise,

$$F_{y|4}(t) \quad = \quad 4F_\epsilon^3(t) - 3F_\epsilon^4(t)$$

from which we can calculate

$$\begin{aligned}
f_{y|4}(0) &= 0 \\
f'_{y|4}(0) &= 0 \\
f''_{y|4}(0) &= 24 f_\epsilon^3(0) \\
f'''_{y|4}(0) &= 144 f_\epsilon^2(0) f'_\epsilon(0) - 72 f_\epsilon^4(0) \\
f''''_{y|4}(0) &= 240 f_\epsilon^2(0) f''_\epsilon(0) + 360 f_\epsilon(0) (f'_\epsilon(0))^2 - 720 f_\epsilon^3(0) f'_\epsilon(0)
\end{aligned}$$

Step 1: recover $f_\epsilon(0)$ and $f_\theta(0)$. If we plug the expressions for $f'_{y|3}(0)$ and $f''_{y|4}(0)$ into the expressions for $f''_{T|3}(0)$ and $f'''_{T|4}(0)$, we get

$$f''_{T|3}(0) \quad = \quad 6 f_\epsilon^2(0) f_\theta(0)$$

$$f'''_{T|4}(0) \quad = \quad 24 f_\epsilon^3(0) f_\theta(0)$$

Dividing one by the other, we find $\frac{f'''_{T|4}(0)}{f''_{T|3}(0)} = 4f_\epsilon(0)$, giving us the value of $f_\epsilon(0)$. Once we know that, we can recover $f_\theta(0)$ from $f''_{T|3}(0) = 6f_\epsilon^2(0)f_\theta(0)$.

Step 2: recover $f'_\epsilon(0)$ and $f'_\theta(0)$. Next, we plug the expressions for $f''_{y|3}(0)$ and $f'''_{y|4}(0)$ into $f'''_{T|3}(0)$ and $f''''_{T|4}(0)$, giving

$$
\begin{aligned}
f'''_{T|3}(0) &= 6f_\epsilon^2(0)f'_\theta(0) + 18f_\epsilon(0)f'_\epsilon(0)f_\theta(0) - 12f_\epsilon^3(0)f_\theta(0) \\
f''''_{T|4}(0) &= 24f_\epsilon^3(0)f'_\theta(0) + 144f_\epsilon^2(0)f'_\epsilon(0)f_\theta(0) - 72f_\epsilon^4(0)f_\theta(0)
\end{aligned}
$$

Since $f_\epsilon(0)$ and $f_\theta(0)$ are already known, we can calculate

$$
\begin{aligned}
\frac{1}{6f_\epsilon^2(0)}\left(f'''_{T|3}(0) + 12f_\epsilon^3(0)f_\theta(0)\right) &= f'_\theta(0) + 3\frac{f'_\epsilon(0)}{f_\epsilon(0)}f_\theta(0) \\
\frac{1}{24f_\epsilon^3(0)}\left(f''''_{T|4}(0) + 72f_\epsilon^4(0)f_\theta(0)\right) &= f'_\theta(0) + 6\frac{f'_\epsilon(0)}{f_\epsilon(0)}f_\theta(0)
\end{aligned}
$$

Subtracting these, we can solve for $f'_\epsilon(0)$ in terms of $f'''_{T|3}(0)$, $f''''_{T|4}(0)$, $f_\epsilon(0)$, and $f_\theta(0)$; we can then recover $f'_\theta(0)$ as the only remaining unknown in either equation.

Step 3: recover $f''_\epsilon(0)$ and $f''_\theta(0)$. Now we start with

$$
\begin{aligned}
f''''_{T|3}(0) &= f'_{y|3}(0)f''_\theta(0) + f''_{y|3}(0)f'_\theta(0) + f'''_{y|3}(0)f_\theta(0) \\
f'''''_{T|4}(0) &= f''_{y|4}(0)f''_\theta(0) + f'''_{y|4}(0)f'_\theta(0) + f''''_{y|4}(0)f_\theta(0)
\end{aligned}
$$

Since we already know $f_\epsilon(0)$, $f'_\epsilon(0)$, $f_\theta(0)$, and $f'_\theta(0)$, we can subtract off all the terms we know, divide by $f'_{y|3}(0)$ and $f''_{y|4}(0)$ respectively, and write

$$
\begin{aligned}
\frac{1}{6f_\epsilon^2(0)}\left(f''''_{T|3}(0) - \text{things we already know}\right) &= f''_\theta(0) + \frac{24f_\epsilon(0)f''_\epsilon(0)}{6f_\epsilon^2(0)}f_\theta(0) \\
\frac{1}{24f_\epsilon^3(0)}\left(f''''_{T|4}(0) - \text{other things we already know}\right) &= f''_\theta(0) + \frac{240f_\epsilon^2(0)f''_\epsilon(0)}{24f_\epsilon^3(0)}f_\theta(0)
\end{aligned}
$$

Subtracting these gives $6f''_\epsilon(0)\frac{f_\theta(0)}{f_\epsilon(0)}$; since we know the values of $f_\theta(0)$ and $f_\epsilon(0)$, we can recover $f''_\epsilon(0)$. $f''_\theta(0)$ can then be recovered from either equation.

Step 4: iterate. Iterating in this way, we can recover each $f_\epsilon^{(k)}(0)$, and then $f_\theta^{(k)}(0)$, from the additional pair of moments $f_{T|3}^{(2+k)}(0)$ and $f_{T|4}^{(3+k)}(0)$. The proof simply requires formalizing the fact that for any (n, n'), this will actually work for each k.

Of course, Lemma 1 also reveals just how strong an assumption analytic distributions are, since the model is identified solely off the distribution of transaction prices in a neighborhood of 0. Of course, this bears no relation to how one would actually choose to estimate such a model, it simply establishes the theoretical certainty of identification. It also underscores the importance that the transaction price data not be truncated by binding reserve prices, as the distribution of T must be observed at the bottom of its support. (That said, if θ and ϵ_i both had bounded support, the analogous procedure could alternatively be run from the top of the support of T instead.)

A more palatable assumption might be to assume f_θ and f_ϵ are *piecewise* analytic. For the sake of illustration, suppose that both distributions were known to be analytic on each interval $[z, z+1)$, $z \in \mathbb{Z}^+$. In that case, we could use the successive derivatives of $f_{T|n}$ and $f_{T|n'}$ at 0 to recover the two distributions on $[0, 1)$, as in Lemma 1. Knowing their values (and their derivatives) on $[0, 1)$, we could then use the right-derivatives of $f_{T|n}$ and $f_{T|n'}$ at 1 to recover the right-derivatives of f_θ and f_ϵ at 1, and therefore recover the distributions on $[1, 2)$; and so on.

However, if analytic distributions is an uncomfortably strong assumption, we can instead establish identification in a completely different way: by considering distributions with discrete support.

3.2 Discrete Valuations

Consider the case where both θ and $\{\epsilon_i\}$ are discrete-valued, each with known support that is bounded below. Thus, assume both θ and $\{\epsilon_i\}$ take values in $\{0, 1, 2, 3, \ldots\}$. I assume zero is in the support of both, but do not require full support above that.[3]

The result is the same as before: with discrete valuations and observable N, the model is identified if there is any exogenous variation in N.

Theorem 2. *If N varies exogenously and takes at least two values, then observation of T and N identifies the model.*

Let $t_k = \Pr(\theta = k)$ and $e_k = \Pr(\epsilon_i = k)$ denote the distributions of θ and ϵ_i. Let $\Pr(T = \cdot|N)$ denote the distribution of transaction prices given N. I prove the following lemma in the appendix, from which Theorem 2 will follow immediately.

Lemma 2. *Fix n and n', with $n' > n \geq 2$.*

1. *The parameters t_0 and e_0 can be recovered from the moments $\Pr(T = 0|n)$ and $\Pr(T = 0|n')$.*
2. *For any $k > 0$, the parameters t_k and e_k can be recovered from the moments $\Pr(T = k|n)$ and $\Pr(T = k|n')$ and the parameters $\{t_j, e_j\}_{j<k}$.*

Given Lemma 2, Theorem 2 follows by induction: t_0 and e_0 can be recovered directly from the data; once these are known, they can be used (along with the data) to recover t_1 and e_1; once these are known, t_2 and e_2 can be recovered; and so on.

[3]The assumption of *common* discrete support is therefore not an additional restriction, as we can think of both θ and ϵ_i as having possible support $\mathrm{supp}(\theta) \cup \mathrm{supp}(\epsilon_i)$, but with some zeroes in the distribution of each.

4 Unobservable N

In some empirical settings, the number of bidders in each auction is not observable, so the approach above will not work.

However, when N is not observable, the model is still identified if the distribution of N is known, and a "participation shifter" – a variable that changes the distribution of N but not the distribution of θ and ϵ_i – is available.

For the continuous and analytic case, I prove this in full generality. For the discrete case, I prove it under the additional assumption that N follows a Poisson distribution. This is the limiting distribution it would take if a large number of potential bidders made independent decisions about whether or not to enter – perhaps a reasonable way to think about many settings, such as eBay auctions, where N is not observed. Under this assumption, knowing the distribution is equivalent to knowing the average number of bidders, which we assume varies in a known way given observables. ("Auctions on Saturdays have 7 bidders on average, while auctions on Wednesdays have 5 bidders on average.")

Formally, let X be a discrete-valued randon variable. Let $\mathbf{p}(x) = (p_2(x), p_3(x), p_4(x), \ldots)$ denote the distribution of N, conditional on $X = x$. (I assume $p_0(x) = p_1(x) = 0$, since auctions with zero or one bidder would not generate positive transaction prices, and would thus be expected to be missing from the data.) The following theorems are proved in the appendix:

Theorem 3. *Suppose f_θ and f_ϵ have continuous support in $\Re^+$, and are both analytic on that support. If $\mathbf{p}(X)$ is known and takes at least two different values, then observation of X and T nonparametrically identifies the model.*

Theorem 4. *Suppose θ and ϵ have known support in $\mathbb{Z}^+$. Suppose $\mathbf{p}(x)$ is the Poisson distribution with parameter $\lambda(x)$, truncated at $N = 2$. If $\lambda(\cdot)$ is a known function of X and takes at least two values, then observation of X and T identifies the model.*

Note that, like with N, we need only the minimal amount of variation in $\mathbf{p}$ – auctions with two different known distributions of N – to identify the model from the distributions of transaction prices.

5 Discussion

Theorems 1, 2, 3, and 4 say that if N or $\mathbf{p}$ take just two values, the model is identified. As a result, if N (or $\mathbf{p}$) takes three or more values, the model is overidentified, and therefore testable.

Perhaps more surprisingly, in the discrete case, even with N (or $\mathbf{p}$) taking just two values, if $F_{T|N}$ has bounded support, the model is overidentified as well. For intuition, focus on the case of observable N, and suppose that the distributions of both θ and ϵ_i have support $\{0, 1, 2, \ldots, M\}$, so that $F_{T|N}$ has support $\{0, 1, \ldots, 2M\}$. Lemma 2 says that the distributions of θ and ϵ_i are identified from the moments $\{\Pr(T = k|N)\}$ for $k = 0, 1, 2, \ldots, M$, that is, the distribution of T only up to

M. So the additional moments $\{\Pr(T = k|N)\}$ for $k = M+1, M+2, \ldots, 2M$ are all additional restrictions that could potentially falsify the model.

Of course, failure of a data set to satisfy these restrictions could be due either to N being correlated with θ or ϵ_i, or to the unobserved heterogeneity not being additively separable. It's not obvious how one would distinguish between these two.

References

1. Aradillas-López, Andrés, Amit Gandhi, and Daniel Quint (2013), "Identification and Inference in Ascending Auctions with Correlated Private Values," *Econometrica* 81(2).

2. Athey, Susan, and Philip Haile (2002), "Identification of Standard Auction Models," *Econometrica* 70(6).

3. Athey, Susan, and Philip Haile (2007), "Nonparametric Approaches to Auctions," in *Handbook of Econometrics* Vol. 6A, J. Heckman and E. Leamer, Eds, Oxford: North-Holland.

4. Athey, Susan, Jonathan Levin, and Enrique Seira (2011), "Comparing Open and Sealed Bid Auctions: Evidence From Timber Auctions," *Quarterly Journal of Economics* 126(1).

5. Haile, Philip, and Elie Tamer (2003), "Inference With an Incomplete Model of English Auctions," *Journal of Political Economy* 111(1).

6. Hu, Yingyao, David McAdams, and Matthew Shum (2013), "Identification of First-Price Auctions with Non-Separable Unobserved Heterogeneity," *Journal of Econometrics* 174.

7. Krasnokutskaya, Elena (2011), "Identification and Estimation in Procurement Auctions under Unobserved Auction Heterogeneity," *Review of Economic Studies* 28.

8. Li, Tong, Isabelle Perrigne, and Quang Vuong (2000) "Conditionally Independent Private Information in OCS Wildcat Auctions," *Journal of Econometrics* 98.

9. Li, Tong, and Quang Vuong (1998), "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis* 65.

Appendix – Omitted Proofs

A.1 Proof of Lemma 1 (continuous valuations, observed N)

Preliminaries (part 1)

As noted in the text, $T = \theta + \epsilon^{(2)}$ implies

$$F_{T|n}(t) \quad = \quad \int_0^t F_{y|n}(t-s) f_\theta(s) ds$$

and therefore

$$f_{T|n}(t) \quad = \quad F_{y|n}(0) f_\theta(t) + \int_0^t f_{y|n}(t-s) f_\theta(s) ds \quad = \quad \int_0^t f_{y|n}(t-s) f_\theta(s) ds$$

It is easily shown by induction that

$$f^{(k)}_{T|n}(t) \quad = \quad \sum_{i=0}^{k-1} f^{(i)}_{y|n}(0) f^{(k-1-i)}_\theta(t) + \int_0^t f^{(k)}_{y|n}(t-s) f_\theta(s) ds$$

and therefore

$$f^{(k)}_{T|n}(0) \quad = \quad \sum_{i=0}^{k-1} f^{(i)}_{y|n}(0) f^{(k-1-i)}_\theta(0)$$

Preliminaries (part 2)

Next, I establish several facts about the derivatives of $f_{y|n}$ at zero. Specifically:

1. For $m < n-2$, $f^{(m)}_{y|n}(0) = 0$.
2. $f^{(n-2)}_{y|n}(0) = n! \cdot (f_\epsilon(0))^{n-1}$.
3. For $m > n-2$, $f^{(m)}_{y|n}(0)$ contains no derivatives of f_ϵ higher than $f^{(m-n+2)}_\epsilon(0)$, and the only term containing $f^{(m-n+2)}_\epsilon(0)$ is
$$n! \sum_{i=1}^{n-1} i \binom{m-i}{m-n+1} (f_\epsilon(0))^{n-2} f^{(m-n+2)}_\epsilon(0)$$
4. Fix $k > 0$ and define $A_n = \sum_{i=1}^{n-1} i \binom{n-2+k-i}{k-1}$; A_n is strictly increasing in n.

To prove all this, we begin with the fact that, since $F_{y|n}(t) = nF^{n-1}_\epsilon(t) - (n-1)F^n_\epsilon(t)$,

$$f_{y|n}(t) \quad = \quad n(n-1)F^{n-2}_\epsilon(t) f_\epsilon(t) - n(n-1)F^{n-1}_\epsilon(t) f_\epsilon(t)$$

If we differentiate this, we get

$$\begin{aligned} f'_{y|n}(t) \quad = \quad & n(n-1)(n-2)F^{n-3}_\epsilon(t) f^2_\epsilon(t) + n(n-1)F^{n-2}_\epsilon(t) f'_\epsilon(t) \\ & -n(n-1)^2 F^{n-2}_\epsilon(t) f^2_\epsilon(t) - n(n-1)F^{n-1}_\epsilon(t) f'_\epsilon(t) \end{aligned}$$

Specifically, we get two terms from differentiating $n(n-1)F_\epsilon^{n-2}(t)f_\epsilon(t)$ – one from taking the derivative of the F_ϵ^{n-2} "part," and one from taking the derivative of the $f_\epsilon(t)$ "part" – and likewise two terms from differentiating $n(n-1)F_\epsilon^{n-1}(t)f_\epsilon(t)$. As we take subsequent derivatives of $f_{y|n}$, we keep getting additional terms, each corresponding to differenating one "piece" of a term from the previous derivative.

Now, the first fact above – $f_{y|n}^{(m)}(0) = 0$ for $m < n-2$ – stems from the fact that until we have taken at least $n-2$ derivatives, every term in $f_{y|n}^{(m)}$ still contains a nonzero power of $F_\epsilon(t)$, which vanishes at 0. Likewise, when we take exactly $n-2$ derivatives, the only term that does not vanish is the one that "used" all $n-2$ derivatives to differentiate the $F_\epsilon^j(t)$ piece of the first term; each time this happens, the term gets multiplied by j (j running from $n-2$ down to 1) and picks up another $f_\epsilon(t)$, so

$$f_{y|n}^{(n-2)}(0) \quad = \quad n \cdot (n-1) \cdot (n-2)! \cdot (f_\epsilon(0))^{n-1} + \text{terms that vanish}$$

Next, suppose we take $m > n-2$ derivatives of $f_{y|n}$. Any term that has a derivative $f_\epsilon^{(m')}$ with $m' > m-n+2$ must have "used" more than $m-(n-2)$ derivatives differentiating $f_\epsilon(t)$ and its subsequent derivatives; this would have left strictly fewer than $n-2$ derivatives to differentiate either F_ϵ^{n-2} or F_ϵ^{n-1}, leaving a positive power of F_ϵ that would therefore vanish at 0. Finally, the only way to have a nonvanishing term containing $f_\epsilon^{(m-n+2)}$ would be to start with the first term of $f_{y|n}$, $n(n-1)F_\epsilon^{n-2}(t)f_\epsilon(t)$, and use exactly $n-2$ derivatives differenating the F_ϵ^{n-2} term and the remaining $m-(n-2)$ derivatives differentiating $f_\epsilon(t)$ and its subsequent derivatives. Each of the $n-2$ derivatives we take of F_ϵ^{n-2} generates an additional f_ϵ term, and we only differentiate one of these, so we're left with $f_\epsilon^{n-2}(0)f_\epsilon^{(m-(n-2))}(0)$. The coefficient on this term is the sum of the coefficients of all the different "ways" we can generate these terms – basically, all the different orders in which we can take $n-2$ derivatives of F_ϵ^{n-2} and $m-(n-2)$ derivatives of f_ϵ.

Now, regardless of the order in which we take the derivatives, at some point, we need to differentiate F_ϵ^{n-2}, generating an $(n-2)$ coefficient; then at some point we differentiate F_ϵ^{n-3}, generating an $(n-3)$; and so on. Combined with the $n(n-1)$ we started with, this gives us a coefficient of $n!$ attached to every nonvanishing term. In addition, at some point, we differentiated $f_\epsilon^i(t)$, which would have generated an i coefficient. The rest of our derivatives were applied to the $f_\epsilon^{(j)}(t)$ term, which never gave any additional multiplicative coefficients.

Now, if we take the derivative of f_ϵ first – when the coefficient on $f_\epsilon^i(t)$ is $i = 1$ – then the coefficient on our eventual non-vanishing term will be $1 \cdot n!$. How many terms like this are there? Well, we still have $m-1$ derivatives left to take, of which $n-2$ need to apply to F_ϵ^{n-2} and the rest to $f_\epsilon^{(j)}$, so there are $\binom{m-1}{n-2}$ different terms corresponding to the choice of differentiating f_ϵ first.

More generally, suppose we differentiate f_ϵ after we have already differentiated F_ϵ^j $i-1$ times, and therefore when the term we're differentiating is $F_\epsilon^{n-2-(i-1)}(t)f_\epsilon^i(t)$. This again provides a new i coefficient. And in addition, we have $m-1-(i-1) = m-i$ derivatives left to take, of which $n-2-(i-1) = n-1-i$ need to be applied to F_ϵ^j; so there are $\binom{m-i}{n-1-i}$ different terms that correspond to this case.

Finally, if we wait to differentiate f_ϵ^i until after we've already taken $n-2$ derivatives of F_ϵ^{n-2}, then we're differentiating f_ϵ^{n-1}, and we get an $i = n-1$ coefficient; but then all remaining derivatives have to be applied to $f_\epsilon^{(j)}$, and there's only one way to do that.

All told, then, the coefficient on $f_\epsilon^{n-2}(0) f_\epsilon^{(m-n+2)}(0)$ in $f_{y|n}^{(m)}$ will be

$$n! \cdot \sum_{i=1}^{n-1} i \cdot \binom{m-i}{n-1-i} \quad = \quad n! \cdot \sum_{i=1}^{n-1} i \cdot \binom{m-i}{m-n+1}$$

Finally, to show that A_n (which is $\frac{1}{n!}$ times this coefficient) is increasing in n, fix k and calculate

$$\begin{aligned}
A_{n+1} - A_n &= \sum_{i=1}^{(n+1)-1} i \binom{(n+1)-2+k-i}{k-1} - \sum_{i=1}^{n-1} i \binom{n-2+k-i}{k-1} \\
&= \sum_{i=1}^{n} i \binom{n-2+k-(i-1)}{k-1} - \sum_{i=1}^{n-1} i \binom{n-2+k-i}{k-1} \\
&= \sum_{i'=0}^{n-1} (i'+1) \binom{n-2+k-i'}{k-1} - \sum_{i=1}^{n-1} i \binom{n-2+k-i}{k-1} \\
&= \sum_{i=0}^{n-1} \binom{n-2+k-i}{k-1} \quad > \quad 0
\end{aligned}$$

This concludes the preliminaries.

Proof of Lemma 1

With these preliminaries established, I prove Lemma 1 by induction on k. For the base case, facts 1 and 2 above, combined with the expansion of $f_{T|n}^{(n-1)}(0)$, give

$$f_{T|n}^{(n-1)}(0) \quad = \quad \sum_{i=0}^{n-2} f_{y|n}^{(i)}(0) f_\theta^{(n-2-i)}(0) \quad = \quad n! \cdot (f_\epsilon(0))^{n-1} \cdot f_\theta(0)$$

and likewise $f_{T|n'}^{(n'-1)}(0) = n'! \cdot (f_\epsilon(0))^{n'-1} \cdot f_\theta(0)$. For $n' > n$, then, we can recover $f_\epsilon(0)$ as

$$f_\epsilon(0) \quad = \quad \left(\frac{\frac{1}{n'!} f_{T|n'}^{(n'-1)}(0)}{\frac{1}{n!} f_{T|n}^{(n-1)}(0)} \right)^{1/(n'-n)}$$

and from there, recover $f_\theta(0)$ as $f_{T|n}^{(n-1)}(0)/(n!(f_\epsilon(0))^{n-1})$.

For the inductive step, we assume we already know $\{f_\epsilon^{(j)}(0), f_\theta^{(j)}(0)\}_{j<k}$. As noted above,

$$f_{T|n}^{(n-1+k)}(0) \quad = \quad \sum_{i=0}^{n-2+k} f_{y|n}^{(i)}(0) f_\theta^{(n-2+k-i)}(0) \quad = \quad \sum_{i=n-2}^{n-2+k} f_{y|n}^{(i)}(0) f_\theta^{(n-2+k-i)}(0)$$

since the first $n-3$ derivatives of $f_{y|n}$ are 0 at 0. Fact 3 implies that for $i < n-2+k$, $f_{y|n}^{(i)}(0)$ contains derivatives no higher than $f_\epsilon^{(k-1)}$, so the only "unknowns" on the right-hand side are

$f_\theta^{(k)}(0)$ and the $f_\epsilon^{(k)}(0)$ term contained in $f_{y|n}^{(n-2+k)}$. Let

$$B(n,k) \quad = \quad \sum_{i=n-1}^{n-3+k} f_{y|n}^{(i)}(0) f_\theta^{(n-2+k-i)}(0)$$

be all but the first and last terms of the sum, and let $C(n,k)$ denote all the terms of $f_{y|n}^{(n-2+k)}(0)$ other than the one containing $f_\epsilon^{(k)}(0)$, both of which depend only on the derivatives $\{f_\epsilon^{(j)}(0), f_\theta^{(j)}(0)\}_{j<k}$ and are therefore known. We can then calculate the value of

$$f_{T|n}^{(n-1+k)}(0) - B(n,k) - C(n,k) f_\theta(0) \quad = \quad f_{y|n}^{(n-2)}(0) f_\theta^{(k)}(0) + \left(n! \cdot A_n (f_\epsilon(0))^{n-2} f_\epsilon^{(k)}(0) \right) f_\theta(0)$$

where

$$A_n \quad = \quad \sum_{i=1}^{n-1} i \binom{(n-2+k)-i}{(n-2+k)-n+1} \quad = \quad \sum_{i=1}^{n-1} i \binom{n-2+k-i}{k-1}$$

Dividing by $f_{y|n}^{(n-2)}(0) = n!(f_\epsilon(0))^{n-1}$ and then by $f_\theta(0)/f_\epsilon(0)$, we get

$$\frac{f_{T|n}^{(n-1+k)}(0) - B(n,k) - C(n,k) f_\theta(0)}{n!(f_\epsilon(0))^{n-1}} \cdot \frac{f_\epsilon(0)}{f_\theta(0)} \quad = \quad \frac{f_\epsilon(0)}{f_\theta(0)} f_\theta^{(k)}(0) + A_n f_\epsilon^{(k)}(0) \tag{A1}$$

So, given our inductive assumption that $f_{T|n}^{(n-1+k)}$, $f_{T|n'}^{(n'-1+k)}$, and $\{f_\epsilon^{(j)}(0), f_\theta^{(j)}(0)\}_{j<k}$ are known, we can calculate the value of the left-hand side of (A1) for both n and n' and subtract, giving us the value of $(A_{n'} - A_n) f_\epsilon^{(k)}(0)$. Fact 4 above says that A_n is strictly increasing in n; since A_n and $A_{n'}$ are known and $A_{n'} - A_n \neq 0$, knowing the value of $(A_{n'} - A_n) f_\epsilon^{(k)}(0)$ allows us to recover $f_\epsilon^{(k)}(0)$. Once we have that, (A1) lets us calculate $f_\theta^{(k)}(0)$ as well, completing the proof. □

A.2 Proof of Lemma 2 (discrete valuations, observed N)

As in the text, let $\epsilon^{(2)}$ denote the second-highest of the ϵ_i terms in a particular auction; and let $\Pr(\epsilon^{(2)} = \cdot | n)$ denote its distribution conditional on $N = n$.

Part 1: recovering t_0 and e_0. For the first part, since θ and ϵ_i both have non-negative support,

$$\Pr(T = 0|N) \quad = \quad \Pr(\theta = 0) \Pr(\epsilon^{(2)} = 0|N) \quad = \quad t_0 \left(N e_0^{N-1} - (N-1) e_0^N \right) \tag{A2}$$

Since we observe both $\Pr(T = 0|n)$ and $\Pr(T = 0|n')$, we can calculate

$$\frac{\Pr(T = 0|n')}{\Pr(T = 0|n)} \quad = \quad \frac{t_0 \left(n' e_0^{n'-1} - (n'-1) e_0^{n'} \right)}{t_0 \left(n e_0^{n-1} - (n-1) e_0^n \right)} \quad = \quad \frac{n' e_0^{n'-1} - (n'-1) e_0^{n'}}{n e_0^{n-1} - (n-1) e_0^n} \tag{A3}$$

This allows us to "eliminate" t_0 – that is, to find a statistic we can calculate from the data that depends only on e_0. Given n and n', define a function $\psi : [0,1] \to \Re^+$ by

$$\psi(x) \quad \equiv \quad \frac{n' x^{n'-1} - (n'-1) x^{n'}}{n x^{n-1} - (n-1) x^n}$$

I'll show that for $n' > n$, ψ is strictly increasing in x. Taking the derivative of its natural log,

$$\begin{aligned}
\frac{d}{dx}\ln\psi(x) &= \frac{d}{dx}\ln\left(n'x^{n'-1}-(n'-1)x^{n'}\right)-\frac{d}{dx}\ln\left(nx^{n-1}-(n-1)x^{n}\right) \\
&= \frac{n'(n'-1)x^{n'-2}-(n'-1)n'x^{n'-1}}{n'x^{n'-1}-(n'-1)x^{n'}}-\frac{n(n-1)x^{n-2}-(n-1)nx^{n-1}}{nx^{n-1}-(n-1)x^{n}} \\
&= \frac{n'(n'-1)x^{n'-2}(1-x)}{x^{n'}+n'x^{n'-1}(1-x)}-\frac{n(n-1)x^{n-2}(1-x)}{x^{n}+nx^{n-1}(1-x)} \\
&= \frac{(n'-1)(1-x)}{\frac{1}{n'}x^2+x(1-x)}-\frac{(n-1)(1-x)}{\frac{1}{n}x^2+x(1-x)}
\end{aligned}$$

For $x \in (0,1)$, this is strictly positive, since $n'-1 > n-1$ (so the first term has the larger numerator) and $\frac{1}{n'} < \frac{1}{n}$ (so the first term has the smaller denominator). Thus, $(\ln\psi)' > 0$ on $(0,1)$, so $\ln\psi$ is strictly increasing on $[0,1]$, so ψ is strictly increasing on $[0,1]$. Writing (A3) as $\frac{\Pr(T=0|n')}{\Pr(T=0|n)} = \psi(e_0)$, this means e_0 can be calculated as

$$e_0 = \psi^{-1}\left(\frac{\Pr(T=0|n')}{\Pr(T=0|n)}\right)$$

Once e_0 has been recovered, (A2) implies

$$t_0 = \frac{\Pr(T=0|n)}{ne_0^{n-1}-(n-1)e_0^{n}}$$

and we're done.

Part 2: recovering e_k and t_k. To prove the second part of the lemma ($k > 0$), note first that

$$\Pr(T=k|N) = \sum_{j=0}^{k}\Pr(\theta=k-j)\Pr(\epsilon^{(2)}=j|N) \tag{A4}$$

As noted above, $\Pr(\epsilon^{(2)}=0|N) = Ne_0^{N-1}-(N-1)e_0^N$. For $j > 0$, define $e_{<j} = \Pr(\epsilon_i < j) = \sum_{j'<j} e_{j'}$. Then

$$\begin{aligned}
\Pr(\epsilon^{(2)}=j|N) &= (e_j+e_{<j})^N - e_{<j}^N - Ne_je_{<j}^{N-1} \\
&\quad +N(1-e_j-e_{<j})\left((e_j+e_{<j})^{N-1}-e_{<j}^{N-1}\right)
\end{aligned}$$

(The first line is the probability that the highest and second-highest of the ϵ_i are both equal to j; the second line is the probability that exactly one ϵ_i is above j, and at least one of the remaining ϵ_i is exactly j.) Note that $\Pr(\epsilon^{(2)}=j|N)$ depends only on $\{e_0, e_1, \dots, e_j\}$. Thus, if we rewrite (A4) as

$$\Pr(T=k|N)-\sum_{j=1}^{k-1}t_{k-j}\Pr(\epsilon^{(2)}=j|N) = t_k\Pr(\epsilon^{(2)}=0|N)+t_0\Pr(\epsilon^{(2)}=k|N)$$

then the left-hand side can be expressed in terms of $\Pr(T=k|N)$ and $\{t_j, e_j\}_{j<k}$, which we assume are already known. Recalling also that $\Pr(\epsilon^{(2)}=0|N)$ depends only on e_0, to prove the lemma, we

can assume that we know the values of

$$\frac{\Pr(T=k|n) - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n)}{\Pr(\epsilon^{(2)}=0|n)} \quad = \quad t_k + t_0\frac{\Pr(\epsilon^{(2)}=k|n)}{\Pr(\epsilon^{(2)}=0|n)}$$

and

$$\frac{\Pr(T=k|n') - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n')}{\Pr(\epsilon^{(2)}=0|n')} \quad = \quad t_k + t_0\frac{\Pr(\epsilon^{(2)}=k|n')}{\Pr(\epsilon^{(2)}=0|n')}$$

Since we know both, we know their difference (which t_k drops out of); and since we also know t_0, we therefore know the value of

$$\frac{1}{t_0}\left[\frac{\Pr(T=k|n') - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n')}{\Pr(\epsilon^{(2)}=0|n')} - \frac{\Pr(T=k|n) - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n)}{\Pr(\epsilon^{(2)}=0|n)}\right]$$

$$\begin{aligned}
&= \frac{\Pr(\epsilon^{(2)}=k|n')}{\Pr(\epsilon^{(2)}=0|n')} - \frac{\Pr(\epsilon^{(2)}=k|n)}{\Pr(\epsilon^{(2)}=0|n)} \\
&= \frac{(e_k+e_{<k})^{n'} - e_{<k}^{n'} - n'e_k e_{<k}^{n'-1} + n'(1-e_k-e_{<k})\left((e_k+e_{<k})^{n'-1} - e_{<k}^{n'-1}\right)}{e_0^{n'-1} + n'(1-e_0)e_0^{n'-1}} \\
&\quad - \frac{(e_k+e_{<k})^{n} - e_{<k}^{n} - ne_k e_{<k}^{n-1} + n(1-e_k-e_{<k})\left((e_k+e_{<k})^{n-1} - e_{<k}^{n-1}\right)}{e_0^{n-1} + n(1-e_0)e_0^{n-1}}
\end{aligned}$$

Now, fixing $e_{<k}$, e_0, n, and n', define another function $\phi : [0,1] \to \Re^+$ by

$$\begin{aligned}
\phi(x) &= \frac{(x+e_{<k})^{n'} - e_{<k}^{n'} - n'x e_{<k}^{n'-1} + n'(1-x-e_{<k})\left((x+e_{<k})^{n'-1} - e_{<k}^{n'-1}\right)}{e_0^{n'-1} + n'(1-e_0)e_0^{n'-1}} \\
&\quad - \frac{(x+e_{<k})^{n} - e_{<k}^{n} - nx e_{<k}^{n-1} + n(1-x-e_{<k})\left((x+e_{<k})^{n-1} - e_{<k}^{n-1}\right)}{e_0^{n-1} + n(1-e_0)e_0^{n-1}}
\end{aligned}$$

so that

$$\frac{1}{t_0}\left[\frac{\Pr(T=k|n') - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n')}{\Pr(\epsilon^{(2)}=0|n')} - \frac{\Pr(T=k|n) - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n)}{\Pr(\epsilon^{(2)}=0|n)}\right] \quad = \quad \phi(e_k)$$

Next, I show that ϕ is invertible. Since we'll be inverting $\phi(e_k) =$ (something), and since e_k and $e_{<k}$ are probabilities of mutually exclusive events, the argument of ϕ will be in $[0, 1-e_{<k}]$; so we want to show $\phi'(x) > 0$ for $x \in (0, 1-e_{<k})$. Noting that the denominators of both terms of ϕ

do not depend on e_k, we can differentiate and get

$$\begin{aligned}
&\phi'(x) \\
= &\quad \frac{n'(x+e_{<k})^{n'-1} - n'e_{<k}^{n'-1} - n'\left((x+e_{<k})^{n'-1} - e_{<k}^{n'-1}\right) + n'(1-x-e_{<k})\left((n'-1)(x+e_{<k})^{n'-2}\right)}{e_0^{n'-1} + n'(1-e_0)e_0^{n'-1}} \\
&- \frac{n(x+e_{<k})^{n-1} - ne_{<k}^{n-1} - n\left((x+e_{<k})^{n-1} - e_{<k}^{n-1}\right) + n(1-x-e_{<k})\left((n-1)(x+e_{<k})^{n-2}\right)}{e_0^{n-1} + n(1-e_0)e_0^{n-1}} \\
= &\quad \frac{n'(1-x-e_{<k})\left((n'-1)(x+e_{<k})^{n'-2}\right)}{e_0^{n'-1} + n'(1-e_0)e_0^{n'-1}} - \frac{n(1-x-e_{<k})\left((n-1)(x+e_{<k})^{n-2}\right)}{e_0^{n-1} + n(1-e_0)e_0^{n-1}} \\
= &\quad (1-x-e_{<k})\left[\frac{(n'-1)(\frac{x+e_{<k}}{e_0})^{n'-2}}{\frac{1}{n'}e_0 + (1-e_0)e_0} - \frac{(n-1)(\frac{x+e_{<k}}{e_0})^{n-2}}{\frac{1}{n}e_0 + (1-e_0)e_0}\right]
\end{aligned}$$

As noted above, $1 - x - e_{<k} > 0$ in the relevant range of x. Since $e_{<k} = \sum_{j<k} e_j \geq e_0$, $\frac{x+e_{<k}}{e_0} \geq 1$; combined with $n' > n$, this means the first term in the square brackets has a strictly higher numerator than the second. And since $\frac{1}{n'} < \frac{1}{n}$, the first term also has a strictly lower denominator. So $\phi' > 0$ on $(0, 1-e_{<k})$, and so ϕ is strictly increasing on $[0, 1-e_{<k}]$, and is therefore invertible.

Thus, we can recover e_k as

$$\begin{aligned}
e_k \quad &= \\
\phi^{-1}&\left(\frac{1}{t_0}\left[\frac{\Pr(T=k|n') - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n')}{\Pr(\epsilon^{(2)}=0|n')} - \frac{\Pr(T=k|n) - \sum_{j=1}^{k-1} t_{k-j}\Pr(\epsilon^{(2)}=j|n)}{\Pr(\epsilon^{(2)}=0|n)}\right]\right)
\end{aligned}$$

Once e_k is known, $\Pr(\epsilon^{(2)} = k|N)$ is known, and so t_k can be recovered from (A4) as

$$t_k \quad = \quad \frac{1}{\Pr(\epsilon^{(2)}=0|n)}\left[\Pr(T=k|n) - \sum_{j=1}^{k} t_{k-j}\Pr(\epsilon^{(2)}=j|n)\right]$$

concluding the proof. □

A.3 Proof of Theorem 3 (continuous valuations, unobserved N)

Let x and x' be two values of X, with $\mathbf{p}(x') \neq \mathbf{p}(x)$. Let $F_{T|X}$ and $f_{T|X}$ denote the distribution and density of transaction prices given X. I'll show that knowledge of $f_{T|x}$ and $f_{T|x'}$ allows us to recover $\{f_\epsilon^{(k)}(0), f_\theta^{(k)}(0)\}_{k=0,1,2,\ldots}$, and that the model is therefore identified by the same logic as Theorem 1.

I begin with a simple case, and then show that we can reduce any case to the simple case.

Lemma 3. *Suppose that $p_2(x) > 0 = p_2(x')$ and $p_3(x') > 0$. Then knowing $\mathbf{p}(x)$ and $\mathbf{p}(x')$...*

1. *$f_\theta(0)$ and $f_\epsilon(0)$ can be recovered from the derivatives $f'_{T|x}(0)$ and $f''_{T|x'}(0)$*

2. *For any $k > 0$, the derivatives $f_\theta^{(k)}(0)$ and $f_\epsilon^{(k)}(0)$ can be recovered from the derivatives $f_{T|x}^{(1+k)}(0)$ and $f_{T|x'}^{(2+k)}(0)$ and the lower derivatives $\{f_\theta^{(j)}(0), f_\epsilon^{(j)}(0)\}_{j<k}$*

Part 1. Let $p_n = p_n(x)$ and $p'_n = p_n(x')$. When $X = x$, the observed distribution of transaction prices will be

$$f_{T|x}(t) \quad = \quad \sum_{n=2}^{\infty} p_n f_{T|n}(t)$$

Using the properties established in the proof of Lemma 1 above, differentiating yields

$$f_{T|x}^{(k)}(0) \quad = \quad \sum_{n=2}^{\infty} p_n \left(\sum_{i=0}^{k-1} f_{y|n}^{(i)}(0) f_\theta^{(k-1-i)}(0) \right) \tag{A5}$$

We know from above that $f_{y|n}^{(m)}(0) = 0$ for $m < n-2$, or $n > m+2$; and that $f_{y|n}^{(n-2)}(0) = n!(f_\epsilon(0))^{n-1}$. So

$$f'_{T|x}(0) \quad = \quad \sum_{n=2}^{\infty} p_n f_{y|n}(0) f_\theta(0) \quad = \quad p_2 \cdot 2 f_\epsilon(0) f_\theta(0)$$

since $f_{y|2}(0) = 2f_\epsilon(0)$ and for $n > 2$, $f_{y|n}(0) = 0$. By the same logic,

$$f''_{T|x'}(0) \quad = \quad \sum_{n=3}^{\infty} p'_n \left(f'_{y|n}(0) f_\theta(0) + f_{y|n}(0) f'_\theta(0) \right) \quad = \quad p'_3 \cdot 6(f_\epsilon(0))^2 f_\theta(0)$$

since $f_{y|n}(0) = 0$ for $n \geq 3$, $f'_{y|3}(0) = 6(f_\epsilon(0))^2$, and $f'_{y|n}(0) = 0$ for $n > 3$. Dividing,

$$\frac{f''_{T|x'}(0)}{f'_{T|x}(0)} \quad = \quad 3\frac{p'_3}{p_2} f_\epsilon(0)$$

and since we observe the left-hand side (and are assumed to know p'_3 and p_2), we can recover $f_\epsilon(0)$; once that is known, we can get $f_\theta(0)$ from $f'_{T|x}(0) = 2p_2 f_\epsilon(0) f_\theta(0)$.

Part 2. From (A5),

$$f_{T|x}^{(1+k)}(0) \quad = \quad \sum_{n=2}^{\infty} p_n f_{T|n}^{(1+k)}(0) \quad = \quad \sum_{n=2}^{\infty} p_n \left(\sum_{i=0}^{k} f_{y|n}^{(i)}(0) f_\theta^{(k-i)}(0) \right)$$

Recall (from the preliminaries in the proof of Lemma 1) that $f_{y|n}^{(m)}$ has no derivatives of f_ϵ higher than $f_\epsilon^{(m-n+2)}(0)$, and that $f_{y|n}^{(m)}(0) = 0$ for $m < n-2$. This means that, knowing $\{p_n\}$ and $\{p'_n\}$, we can write this as

$$f_{T|x}^{(1+k)}(0) \quad = \quad p_2 \cdot 2 f_\epsilon^{(k)}(0) f_\theta(0) + p_2 \cdot 2 f_\epsilon(0) f_\theta^{(k)}(0) + \text{things we already know}$$

where "things we already know" includes the pieces of $f_{y|n}^{(k)}(0)$ other than $2f_\epsilon^{(k)}(0)$ (multiplied by p_2 and $f_\theta(0)$); the terms of $p_2 f_{T|2}^{(1+k)}(0)$ other than the "first" and "last" terms $p_2 f_{y|2}^{(k)}(0) f_\theta(0)$ and $p_2 f_{y|2}(0) f_\theta^{(k)}(0)$; and all the terms in $p_n f_{y|n}^{(1+k)}(0)$ for $n > 2$, which all contain derivatives no higher than $f_\epsilon^{(k-1)}(0)$ and $f_\theta^{(k-1)}(0)$.

Similarly,

$$
\begin{aligned}
f_{T|x'}^{(2+k)}(0) &= \sum_{n=3}^{\infty} p'_n \left(\sum_{i=0}^{k+1} f_{y|n}^{(i)}(0) f_{\theta}^{(k+1-i)}(0) \right) \\
&= p'_3 \cdot 6(k+2) f_{\epsilon}(0) f_{\epsilon}^{(k)}(0) f_{\theta}(0) + p'_3 \cdot 6(f_{\epsilon}(0))^2 f_{\theta}^{(k)}(0) + \text{other things we know}
\end{aligned}
$$

Letting A and B denote the two massive collections of terms we already know the value of, then, we can write

$$
\begin{aligned}
\frac{f_{T|x}^{(1+k)}(0) - A}{2p_2 f_{\theta}(0)} &= f_{\epsilon}^{(k)}(0) + \frac{f_{\epsilon}(0)}{f_{\theta}(0)} f_{\theta}^{(k)}(0) \\
\frac{f_{T|x'}^{(2+k)}(0) - B}{6p'_3 f_{\epsilon}(0) f_{\theta}(0)} &= (k+2) f_{\epsilon}^{(k)}(0) + \frac{f_{\epsilon}(0)}{f_{\theta}(0)} f_{\theta}^{(k)}(0)
\end{aligned}
$$

and, subtracting,

$$
\frac{1}{k+1} \left[\frac{f_{T|x'}^{(2+k)}(0) - B}{6p'_3 f_{\epsilon}(0) f_{\theta}(0)} - \frac{f_{T|x}^{(1+k)}(0) - A}{2p_2 f_{\theta}(0)} \right] = f_{\epsilon}^{(k)}(0)
$$

Once $f_{\epsilon}^{(k)}(0)$ is known, $f_{\theta}^{(k)}(0)$ is the only remaining unknown in the expression for $f_{T|x}^{(1+k)}(0)$, and can be recovered from that, concluding the proof of the lemma. □

This establishes identification if $p_2 > 0 = p'_2$ and $p_3 > 0$. The point is that all the "new" (unknown) terms at each level iteration come only from the "leading" term – $n = 2$ in the case of $p(x)$, and $n = 3$ in the case of $p(x')$. Obviously, if $p_2 > 0$ and $p'_2 = p'_3 = 0 < p'_4$, the same would work, taking one higher derivative of $f_{T|x'}$; and if $p_2 = 0 = p'_2$ but the two had different leading terms, the same would work, but taking higher derivatives of both.

The remaining challenge, then, is if the "leading" term of both $f_{T|x}$ and $f_{T|x'}$ – i.e., the lowest values of n for which $p_n(x) > 0$ and $p_{n'}(x) > 0$, respectively – are the same, for example if $p_2 > 0$ and $p'_2 > 0$. In that case, we synthesize a new distribution with no p_2 term. Specifically, define

$$
g(t) = p'_2 f_{T|x}(t) - p_2 f_{T|x'}(t)
$$

Letting $q_j = p'_2 p_j - p_2 p'_j$,

$$
g(t) = \sum_{n=3}^{\infty} q_n f_{T|n}(t)
$$

If $p'_2 p_3 \neq p_2 p'_3$, then $q_3 \neq 0$, and we can proceed as in the lemma above, with g replacing $f_{T|x'}$. If $p'_2 p_3 = p_2 p'_3$, then $q_3 = 0$, and the "leading term" of g is higher. If $p'_2 p_j = p'_j p_2$ for *every* j, then $\mathbf{p}(x) = \mathbf{p}(x')$, which we already assumed was not true; so g must have some leading term $q_n \neq 0$, giving us our starting point. Thus, for any two distributions of N, $\mathbf{p}(x)$ and $\mathbf{p}(x')$, we can recover $\{f_{\theta}^{(k)}(0), f_{\epsilon}^{(k)}(0)\}_{k=0,1,2,\ldots}$ from the derivatives of $f_{T|x}$ and g; following the logic behind Theorem 1, this means f_{θ} and f_{ϵ} are uniquely determined if they are assumed to be analytic. □

A.4 Proof of Theorem 4 (discrete valuations, unobserved N)

The outline of the proof is identical to the case of observable N. Again, the theorem follows easily by induction once I first prove the following lemma:

Lemma 4. *Fix λ and λ', with $\lambda' > \lambda > 0$.*

1. *The parameters t_0 and e_0 can be recovered from the moments* $\Pr(T = 0|\lambda)$ *and* $\Pr(T = 0|\lambda')$.
2. *For any $k > 0$, the parameters t_k and e_k can be recovered from the moments* $\Pr(T = k|\lambda)$ *and* $\Pr(T = k|\lambda')$ *and the parameters $\{t_j, e_j\}_{j<k}$.*

The proof follows the same outline as that of Lemma 2, just with adjustments to account for N having a known probability distribution rather than being constant.

Part 1: recovering e_0 and t_0. To prove the first part, note that

$$\Pr(T = 0|\lambda) = t_0 \Pr(\epsilon^{(2)} = 0|\lambda) = t_0 \sum_{n=2}^{\infty} \Pr(N = n|\lambda) \Pr(\epsilon^{(2)} = 0|N = n)$$

Given λ, the distribution of N, conditional on being at least 2, is

$$\Pr(N = n|\lambda) = \frac{1}{1 - e^{-\lambda} - \lambda e^{-\lambda}} \frac{e^{-\lambda}\lambda^n}{n!}$$

and so

$$\Pr(T = 0|\lambda) = t_0 \frac{1}{1 - e^{-\lambda} - \lambda e^{-\lambda}} \sum_{n=2}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(e_0^n + n(1 - e_0)e_0^{n-1}\right) \tag{A6}$$

As before, if we take the ratio $\frac{\Pr(T=0|\lambda')}{\Pr(T=0|\lambda)}$, the t_0 term drops out, so given λ' and λ (assumed to be known), $\frac{\Pr(T=0|\lambda')}{\Pr(T=0|\lambda)}$ depends only on e_0. Next, I show that it is strictly increasing in e_0, allowing us to recover e_0. Once e_0 is known, t_0 can be calculated from (A6).

To see that (for $\lambda' > \lambda$) $\frac{\Pr(T=0|\lambda')}{\Pr(T=0|\lambda)}$ is strictly increasing in e_0, note that

$$\frac{\partial}{\partial e_0} \ln \left(\frac{\Pr(T = 0|\lambda')}{\Pr(T = 0|\lambda)} \right) = \frac{\partial}{\partial e_0} \ln \Pr(T = 0|\lambda') - \frac{\partial}{\partial e_0} \ln \Pr(T = 0|\lambda)$$

so it suffices to show that $\frac{\partial}{\partial e_0} (\ln \Pr(T = 0|\lambda))$ is strictly increasing in λ. To show this, first note that

$$
\begin{aligned}
\Pr(T=0|\lambda) &= t_0 \sum_{n=2}^{\infty} \Pr(N=n|\lambda) \Pr(\epsilon^{(2)}=0|N=n) \\
&= t_0 \sum_{n=2}^{\infty} \frac{1}{1-e^{-\lambda}-\lambda e^{-\lambda}} \frac{e^{-\lambda}\lambda^n}{n!} \left(e_0^n + n(1-e_0)e_0^{n-1}\right) \\
&= \frac{t_0}{1-e^{-\lambda}-\lambda e^{-\lambda}} \frac{e^{-\lambda}}{e^{-\lambda e_0}} \left[\sum_{n=2}^{\infty} \frac{e^{-\lambda e_0}(\lambda e_0)^n}{n!} + \lambda(1-e_0) \sum_{n=2}^{\infty} \frac{e^{-\lambda e_0}(\lambda e_0)^{n-1}}{(n-1)!}\right] \\
&= \frac{t_0}{1-e^{-\lambda}-\lambda e^{-\lambda}} \frac{e^{-\lambda}}{e^{-\lambda e_0}} \left[(1-e^{-\lambda e_0}-\lambda e_0 e^{-\lambda e_0}) + \lambda(1-e_0)(1-e^{-\lambda e_0})\right] \\
&= \frac{t_0 e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \left[e^{\lambda e_0} - 1 - \lambda e_0 + \lambda(1-e_0)(e^{\lambda e_0}-1)\right] \\
&= \frac{t_0 e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \left[e^{\lambda e_0} - 1 + \lambda e^{\lambda e_0} - \lambda - \lambda e_0 e^{\lambda e_0}\right]
\end{aligned}
$$

This means that

$$
\begin{aligned}
\frac{\partial}{\partial e_0} \ln \Pr(T=0|\lambda) &= \frac{\lambda e^{\lambda e_0} + \lambda^2 e^{\lambda e_0} - \lambda e^{\lambda e_0} - \lambda^2 e_0 e^{\lambda e_0}}{e^{\lambda e_0} - 1 + \lambda e^{\lambda e_0} - \lambda - \lambda e_0 e^{\lambda e_0}} \\
&= \frac{\lambda^2 e^{\lambda e_0}(1-e_0)}{e^{\lambda e_0} - 1 + \lambda e^{\lambda e_0} - \lambda - \lambda e_0 e^{\lambda e_0}} \\
&= \frac{\lambda^2 (1-e_0)}{1 - e^{-\lambda e_0} + \lambda - \lambda e^{-\lambda e_0} - \lambda e_0}
\end{aligned}
$$

To show that's increasing in λ, we calculate (via the quotient rule)

$$
\begin{aligned}
\frac{\partial}{\partial \lambda}\left(\frac{\partial}{\partial e_0} \ln \Pr(T=0|\lambda)\right) &\propto \left(1 - e^{-\lambda e_0} + \lambda - \lambda e^{-\lambda e_0} - \lambda e_0\right) 2\lambda(1-e_0) \\
&\quad -\lambda^2(1-e_0)\left(e_0 e^{-\lambda e_0} + 1 - e^{-\lambda e_0} + \lambda e_0 e^{-\lambda e_0} - e_0\right) \\
&\propto 2 - 2e^{-\lambda e_0} + 2\lambda - 2\lambda e^{-\lambda e_0} - 2\lambda e_0 \\
&\quad -\lambda e_0 e^{-\lambda e_0} - \lambda + \lambda e^{-\lambda e_0} - \lambda^2 e_0 e^{-\lambda e_0} + \lambda e_0 \\
&= 2 - 2e^{-\lambda e_0} + \lambda - \lambda e^{-\lambda e_0} - \lambda e_0 - \lambda e_0 e^{-\lambda e_0} - \lambda^2 e_0 e^{-\lambda e_0} \\
&\propto 2e^{\lambda e_0} - 2 + \lambda e^{\lambda e_0} - \lambda - \lambda e_0 e^{\lambda e_0} - \lambda e_0 - \lambda^2 e_0
\end{aligned}
$$

For $x \geq 0$, we know from the Taylor expansion that $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \geq 1 + x + \frac{x^2}{2}$; plugging this in,

$$
\begin{aligned}
(2+\lambda-\lambda e_0)\, e^{\lambda e_0} - 2 - \lambda - \lambda e_0 - \lambda^2 e_0 \quad &\geq \quad (2+\lambda-\lambda e_0)\left(1+\lambda e_0 + \tfrac{1}{2}\lambda^2 e_0^2\right) - 2 - \lambda - \lambda e_0 - \lambda^2 e_0 \\
&= \quad 2 + 2\lambda e_0 + \lambda^2 e_0^2 + \lambda + \lambda^2 e_0 + \tfrac{1}{2}\lambda^3 e_0^2 \\
&\qquad -\lambda e_0 - \lambda^2 e_0^2 - \tfrac{1}{2}\lambda^3 e_0^3 - 2 - \lambda - \lambda e_0 - \lambda^2 e_0 \\
&= \quad \tfrac{1}{2}\lambda^3 e_0^2 (1 - e_0) \\
&> \quad 0
\end{aligned}
$$

Thus, $\frac{\partial}{\partial e_0} \ln \Pr(T=0|\lambda)$ is increasing in λ, so as argued above, $\frac{\Pr(T=0|\lambda')}{\Pr(T=0|\lambda)}$ is strictly increasing in e_0 for $\lambda' > \lambda$. This means we can recover e_0 from the observed value of $\frac{\Pr(T=0|\lambda')}{\Pr(T=0|\lambda)}$, then recover t_0 from (A6).

Part 2: recovering e_k and t_k. For the second part, we are assumed to already know $\Pr(T=k|\lambda)$, $\Pr(T=k|\lambda')$, and the parameters $\{t_j, e_j\}_{j<k}$. Now,

$$
\Pr(T=k|\lambda) \quad = \quad \sum_{j=0}^{k} t_{k-j} \Pr(\epsilon^{(2)} = j|\lambda) \tag{A7}
$$

where now

$$
\Pr(\epsilon^{(2)} = j|\lambda) \quad = \quad \textstyle\sum_{n\geq 2} \Pr(N=n|\lambda) \Pr(\epsilon^{(2)} = j|N=n) \quad =
$$

$$
\frac{1}{1-e^{-\lambda}-\lambda e^{-\lambda}} \sum_{n=2}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left((e_j + e_{<j})^n - e_{<j}^n - n e_j e_{<j}^{n-1} + n(1-e_j-e_{<j}) \left((e_j+e_{<j})^{n-1} - e_{<j}^{n-1} \right) \right)
$$

which depends only on $\{e_0, \ldots, e_j\}$. We can rearrange (A7) to

$$
\frac{1}{\Pr(\epsilon^{(2)}=0|\lambda)} \left[\Pr(T=k|\lambda) - \sum_{j=1}^{k-1} t_{k-j} \Pr(\epsilon^{(2)}=j|\lambda) \right] \quad = \quad t_k + t_0 \frac{\Pr(\epsilon^{(2)}=k|\lambda)}{\Pr(\epsilon^{(2)}=0|\lambda)} \tag{A8}
$$

where everything on the left-hand side is already known. If we define $G_k(\lambda)$ as the left-hand side of (A8), then

$$
\frac{1}{t_0} \left[G_k(\lambda') - G_k(\lambda) \right] \quad = \quad \frac{\Pr(\epsilon^{(2)}=k|\lambda')}{\Pr(\epsilon^{(2)}=0|\lambda')} - \frac{\Pr(\epsilon^{(2)}=k|\lambda)}{\Pr(\epsilon^{(2)}=0|\lambda)} \tag{A9}
$$

with the right-hand side depending only on e_k and parameters we already know. Next, I show that the right-hand side of (A9) is strictly increasing in e_k, and therefore invertible, so that e_k is pinned down by (A9). Once e_k is known, t_k can be recovered from (A8), finishing the proof.

To see that (for $\lambda' > \lambda$) the right-hand side of (A9) is strictly increasing in e_k, note that this is equivalent to showing that given $\{e_j\}_{j<k}$,

$$
\frac{\partial}{\partial e_k} \frac{\Pr(\epsilon^{(2)}=k|\lambda')}{\Pr(\epsilon^{(2)}=0|\lambda')} \quad > \quad \frac{\partial}{\partial e_k} \frac{\Pr(\epsilon^{(2)}=k|\lambda)}{\Pr(\epsilon^{(2)}=0|\lambda)}
$$

or that $\frac{\partial}{\partial e_k}\frac{\Pr(\epsilon^{(2)}=k|\lambda)}{\Pr(\epsilon^{(2)}=0|\lambda)}$ is strictly increasing in λ.

Now,

$$
\begin{aligned}
&\Pr(\epsilon^{(2)} = k|\lambda) \\
&= \sum_{n=2}^{\infty} \frac{1}{1-e^{-\lambda}-\lambda e^{-\lambda}} \frac{e^{-\lambda}\lambda^n}{n!} \Pr(\epsilon^{(2)} = k|N = n) \\
&= \frac{1}{1-e^{-\lambda}-\lambda e^{-\lambda}} \sum_{n=2}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left((e_k + e_{<k})^n - e_{<k}^n - n e_k e_{<k}^{n-1} + n(1-e_k-e_{<k})\left((e_k+e_{<k})^{n-1} - e_{<k}^{n-1}\right)\right) \\
&= \frac{1}{1-e^{-\lambda}-\lambda e^{-\lambda}} \sum_{n=2}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left((e_k + e_{<k})^n - e_{<k}^n + n(1-e_k-e_{<k})(e_k+e_{<k})^{n-1} - n(1-e_{<k})e_{<k}^{n-1}\right) \\
&= \frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \Bigg[\sum_{n=2}^{\infty} \frac{\lambda^n}{n!}(e_k+e_{<k})^n - \sum_{n=2}^{\infty} \frac{\lambda^n}{n!} e_{<k}^n \\
&\quad + (1-e_k-e_{<k}) \sum_{n=2}^{\infty} \frac{\lambda^n}{n!} n(e_k+e_{<k})^{n-1} - (1-e_{<k}) \sum_{n=2}^{\infty} \frac{\lambda^n}{n!} n e_{<k}^{n-1}\Bigg] \\
&= \frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \Bigg[\frac{1}{e^{-\lambda(e_k+e_{<k})}} \sum_{n=2}^{\infty} \frac{e^{-\lambda(e_k+e_{<k})}(\lambda(e_k+e_{<k}))^n}{n!} - \frac{1}{e^{-\lambda e_{<k}}} \sum_{n=2}^{\infty} \frac{e^{-\lambda e_{<k}}(\lambda e_{<k})^n}{n!} \\
&\quad + \frac{\lambda(1-e_k-e_{<k})}{e^{-\lambda(e_k+e_{<k})}} \sum_{n=2}^{\infty} \frac{e^{-\lambda(e_k+e_{<k})}(\lambda(e_k+e_{<k}))^{n-1}}{(n-1)!} - \frac{\lambda(1-e_{<k})}{e^{-\lambda e_{<k}}} \sum_{n=2}^{\infty} \frac{e^{-\lambda e_{<k}}(\lambda e_{<k})^{n-1}}{(n-1)!}\Bigg]
\end{aligned}
$$

Continuing to simplify,

$$
\begin{aligned}
&\Pr(\epsilon^{(2)} = k|\lambda) \\
&= \frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \Bigg[\frac{1}{e^{-\lambda(e_k+e_{<k})}} \left(1 - e^{-\lambda(e_k+e_{<k})} - \lambda(e_k+e_{<k})e^{-\lambda(e_k+e_{<k})}\right) \\
&\quad - \frac{1}{e^{-\lambda e_{<k}}} \left(1 - e^{-\lambda e_{<k}} - \lambda e_{<k} e^{-\lambda e_{<k}}\right) + \frac{\lambda(1-e_k-e_{<k})}{e^{-\lambda(e_k+e_{<k})}} \left(1 - e^{-\lambda(e_k+e_{<k})}\right) - \frac{\lambda(1-e_{<k})}{e^{-\lambda e_{<k}}} \left(1 - e^{-\lambda e_{<k}}\right)\Bigg] \\
&= \frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \Bigg[\left(e^{\lambda(e_k+e_{<k})} - 1 - \lambda(e_k+e_{<k})\right) - \left(e^{\lambda e_{<k}} - 1 - \lambda e_{<k}\right) \\
&\quad + \lambda(1-e_k-e_{<k})\left(e^{\lambda(e_k+e_{<k})} - 1\right) - \lambda(1-e_{<k})\left(e^{\lambda e_{<k}} - 1\right)\Bigg] \\
&= \frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \Bigg[e^{\lambda(e_k+e_{<k})} - e^{\lambda e_{<k}} \\
&\quad + \lambda(1-e_k-e_{<k})e^{\lambda(e_k+e_{<k})} - \lambda(1-e_{<k})e^{\lambda e_{<k}}\Bigg] \\
&= \frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}} \left[\left(1+\lambda(1-e_k-e_{<k})\right) e^{\lambda(e_k+e_{<k})} - \left(1+\lambda(1-e_{<k})\right) e^{\lambda e_{<k}}\right]
\end{aligned}
$$

Plugging our expression for $\Pr(\epsilon^{(2)} = 0|\lambda)$ from earlier in for the denominator, this means

$$\begin{aligned}
\frac{\Pr(\epsilon^{(2)} = k|\lambda)}{\Pr(\epsilon^{(2)} = 0|\lambda)} &= \frac{\frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}}\left[\left(1+\lambda(1-e_k-e_{<k})\right)e^{\lambda(e_k+e_{<k})} - \left(1+\lambda(1-e_{<k})\right)e^{\lambda e_{<k}}\right]}{\frac{e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}}\left[(1+\lambda(1-e_0))e^{\lambda e_0} - (1+\lambda)\right]} \\
&= \frac{\left(1+\lambda(1-e_k-e_{<k})\right)e^{\lambda(e_k+e_{<k})} - \left(1+\lambda(1-e_{<k})\right)e^{\lambda e_{<k}}}{(1+\lambda(1-e_0))e^{\lambda e_0} - (1+\lambda)}
\end{aligned}$$

Since the denominator does not depend on e_k,

$$\begin{aligned}
\frac{\partial}{\partial e_k}\frac{\Pr(\epsilon^{(2)} = k|\lambda)}{\Pr(\epsilon^{(2)} = 0|\lambda)} &= \frac{-\lambda e^{\lambda(e_k+e_{<k})} + \left(1+\lambda(1-e_k-e_{<k})\right)\lambda e^{\lambda(e_k+e_{<k})}}{(1+\lambda(1-e_0))e^{\lambda e_0} - (1+\lambda)} \\
&= \frac{\lambda^2(1-e_k-e_{<k})e^{\lambda(e_k+e_{<k})}}{(1+\lambda(1-e_0))e^{\lambda e_0} - (1+\lambda)}
\end{aligned}$$

To show this is increasing in λ, we calculate via the quotient rule

$$\begin{aligned}
&\frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial e_k}\frac{\Pr(\epsilon^{(2)} = k|\lambda)}{\Pr(\epsilon^{(2)} = 0|\lambda)}\right) \\
\propto\quad & \left((1+\lambda(1-e_0))e^{\lambda e_0} - (1+\lambda)\right)\left(2\lambda(1-e_k-e_{<k})e^{\lambda(e_k+e_{<k})} + \lambda^2(e_k+e_{<k})(1-e_k-e_{<k})e^{\lambda(e_k+e_{<k})}\right) \\
& \quad - \left((1-e_0)e^{\lambda e_0} + e_0(1+\lambda(1-e_0))e^{\lambda e_0} - 1\right)\left(\lambda^2(1-e_k-e_{<k})e^{\lambda(e_k+e_{<k})}\right) \\
\propto\quad & \left((1+\lambda(1-e_0))e^{\lambda e_0} - (1+\lambda)\right)\left(2+\lambda(e_k+e_{<k})\right) \\
& \quad - \left((1-e_0)e^{\lambda e_0} + e_0(1+\lambda(1-e_0))e^{\lambda e_0} - 1\right)(\lambda) \\
=\quad & (2+2\lambda(1-e_0))e^{\lambda e_0} - 2(1+\lambda) + \lambda(e_k+e_{<k})(1+\lambda(1-e_0))e^{\lambda e_0} - \lambda(e_k+e_{<k})(1+\lambda) \\
& \quad -\lambda(1-e_0)e^{\lambda e_0} - \lambda e_0(1+\lambda(1-e_0))e^{\lambda e_0} + \lambda \\
=\quad & (2+\lambda(1-e_0))e^{\lambda e_0} + \lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0))e^{\lambda e_0} \\
& \quad -2-\lambda-\lambda(e_k+e_{<k})(1+\lambda)
\end{aligned}$$

Again using the fact that $e^{\lambda e_0} \geq 1 + \lambda e_0 + \frac{1}{2}(\lambda e_0)^2$ from the Taylor expansion of e^x, this is

$$
\begin{aligned}
\geq\quad & (2+\lambda(1-e_0))\left(1+\lambda e_0+\tfrac{1}{2}\lambda^2 e_0^2\right)+\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0))\left(1+\lambda e_0+\tfrac{1}{2}\lambda^2 e_0^2\right) \\
& \quad -2-\lambda-\lambda(e_k+e_{<k})(1+\lambda) \\
=\quad & 2+\lambda(1-e_0)+2\lambda e_0+\lambda^2 e_0(1-e_0)+\lambda^2 e_0^2+\tfrac{1}{2}\lambda^3 e_0^2(1-e_0) \\
& \quad +\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0)) \\
& \quad +\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0))\lambda e_0 \\
& \quad +\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0))\tfrac{1}{2}\lambda^2 e_0^2 \\
& \quad -2-\lambda-\lambda(e_k+e_{<k})(1+\lambda) \\
=\quad & \tfrac{1}{2}\lambda^3 e_0^2(1-e_0)+\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0))\tfrac{1}{2}\lambda^2 e_0^2 \\
& \quad +\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0)+\lambda e_0+\lambda^2 e_0(1-e_0)) \\
& \quad -\lambda(e_k+e_{<k}-e_0)(1+\lambda) \\
=\quad & \tfrac{1}{2}\lambda^3 e_0^2(1-e_0)+\lambda(e_k+e_{<k}-e_0)(\lambda^2 e_0(1-e_0))+\lambda(e_k+e_{<k}-e_0)(1+\lambda(1-e_0))\tfrac{1}{2}\lambda^2 e_0^2 \\
>\quad & 0
\end{aligned}
$$

and so $\frac{\partial}{\partial e_k}\frac{\Pr(\epsilon^{(2)}=k|\lambda)}{\Pr(\epsilon^{(2)}=0|\lambda)}$ is strictly increasing in λ. This means that for $\lambda' > \lambda$, $\frac{\Pr(\epsilon^{(2)}=k|\lambda')}{\Pr(\epsilon^{(2)}=0|\lambda')} - \frac{\Pr(\epsilon^{(2)}=k|\lambda)}{\Pr(\epsilon^{(2)}=0|\lambda)}$ is strictly increasing in e_k, exactly what we needed to be able to recover e_k from (A9) in the text; t_k is then recovered from (A8), concluding the proof. □