Improving Estimation of Labor Market Disequilibrium Through Inclusion of Shortage Indicators

Matthew D. Baird, Lindsay Daugherty, and Krishna B. Kumar

RAND Labor & Population

WR-1175
December 2016
This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the RAND Labor and Population Unit.

RAND working papers are intended to share researchers’ latest findings and to solicit informal peer review. They have been approved for circulation by RAND Labor and Population but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND’s publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.
Improving Estimation of Labor Market Disequilibrium Through Inclusion of Shortage Indicators

Matthew Baird, RAND, mmbaird@rand.org
Lindsay Daugherty, RAND, ldaugher@rand.org
Krishna B. Kumar, RAND and IZA, kumar@rand.org

October 2016

Abstract

While economic studies often assume that labor markets are in equilibrium, there may be specialized labor markets likely in disequilibrium. We develop a new methodology to improve the estimation of a disequilibrium model that incorporates a survey-based shortage indicator into the model and estimation strategy. We demonstrate the gains in information provided by the methodology. We apply the model to the labor market of anesthesiologists, the outcomes of which would be of independent interest. We find improved accuracy in the estimation as well as useful information revealed by the expanded model.

1 The authors are listed in alphabetic order. The American Society of Anesthesiologists (ASA) supported this study. We thank the ASA advisory committee, participants at the RAND Labor & Population brownbag, and Misha Dworsky for comments.
1. Introduction

Labor markets are often assumed to be relatively flexible, with workers receiving wages close to the value of their marginal product of labor, and these wages adjusting in the aggregate to ensure that the supply and demand of labor are equilibrated. However, assumptions of wage flexibility and the resulting equilibrium are difficult to defend in some markets. For example, consider the case of highly specialized segments of the labor market that require years of training and subsequent licensing, resulting in very thin markets. Medical specialties are one important example of such exceptions. Barriers to entry to the profession are both natural, arising from the rigors of qualifying, and also regulated by the relevant associations of professionals, which restrict the supply of labor. Moreover, government involvement in the reimbursement for services and the regulation of the provision of these services and of the facilities that provide them places restrictions on the demand for labor.

However, econometric methodology in measuring disequilibrium has been largely ignored in the microeconomic literature since the models of Maddala and Nelson (1974) and Gourieroux, Laffont, and Monfort, (1980), reemphasized in Gourieroux (2000). We build on these approaches by using an innovative strategy that incorporates additional information from surveys into the likelihood function. Specifically, we propose a disequilibrium model that directly uses indicators correlated with shortage to improve estimation. As a by-product of our model, we also get useful information about the relationship between the shortage indicators and actual shortage, such as the average level of the shortage indicator in equilibrium and how much increased shortage affects the shortage indicator. The former gives insight into the “natural rate” of the indicator (e.g., proportion of workers whose employees are actively attempting to hire more of the same type of workers) in a given industry. The latter helps researchers and policy makers understand how indicators of interest (e.g. fraction of workers working in an office where production is being hampered by insufficient labor supply, such as delayed medical procedures) might be expected to fluctuate with changes in economic trends such as recessions and policy inputs, such as number of medical residencies in the country.

We then demonstrate an application of the model to the labor market for anesthesiologists. The anesthesiology labor market provides an excellent context in which we
might consider the typical assumptions of flexible labor markets to not be valid and where our ability to better evaluate labor market conditions is likely to be important for policy decisions. Shortages in such a critical specialty would have important implications for access to care, leading to waits in hiring, delaying necessary medical procedures, and potentially increasing medical expenditures. On the other hand, a surplus of medical specialists can lead to highly capable, trained, and productive physicians being underutilized, leading to inefficient allocation of human capital, without necessarily improving health outcomes (Baiker and Chandra 2004, Phillips et al. 2005).

There is an open discussion concerning the direction and extent of shortages of medical specialties. Dall et al. (2013) project future demand and supply among various medical specialties, and predict a substantial increase in demand for physician services: a 14% increase in demand for FTE primary care physicians from 2013 to 2025, with an even larger increase for most specialties they examined. They contend that insufficient attention to expanding supply of medical specialists could lead to shortages, causing longer wait times and reduced access to care.

Schubert et al. (2012) estimated a shortage of 2,000 anesthesiologists (the specialists of interest in this paper) in 2007. Schubert et al. (2012) further conclude that there is evidence for persistence shortage at the national level, which seems to have been diminishing over the last 2 decades. They suggest, albeit without making any quantitative statements, that increases in the number of new anesthesiologists, lower compensation and decreased demand due to the recession have led to the decrease in earlier estimated shortages, but warn that smaller residency graduation in the future along with demographic shifts (gender and age) related to willingness to work may exacerbate shortage in the future. Our model allows for each of these factors to affect regional labor market conditions. Using these and other variables to quantify shortage or surplus by state, using econometric methods is the novelty of our approach. In agreement with Schubert et al. (2012), Baird et al. (2015) predicts that incoming residents will not keep pace with retirees, leading to an overall decrease in the number of anesthesiologists over the next decade, despite lack of any indication of decreased demand.

---

2 See Baird et al. (2015) for a more detailed literature review and discussion on this topic.
3 Schubert et al. (2012) also provide a good review of the literature regarding evidence for disequilibrium in the market for anesthesiologists. See also Daugherty et al. (2010).
We administered two surveys, approximately 5 years apart, to anesthesiologists in the United States. We combine this with other data, primarily from the Area Health Resource Files, to evaluate this market. We use Hospital Referral Regions (HRR) as the labor market unit of analysis. HRRs are geographical regions in the United States defined as part of the Dartmouth Atlas Project. They represent regional health care markets with at least one hospital that performs major cardiovascular procedures and neurosurgery. There are currently 304 HRRs in the United States.

We find that, while the baseline disequilibrium model estimates a statistically significant surplus of anesthesiologists in both 2007 and 2013, the expanded disequilibrium model is closer to national equilibrium for both years, with a statistically significant surplus in 2007 but a non-statistically significant surplus or shortage in 2013. We document evidence that the expanded disequilibrium model is not simply estimating differently (the two models are statistically different from each other), but that the additional information yields better out of sample predictions and somewhat improved correlations with other shortage indicators.

The rest of the paper proceeds as follows. Section 2 presents the economic model. Section 3 discusses details of our econometric approach to model disequilibrium and contrasts it with the earlier disequilibrium models. Section 4 applies the model to data from our surveys and secondary data sources to analyze the labor market for anesthesiologists. Section 5 discusses the results, and Section 6 concludes.

2. Economic Model

In the labor for market in year , let quantity of labor demand and quantity of labor supplied be given by:

\[ Q^D_{mt} = X^D_{mt} \beta^D + \epsilon^D_{mt} \]  \hspace{1cm} (1)

\[ Q^S_{mt} = X^S_{mt} \beta^S + \epsilon^S_{mt} \]  \hspace{1cm} (2)

Here, \( Q^D \) and \( Q^S \) denote the total number of full-time equivalent (FTE) workers demanded and supplied respectively. \( X^D \) and \( X^S \) include factors influencing demand and supply, most importantly the wage. Under equilibrium, \( Q^D = Q^S = Q \) is observed, and we

\[ \text{http://www.dartmouthatlas.org/tools/faq/researchmethods.aspx} \]
can estimate the unknown coefficients using standard regression techniques, such as 3SLS, instrumenting for the endogenous wage using excluded variables in each equation. However, if the market is in disequilibrium, then we are unable to observe both the quantity demanded and quantity supplied jointly for a given market and year; we only observe the minimum of the labor demand and labor supply. This situation is illustrated in Figure 1.

Figure 1: Disequilibrium in Labor Market (Daugherty et al. 2010)

In disequilibrium, only \( Q = \min(Q^D, Q^S) \) is observed, denoted by the thick black line in Figure 1. If \( Q^D > Q^S \) then there is excess demand (shortage), and only \( Q^S \) is observed. If \( Q^D < Q^S \) then there is excess supply (surplus), and only \( Q^D \) is observed. The problem is that, while \( Q \) is observed, we do not know whether we have observed \( Q^D \) or \( Q^S \). This uncertainty must be accounted for in the estimation strategies.

Further assume that we observe at least one indicator \( A \), which we designate as a “shortage indicator” given its posited relationship with the true (unobserved) shortage:

\[
A_{mt} = \gamma_0 + \gamma_1 (Q_{mt}^D - Q_{mt}^S)/\rho_m + \nu_{mt} \tag{3}
\]

We will discuss the actual shortage indicator we use from our surveys below. Here, \( \rho_m \) is the population in the labor market \( m \), so that the shortage indicator is normalized to depend on per capita labor shortage. \( \gamma_0 \) captures the value of \( A \) expected in equilibrium. \( \gamma_1 \) describes the relationship between labor shortage and the observed indicator. It measures the average increase in the shortage indicator for each additional FTE worker demanded in
excess of supply, per capita. There may be more than one such shortage indicator available to the researcher.

3. Econometric Models

As mentioned above, the primary econometric challenge is to discern whether the observed quantity of anesthesiologist labor is supply or demand or both (in the case of equilibrium). If $Q^D > Q^S$ then there is a situation of excess demand or shortage, and only $Q^S$ is observed. If $Q^D < Q^S$ then there is a situation of excess supply or surplus, and only $Q^D$ is observed. Assigning probabilities to the observed quantity being supply or demand is the primary concern of the models discussed in this section.

We consider three models: equilibrium, basic disequilibrium models after the example of Maddala and Nelson (1974) and Gourieroux, Laffont, and Monfort, (1980), reemphasized in Gourieroux (2000), and our expanded disequilibrium models. Our primary comparison is between the basic class of disequilibrium models (hereafter referred to as MN, after Maddala and Nelson 1974) and our expanded model that includes information from shortage indicators (hereafter, the Expanded Disequilibrium likelihood, or ED). In the literature on basic disequilibrium models, the focus is on the test of whether the market is in equilibrium or not; as this is not the focus of this paper, we rely simply on whether the expected aggregate excess demand is significantly different from zero as our inference for the hypothesis of overall equilibrium. Individual labor markets may be in disequilibrium even while the national market is on average in equilibrium.

We summarize the likelihood of the three major dependent variables $q$ (observed labor), $w$ (observed log wages) and $a$ (observed shortage indicators) in Equation 4 (with subscripts suppressed):

$$
\Pr(Q = q, A = a, W = w) = \Pr (Q = q|A = a, W = w) \Pr (A = a|W = w) \Pr (W = w)
$$

The equilibrium model and MN assume that $A$ is unobserved or does not impact quantity, and so it drops out of the likelihood. Wage is endogenous to the system, and as such, it belongs in the likelihood function. However, to reduce the dimension of the
parameter space for the maximum likelihood search algorithm, we estimate the first stage regressions of log wages on the determinants of wages (which include the excluded instruments of the demand and supply functions), and use the predicted wages in the maximum likelihood of observed quantities. Doing so with valid excluded variables (demand and supply shifters) removes the endogeneity that arises from the relationship between wages and hours. We estimate the standard errors using block bootstrapping, and given that the first stage is included in the bootstrapping procedure, the standard errors are still correct with this two-stage estimation procedure. From now on, we will suppress the dependency on wages and consider it implied, and suppress the additional estimation of \( \Pr(W = w) \), which is the same across the three classes of models.

3.1. Equilibrium

If the markets are in equilibrium, then \( Q_{mt} = Q^D_{mt} = Q^S_{mt} \) and the model may be estimated by full-information maximum likelihood. However, it is more direct to pursue the closed-form solutions that GMM offers through Two Stage Least Squares (2SLS) and Three Stage Least Squares (3SLS). The usual approach of instrumenting is necessary to overcome the simultaneity of wage in the demand and supply equation. We estimate the equilibrium model as a frame of reference for the parameter values of the disequilibrium models rather than as a test of whether the markets are in equilibrium.

3.2. Basic Disequilibrium Model (MN)

The basic disequilibrium model assumes that \( Q^D_{mt} \neq Q^S_{mt} \). We use our own notation to the problem as set forth by Madalla and Nelson (1974) and others. The likelihood function (suppressing \( m \) and \( t \) subscripts, and the conditioning on observed \( X \), including log wages) without incorporating the shortage indicator may be expressed as:

\[
\Pr(Q = q) = \Pr(Q = q|Q^D > Q^S) \Pr(Q^D > Q^S) + \Pr(Q = q|Q^D < Q^S) \Pr(Q^D < Q^S)
\]

\[
= \frac{\Pr(Q = q, Q^D > Q^S)}{\Pr(Q^D > Q^S)} \Pr(Q^D > Q^S) + \frac{\Pr(Q = q, Q^D > Q^S)}{\Pr(Q^D > Q^S)} \Pr(Q^D < Q^S)
\]
\[ \text{Pr}(Q = q, Q^D > Q^S) + \text{Pr}(Q = q, Q^D < Q^S) \]
\[ = \text{Pr}(Q^S = q, Q^D > q) + \text{Pr}(Q^D = q, Q^S > q) \]
\[ = \text{Pr}(Q^S = q) \text{Pr}(Q^D > q) + \text{Pr}(Q^D = q) \text{Pr}(Q^S > q) \]  \hspace{1cm} (5)

Substituting Equations 1 and 2 into Equation 5, and assuming the error terms are normally distributed, the likelihood becomes:

\[ \frac{1}{\sigma_{eS}} \phi \left( \frac{q - X^S \beta^S}{\sigma_{eS}} \right) \left(1 - \phi \left( \frac{q - X^D \beta^D}{\sigma_{eD}} \right) \right) + \frac{1}{\sigma_{eD}} \phi \left( \frac{q - X^D \beta^D}{\sigma_{eD}} \right) \left(1 - \phi \left( \frac{q - X^S \beta^S}{\sigma_{eS}} \right) \right) \]  \hspace{1cm} (6)

We then choose the parameters that maximize the log-likelihood function \( \sum_{i=1}^{N} \ln (\text{Pr}(Q = q)) \). After we estimate the parameters of the model, we are able to estimate the probability of each market being in shortage as well as the expected shortage for each market as well as in aggregate. These derivations are provided in the Appendix.

### 3.3. Expanded Disequilibrium Model (ED)

Alternatively, we can incorporate information from shortage indicators into our likelihood. Recall from Equation 3 that the shortage indicator is posited to be a function of the actual excess demand. The likelihood function expands on MN as follows:

\[ \text{Pr}(Q = q, A = a) = \text{Pr}(Q = q \mid A = a) \text{Pr}(A = a) \]
\[ = \left[\text{Pr}(Q^S = q \mid A = a) \text{Pr}(Q^D > q \mid A = a) \right. \]
\[ + \text{Pr}(Q^D = q \mid A = a) \text{Pr}(Q^S > q \mid A = a)] \times \text{Pr} \]  \hspace{1cm} (7)

Considering the first element, and substituting in from Equations 2 and 3, we have

\[ \text{Pr}(Q^S = q \mid A = a) = \text{Pr}(X^S \beta^S + e^S = q |_\gamma_0 + \gamma_1 (Q^D - Q^S)/p + v = a) \]  \hspace{1cm} (8)

Given our normality assumption and independence of the error terms, the conditional normal distribution takes on the form:
\[ Q|A \sim N \left( \mu_Q + \frac{\sigma_{QA}}{\sigma_A^2} (a - \mu_A), \sigma_Q^2 \left( 1 - \frac{\sigma_{QA}^2}{\sigma_Q^2 \sigma_A^2} \right) \right) \] (10)

In our case,

\[ \mu_Q = X^S \beta^S \] (11)
\[ \sigma_{QA} = \text{Cov}(X^S \beta^S + \epsilon^S, \gamma_0 + \gamma_1 (X^D \beta^D + \epsilon^D - X^S \beta^S - \epsilon^S)/p + v) = -\gamma_1 \sigma_{\epsilon_S}/p \] (12)
\[ \sigma_A^2 = \frac{\gamma_1^2 (\sigma_{\epsilon_S}^2 + \sigma_{\epsilon_D}^2)}{p^2} + \sigma_v^2 \] (13)
\[ \mu_A = \gamma_0 + \gamma_1 (X^D \beta^D - X^S \beta^S)/p \] (14)

Then it follows that,

\[ \text{Pr}(Q^S = q | A = a) = \phi \left( \frac{q - X^S \beta^S - \mu_{CS}}{\sigma_{CS}} \right) / \sigma_{CS} \] (15)

where,

\[ \mu_{CS} = \frac{-\gamma_1 \sigma_{\epsilon_S}^2}{\gamma_1^2 (\sigma_{\epsilon_S}^2 + \sigma_{\epsilon_D}^2)/p^2 + \sigma_v^2} (a - \gamma_0 - \gamma_1 (X^D \beta^D - X^S \beta^S)/p) \] (16)
\[ \sigma_{\epsilon_S}^2 = \sigma_{\epsilon_S}^2 \left( 1 + \frac{\gamma_1^2 \sigma_{\epsilon_S}^2}{p^2} \right) \] (17)

We also may define:

\[ \mu_{CD} = \frac{\gamma_1 \sigma_{\epsilon_D}^2}{\gamma_1^2 (\sigma_{\epsilon_S}^2 + \sigma_{\epsilon_D}^2)/p^2 + \sigma_v^2} (a - \gamma_0 - \gamma_1 (X^D \beta^D - X^S \beta^S)/p) \] (18)
\[ \sigma_{\epsilon_D}^2 = \sigma_{\epsilon_S}^2 \left( 1 - \frac{\gamma_1^2 \sigma_{\epsilon_D}^2}{p^2} \right) \] (19)

Following a similar approach for the other elements, the likelihood is given by

\[ \text{Pr}(Q = q, A = a) = \left[ \frac{1}{\sigma_{CS}} \phi \left( \frac{q - X^S \beta^S - \mu_{CS}}{\sigma_{CS}} \right) \left( 1 - \phi \left( \frac{q - X^D \beta^D - \mu_{CD}}{\sigma_{CD}} \right) \right) \right] \]
\[
\frac{1}{\sigma_{CD}} \phi \left( \frac{q - X^D \beta^D - \mu_{CD}}{\sigma_{CD}} \right) \left( 1 - \varphi \left( \frac{q - X^S \beta^S - \mu_{CS}}{\sigma_{CS}} \right) \right)
\times \varphi \left( \frac{a - \gamma_0 - \gamma_1 (X^D \beta^D - X^S \beta^S)/p}{\sigma_A} \right) / \sigma_A
\] 

(20)

The expanded likelihood bears a lot of similarity to the likelihood of MN, Equation 6. The additional information from the shortage indicator adjusts the likelihood in an intuitive way. For example, consider the element of the likelihood function representing shortage:

\[
1 - \phi \left( \frac{q - X^D \beta^D - \mu_{CD}}{\sigma_{CD}} \right)
\] 

(21)

This is the probability that demand exceeds (observed quantity) supply. Higher values of \(X^D \beta^D\) increase this probability, which is true for both MN and EDL. For EDL, so do higher values of \(\mu_{CD}\). This doesn’t necessarily occur when \(a\), the shortage indicator, is high, but only if it exceeds the predicted shortage indicator conditional on excess demand. In fact, if the shortage indicator is equal to the expected shortage indicator, then this portion of the likelihood is identical to the one for MN. However, if they are not equal, the there is additional information to be gleaned and the likelihood is adjusted. In other words, the shortage indicator contains no new information that is not already contained in the excess demand. For example, consider the case when \(a\) exceeds the expected \(a\); there will be a higher value for Expression 21 and therefore more weight will be put on matching \(X^S \beta^S\) to \(q\). This is because it will be assumed that the labor market is more likely to be in the state of excess demand, and the observed \(q\) is the labor supply, rather than the labor demand.

Each element of the likelihood function Equations 6 and 20 have a similar potential adjustment depending on the shortage indicator for each labor market. If in fact the shortage indicator has the posited relationship with excess demand, then the expansion of the likelihood function from Equation 6 to Equation 20 will yield more accurate measurements of the parameters of the model by more accurately discriminating between cases of shortage and surplus in each labor market, and matching the observed quantity to the appropriate independent variables.

There is also the last element of ED in Equation 20, which differs from MN in Equation 6. This serves to estimate the parameters of the shortage indicator function.
After we estimate the parameters, we also adjust the expectations of demand and supply, and thus the expected shortage. When we incorporate the shortage indicators, and using the identity in Equation 10 again, we have

$$E[Q^D | q, x^D, x^S, A] = E[Q^D | q, x^D, x^S] + \frac{\sigma_{Q^d A}}{\sigma^2_A} (\alpha - E[A | q, x^D, x^S]) \quad (22)$$

Note that the first element is just the expectation under MN. The second element shifts the expectation depending on how the shortage indicator varies from the predicted shortage indicator. For example, if the observed shortage indicator exceeds the predicted one, and given we expect a positive correlation to exist between quantity demanded and the shortage indicator, then we would increase the expectation above that given by MN in this case. Similar to the likelihood function, if the observed shortage indicator is exactly equal to the predicted one by MN, then it contains no new information and we don’t shift the expectation at all.

Equation 23 gives expected labor supply. The appendix contains details on estimation of Equation 22.

$$E[Q^S | q, x^D, x^S, A] = E[Q^S | q, x^D, x^S] - \frac{\sigma_{Q^s A}}{\sigma^2_A} (\alpha - E[A | q, x^D, x^S]) \quad (23)$$

Thus far, we have only considered the case where there is a single shortage indicator. This model may be expanded to include more than one shortage indicator. That can either be done as a vector of shortage indicators that expand the likelihood, or by creating an index that collapses the shortage indicators into one.

4. Application to the Labor Market of Anesthesiologists

We use the example of the labor market of anesthesiologists to examine the differences between MN and ED. Anesthesiology provides an appropriate context because there are market imperfections that can lead to the anesthesiologist labor markets to not be in equilibrium at any given point in time. One reason is the lag with which supply is able to respond to perceived needs; it takes years for a potential anesthesiologist to go through medical school and complete a residency in anesthesiology. Should the market have a large demand shock for anesthesiologist services, even if hospitals can offer higher wages, it will
not speed up the process of reaching a new equilibrium. Likewise, adjustment might be slow in the face of negative demand shocks, as anesthesiologists may be protected by long-term contracts. Adjustment is likely to be slow on the intensive margin as well. Our surveys reveal numerous cases where anesthesiologists were unable or unwilling to increase their number of hours, even with an increase in their pay. Around 27% of anesthesiologists we surveyed responded that they would not increase their hours because they did not have any more time available. Only around 37% said they would be willing to increase their hours if their compensation was high enough. When asked why they would not increase hours for any compensation, answers included personal reasons such as family reasons and the need for work-life balance. However, some of the replies indicated inability to increase hours due to institutional restrictions, including already operating at the maximum allowed number of hours.

There are also potential barriers to equilibrium being attained from the demand side. Hospitals and the medical industry in general operate under heavy regulation. Furthermore, HMOs and pay-for-service arrangements, such as fixed or capped prices for healthcare services, create wedges between market clearing wages and what can actually be offered to anesthesiologists.

Data for the variables contained in labor supply comes primarily from two surveys we administered to anesthesiologists, first in 2007 and then in 2013. We refer to these surveys as the RAND Surveys. They are described in more detail in Section 4.4, as well as in Baird et al. 2015. We aggregate the data to the HRR labor market level by year. Data for variables contained in labor demand come primarily from external data sources, and in particular the Area Health Resource File (AHRF) which we crosswalk to the HRR. Given our small number of observations (180 labor markets for which we have sufficient data in 2 different years for 360 observations), we aimed for parsimony in constructing the labor demand and labor supply functions. The results are not very sensitive to the inclusion of more covariates, and the variables included seemed a priori to be the more relevant factors.

---

7 Robinson et al. (2004), Madison (2004), Hillman (1987)
4.1 Labor Demand Function

The primary variable affecting demand is the average log wage of anesthesiologists in the HRR. Increased wages make anesthesiologists more expensive, decreasing demand for their services. The RAND Surveys provide us with wage data. We also include the log of the total number of surgeries in the HRR (irrespective of whether an anesthesiologist participated in the surgery or not). This is a good measure of demand for health services for which anesthesiologists would be required. We include the log of the population in the geography covered by the HRR as well as the log of median household income of that population. Increases in either population or income in the population in the market should increase the demand for anesthesiologist hours. Demand is also modeled as a function of the number of Certified Registered Nurse Anesthetists (NAs), interacted with the opt-out status of state (states where NAs are able to perform anesthesia unsupervised). By including NAs we account for the complementarity in the production of anesthesia services. In opt-out states, NAs may serve as more of substitutes for anesthesiologists.\(^8\) The number of NAs is derived from the AHRF. Surgeries, population, income, and the number of NAs are taken from the AHRF.

We also include the local unemployment rate, available from the Bureau of Labor Statistics Local Area Unemployment Statistics files. Finally, we also include a year dummy for 2013 to allow for different baseline aggregate demand.

4.2 Labor Supply Function

As with the labor demand function, the labor supply function for anesthesiologists contains variables that affect supply on the intensive or extensive margin. The first and primary variable is the average wage in the market. Higher wages induce current anesthesiologists to work more hours, and for more anesthesiologists to move to areas of high demand. The coefficient on log wages is a function of the labor supply elasticity. In the RAND Surveys, we asked each respondent for the wage increase necessary to induce a 10% increase in work, from which we can estimate an individual labor supply elasticity. See Daugherty et al. (2010) for details concerning how we estimate the elasticity from the

---

\(^8\) Kalist et al. (2011), Kane and Smith (2004)
questions and the resulting distribution of elasticities by state. Rather than estimate the 
coefficient on log wages in the supply function using appropriate demand shifters as 
instrunents, we decide to use the HRR-averaged survey elasticities directly. Given our labor 
supply model is a level-log model, the elasticity is equal to the coefficient on log wages 
divided by the quantity, or equivalently, the coefficient on log wage is equal to the elasticity 
multiplied by the quantity. Thus, we multiply the elasticity, the quantity, and log-wages and 
subtract this product from the observed quantity for labor supply. This simplifies the 
analysis by requiring fewer assumptions on valid instruments.

From the RAND surveys we include additional labor supply factors: the fraction of 
anesthesiologists that are male (male anesthesiologists are more likely to work more hours 
than females), the fraction of anesthesiologists working fewer than 30 hours (which reveals 
both work preferences of the local anesthesiologist population and the available capacity to 
increase labor hours), the fraction of anesthesiologists working in an urban area (making 
labor hour increases easier with smaller transportation costs, and also potentially related to 
anesthesiologist living preferences and thus the labor supply extensive margin). We also 
include the local unemployment rate, the log population in the HRR, and a dummy for the 
year 2013 to allow for overall time-dependent shifts in labor supply.

4.3 Identification of the Elasticity of Labor Demand

Wages and labor demand are jointly determined, so that estimation of the coefficient on 
wages in the labor demand function (which is proportional to the underlying labor demand 
elasticity) is endogenous. We identify the coefficient based on the supply shifters: excluded 
variables that are in the labor supply function that serve to map out the slope of the labor 
demand function with respect to wages, and hence the coefficient on wages.

Our excluded instruments for the elasticity of labor demand are the fraction of 
anesthesiologists in the market that are working part-time, the fraction that are female, and 
the fraction working in an urban area. We argue that each of these has an effect on labor 
supply, as described in Section 4.2, but has no independent effect on labor demand. It is 
hard to imagine how the gender of those providing services or the part-time nature of work

---

9 Baird et al. (2015)
would directly affect the demand for anesthesiologist services. Justifying the fraction working in an urban area is potentially more difficult, as a higher urban concentration might increase demand for services. However, the primary avenues through which it would affect demand—higher population and lower income—are already included in the demand function. Furthermore, we find the results not sensitive to the inclusion of this instrument. In the 3SLS equilibrium model, the average elasticity with urban included as an excluded instrument is -2.8. If it is included in both demand supply functions, the average elasticity is estimated to be -2.9.

4.4. Data

We conducted detailed surveys of members of the American Society of Anesthesiologists (ASA) in 2007 and then again in 2013. Of the 29,158 ASA members (who were not residents) invited to respond for the 2013 survey, 6,825 of did so, which yielded a response rate of 23.41%. The 6,825 respondents represent a sample of the total of 42,230 anesthesiologists practicing in the United States. To correct for non-response bias in the survey and differences between ASA members and the larger anesthesiologist population, we condition non-response on observed covariates, and create weights to aggregate to the state and national levels. Details about the survey respondents and their characteristics can be found in Baird et al. (2015).

Although there are 304 HRRs in the United States, we only include those for which we have sufficient number of observations to estimate the averages within the HRR. For our purposes, we only include HRRs for which we have at least 5 survey respondents or over 25% of all anesthesiologists in the HRR responding to our surveys. We only keep HRRs for which we have data for both years of the survey. This leaves us with a final sample of 180 markets in 2 years, or 360 total market/year observations. We have their working zip code in the RAND Surveys; HRRs are defined as collections of zip codes, so we can easily aggregate the values up to the HRR level for each of these variables.

For the AHRF and Local Area Unemployment Series data, variables are defined at the county level. We crosswalk each HRR zip codes to the counties, and create a weighted aggregation depending on the relative populations of the counties included in the zip codes.
We also will examine shortage indicators derived from our survey for our ED model. While our estimation procedure will only investigate one indicator, we will later contrast it with three others. Our primary shortage indicator is the fraction of ANs that work in a facility that are actively trying to hire more ANs. The values range from 0 (no ANs in that HRR work in a facility trying to hire more ANs) to 1 (all ANs in that HRR work in a facility trying to hire more ANs). The average is about half, which is to say that the average HRR has half of the ANs working in such a facility. The standard deviation is relatively large as well at around 0.2, suggesting a significant amount of variation in this variable across HRRs, providing good variation for our analysis and suggesting that there may be differences in the likelihood of a given market being in equilibrium, shortage, or surplus. Figure 1 presents the by-HRR distribution of this shortage indicator, showing considerable amount of variation.

Figure 1: Proportion of ANs working in facilities trying to hire more ANs, by HRR

The other three variables we look at are the fraction of ANs that work in a facility that would prefer more ANs hired to cover current workload, the fraction of ANs in the HRR that report that they have increased hours worked appreciably in the past 3 years, and the fraction of ANs that would increase their hours for a sufficient increase in pay.

Table 1 presents the market-level summary statistics.
Table 1: HRR-level Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total AN FTEs</td>
<td>210.5</td>
<td>223</td>
<td>15.67</td>
<td>1685</td>
</tr>
<tr>
<td>Wage</td>
<td>142.7</td>
<td>21.11</td>
<td>81.45</td>
<td>226.4</td>
</tr>
<tr>
<td>Total surgeries (1,000s)</td>
<td>135.4</td>
<td>111.8</td>
<td>17.36</td>
<td>645.2</td>
</tr>
<tr>
<td>Median Household Income</td>
<td>27423</td>
<td>8787</td>
<td>8707</td>
<td>71990</td>
</tr>
<tr>
<td>NAs x opt out state</td>
<td>0.034</td>
<td>0.0968</td>
<td>0</td>
<td>0.914</td>
</tr>
<tr>
<td>NAs x opt in state</td>
<td>0.163</td>
<td>0.189</td>
<td>0</td>
<td>1.094</td>
</tr>
<tr>
<td>Fraction working under 30 hours/week</td>
<td>0.0592</td>
<td>0.0649</td>
<td>0</td>
<td>0.348</td>
</tr>
<tr>
<td>Fraction female</td>
<td>0.211</td>
<td>0.121</td>
<td>0</td>
<td>0.509</td>
</tr>
<tr>
<td>Fraction working urban</td>
<td>0.929</td>
<td>0.14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Population</td>
<td>1.45E+06</td>
<td>1.30E+06</td>
<td>229360</td>
<td>1.02E+07</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>3.488</td>
<td>1.542</td>
<td>0.512</td>
<td>9.206</td>
</tr>
<tr>
<td>Elasticity of labor supply</td>
<td>0.355</td>
<td>0.146</td>
<td>0</td>
<td>0.599</td>
</tr>
<tr>
<td>Work in facility trying to hire more ANs</td>
<td>0.482</td>
<td>0.194</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Work in facility that would prefer more ANs to cover current workload</td>
<td>0.371</td>
<td>0.191</td>
<td>0.914</td>
<td></td>
</tr>
<tr>
<td>Have increased hours in past 3 years</td>
<td>0.487</td>
<td>0.191</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Would increase hours for sufficient increase in pay</td>
<td>0.397</td>
<td>0.162</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*360 Observations

4.5. Results

We estimate the models using Maximum Likelihood. We use the Nelder-Mead simplex search algorithm, starting once from the equilibrium 3SLS parameter values and once from a perturbation of these starting values. We tested starting from up to 10 different initial starting values but found no changes in the convergence points. In fact, in almost all cases across many different versions and data pulls, the second initial values starting yields the same converged parameters as the first. We include the second only as back up against a local maximum in the bootstrapping. We bootstrap all of the parameters by taking random draws of the RAND Survey respondents and reconstructing the HRRs with those respondents, following the same inclusion rules as before.

Table 2 presents the estimated coefficients of the demand and supply models.
For the ED model, we have additional parameters related to the shortage indicator as seen in Equation 3. The equation relates the shortage indicator to the excess demand per 1,000 residents. These parameter estimates are presented in Table 3. γ₀ estimates the expected value of the shortage indicator for a labor market in equilibrium. The value we

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2SLS</th>
<th>3SLS</th>
<th>MN</th>
<th>ED</th>
</tr>
</thead>
<tbody>
<tr>
<td>log wage</td>
<td>-583.63</td>
<td>-574.02</td>
<td>-539.57</td>
<td>-676.68</td>
</tr>
<tr>
<td>log surgery</td>
<td>(-716.95,129.7)</td>
<td>(-669.28,149.02)</td>
<td>(-1228.4,95.36)</td>
<td>(-1149.1,174.92)</td>
</tr>
<tr>
<td>log HH income</td>
<td>53.46</td>
<td>49.17</td>
<td>81.63</td>
<td>89.94</td>
</tr>
<tr>
<td></td>
<td>(15.03,93.27)</td>
<td>(17.86,86.75)</td>
<td>(-5.7884,123.86)</td>
<td>(-20.165,127.82)</td>
</tr>
<tr>
<td>NAs x opt out</td>
<td>-6.396</td>
<td>7.80</td>
<td>-34.83</td>
<td>-57.03</td>
</tr>
<tr>
<td></td>
<td>(-36.12,31.48)</td>
<td>(-22.48,39.37)</td>
<td>(-95.17,23.23)</td>
<td>(-103.33,30.32)</td>
</tr>
<tr>
<td>NAs x opt in</td>
<td>(-247.14,-24.09)</td>
<td>(-218.38,22.08)</td>
<td>(-309.66,101.25)</td>
<td>(-333.18,125.77)</td>
</tr>
<tr>
<td>log population</td>
<td>212.52</td>
<td>210.37</td>
<td>184.67</td>
<td>161.42</td>
</tr>
<tr>
<td></td>
<td>(174.70,247.42)</td>
<td>(175.59,239.39)</td>
<td>(144.16,264.84)</td>
<td>(120.03,265.95)</td>
</tr>
<tr>
<td>Unemp. rate</td>
<td>-17.57</td>
<td>-19.70</td>
<td>-19.18</td>
<td>-14.05</td>
</tr>
<tr>
<td></td>
<td>(-26.69,-12.06)</td>
<td>(-28.53,-13.21)</td>
<td>(-34.37,-5.86)</td>
<td>(-32.21,-3.30)</td>
</tr>
<tr>
<td>Year 2013</td>
<td>-36.28</td>
<td>-33.87</td>
<td>-14.70</td>
<td>-23.75</td>
</tr>
<tr>
<td></td>
<td>(-47.00,16.03)</td>
<td>(-45.14,16.93)</td>
<td>(-76.18,35.13)</td>
<td>(-73.91,39.46)</td>
</tr>
<tr>
<td>Constant</td>
<td>-310.88</td>
<td>-425.33</td>
<td>-75.13</td>
<td>1053.60</td>
</tr>
<tr>
<td></td>
<td>(-4124.6,178.44)</td>
<td>(-4139.8,15.71)</td>
<td>(-3492.36,83.2)</td>
<td>(-3713.5,3320.4)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Supply</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>log population</td>
<td>-204.27</td>
<td>-206.21</td>
<td>-32.66</td>
<td>-24.65</td>
</tr>
<tr>
<td></td>
<td>(-236.67,-171.09)</td>
<td>(-238.62,-171.95)</td>
<td>(-55.46,13.39)</td>
<td>(-50.92,18.71)</td>
</tr>
<tr>
<td>Unemp. rate</td>
<td>14.60</td>
<td>15.72</td>
<td>8.58</td>
<td>6.78</td>
</tr>
<tr>
<td></td>
<td>(5.17,29.04)</td>
<td>(5.98,29.97)</td>
<td>(-5.41,23.78)</td>
<td>(-5.90,19.90)</td>
</tr>
<tr>
<td></td>
<td>(-60.31,10.00)</td>
<td>(-61.00,8.28)</td>
<td>(-69.13,7.87)</td>
<td>(-64.36,8.22)</td>
</tr>
<tr>
<td>Constant</td>
<td>2668</td>
<td>2651.1</td>
<td>372.62</td>
<td>285.97</td>
</tr>
<tr>
<td></td>
<td>(2204.5,3054.4)</td>
<td>(2207.6,3036.7)</td>
<td>(-189.8,693.36)</td>
<td>(-237.89,651.45)</td>
</tr>
</tbody>
</table>

*Bootstrapped 95% confidence interval in parentheses
obtain of 0.472 is slightly lower than the observed value in Table 1 of 0.482, giving us our first indication that this labor market might in aggregate be in surplus, or have excess supply. The parameter is significantly different from zero.

\( \gamma_1 \) tells us that for each additional FTE shortage of ANs per 1,000 residents, we expect the shortage indicator to decrease by around 0.27. Note that the average number of FTE ANs in an HRR per 1,000 residents is 0.14, with a minimum of 0.03 and a maximum of 0.47. Thus, a unit increase in the demand for number of FTE ANs per 1,000 is very large. An increase of 5% of the AN per 1,000 average (a moderate shift up in demand) is 0.007, for a marginal effect of an increase in the shortage indicator of 0.002. This seems a reasonable value. However, the result is not statistically significant, as the confidence interval crosses zero. This is an issue we will have in other cases as well, given the small number of observations, and an even smaller number of ANs in a given HRR yielding large variation in the sample which, gives us large confidence intervals.

Table 3: Equation 3 Estimated Parameters

| Work in facility trying to hire more ANs | \( y_0 \) | 0.472 |
| | | (0.459, 0.497) |
| | \( y_1 \) | 0.269 |
| | | (-0.464, 0.395) |

*Bootstrapped 95% confidence interval in parentheses

With the estimated coefficients and other parameters, we are able to estimate elasticities and expected excess demand (see Table 4). We do so by year. Note that both models estimate a surplus of ANs in 2007, with both being significantly different from zero. However, the MN model predicts a substantially larger surplus. With a total national working population of over 30,000, both models predict that in 2007 there was a surplus that was substantial. For 2013, the MN model continues to estimate a statistically significant surplus of ANs, while the ED model estimates a very small and not statistically different from equilibrium shortage of ANs. Supply elasticities, coming from the survey directly, are accurately measured. Demand elasticities, coming from our MLE, are relatively large although not statistically different from zero.
Table 4: Estimated Parameters from Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Excess Demand 2007</th>
<th>Excess Demand 2013</th>
<th>Demand Elasticity</th>
<th>Supply Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MN</td>
<td>-4870.30</td>
<td>-2217.10</td>
<td>-2.56</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>(-19653, -1426.1)</td>
<td>(-16509, -792.72)</td>
<td>(-5.69, 0.44)</td>
<td>(0.34, 0.37)</td>
</tr>
<tr>
<td>ED</td>
<td>-2995.90</td>
<td>289.13</td>
<td>-3.22</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>(-18505, -1054.7)</td>
<td>(-16470, 937.12)</td>
<td>(-5.15, 0.82)</td>
<td>(0.34, 0.37)</td>
</tr>
</tbody>
</table>

Figures 2 and 3 present the estimated shortage by HRR for the two models. The results are very similar, but do differ from each other. The same general trends are present, but there is a lower level of surplus estimated and some sorting changes.

Figure 2: MN Estimated Expected Excess Demand for 2013

![Figure 2](image1.png)

Figure 3: ED Estimated Expected Excess Demand for 2013

![Figure 3](image2.png)
In addition to comparing the coefficients and predictions of the model, we can implement three different tests to compare the models after estimation.

First, we do inference on whether the two models differ from each other empirically. We can do this in two ways. MN is a special case of ED where $\gamma_1 = 0$. In that case, the shortage indicator contains no additional information and the likelihood collapses to MN.\(^{10}\) Thus, we can do a likelihood ratio test of the restricted (MN) and unrestricted (ED) models. Doing so yields a likelihood ratio $\chi^2$ statistic of 165.48. With one degree of freedom, the difference between the two models is statistically different at the 0.01 level.

Second, we compare how the HRR-level estimates of expected shortage covary with the shortage indicators, including those not included in our estimation but shown in Table 1. There is no guarantee that, even if ED is a better estimator than MN, that it will correlate better with other shortage indicators. The results are given in Table 5. ED does significantly better for indicator 1, the fraction working in facilities trying to hire more ANs. This is almost by construction, as the likelihood function for ED includes this shortage indicator. Note that the correlation coefficient is not very large, at between 0.1 and 0.2. This is suggests that there is additional information coming from the shortage indicator, but that as expected the fundamentals of the labor demand and supply equations can’t be ignored and provide a lot more information. Further, that there was about a 50% increase in the correlation coefficient going from MN to ED, but that the end correlation coefficient of ED is still below 0.2 suggests that the additional information is valuable, but that it doesn’t eliminate all of the other information contained in the demand and supply functions.

The correlation coefficients for the second and third shortage indicators are much smaller, and the two methods have approximately the same correlation. The fourth indicator is much more highly correlated with the expected excess demand, and here we have somewhat higher correlation for MN.

\(^{10}\) As a technical note, we still need to add on to the likelihood the estimation of the mean and standard deviation of the shortage indicator to make the two comparable.
Table 5: Correlation Coefficients of Estimated Excess Demand and Shortage Indicators

<table>
<thead>
<tr>
<th></th>
<th>MN</th>
<th>ED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Fraction working in facilities trying to hire more ANs</td>
<td>0.13</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(-0.04, 0.19)</td>
<td>(-0.08, 0.21)</td>
</tr>
<tr>
<td>2. Fraction working in facilities that could handle more ANs</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(-0.11, 0.12)</td>
<td>(-0.12, 0.12)</td>
</tr>
<tr>
<td>3. Fraction that increased working hours in past 3 years</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(-0.10, 0.12)</td>
<td>(-0.10, 0.11)</td>
</tr>
<tr>
<td>4. Negative of Fraction that would increase hours for increased pay</td>
<td>0.46</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(0.28, 0.59)</td>
<td>(0.29, 0.59)</td>
</tr>
</tbody>
</table>

The final post-estimation test we do is to estimate MN and ED only on 2007 data, and then use that model to predict what the labor demand and labor supply will be given observables we see in 2013, and hence what the observed labor quantity (as the minimum of the two) is compared to the actual 2013 observed quantity. If ED estimates the supply and demand functions better, than we would expect better predictions. We estimate the average absolute bias as well as the Mean Square Error of the two predictions. Table 6 presents these results. ED has slightly better mean absolute bias as well as MSE, suggesting it does a better job of out of sample prediction and has better captured the true demand and supply functions. However, the gains are small.

Table 6: Comparisons of out of sample predictions

<table>
<thead>
<tr>
<th></th>
<th>MN</th>
<th>ED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Absolute Bias</td>
<td>95.14</td>
<td>94.05</td>
</tr>
<tr>
<td>MSE</td>
<td>15,929</td>
<td>15,718</td>
</tr>
</tbody>
</table>

5. Discussion

In this paper, we develop a new disequilibrium estimation technique that uses shortage indicators as sources of additional information for shedding light on excess demand in labor markets. The extended model has an intuitive explanation, wherein markets with higher (lower)-than-expected values of the shortage indicator put more weight on estimating observed quantity as labor supply (demand), and adjust the expected labor demand (supply) upwards and labor supply (demand) downwards.
We estimate the model on the labor market for anesthesiologists, and find changes in the estimated parameters. The expanded model is statistically different from the base model using a likelihood ratio test, and show that it also has more predictive power. There are also interesting by-products to our new approach, including estimated information about the shortage indicator such as its quantitative relationship with changes in shortage or surplus per capita, as well as what the equilibrium level of the shortage indicator is. This additional information may be useful in many settings when analyzing labor markets for disequilibrium.

There are potential extensions of this model not included in this paper. For example, researchers may alter the model to allow for more than one shortage indicator, either as a vector of indicators or as an indexed scalar, with the weights used in the indexation also needing estimation. Another extension is to leverage the panel nature of our data, and allow for market level fixed effects. There are several complications and assumptions that would need to be made in the process, as with large N and small T there are very few observations for which the fixed effects could be estimated.

6. Conclusion

Understanding and estimating potential disequilibrium in labor markets continues to be important. This may be especially true in health labor markets, which may be more prone to disequilibrium given the long training time and the rules and institutes surrounding health care delivery. There are important implications of shortages and surpluses for access to care and the effects from health industry consolidation. In this paper, we develop a new methodology to estimate shortage in specialized labor markets by using auxiliary information regarding probabilities of shortage and surplus to inform estimation of labor supply and labor demand. We apply this methodology to the labor market for anesthesiologists.

This model would be useful in examining labor markets for other specializations. The gains from the models of Maddala and Nelson (1974), and explained by Gourieroux (2000) are important, in terms of additional information as well as some accuracy gains.

This paper contributes to the literature where disequilibrium is likely in labor markets, but it is challenging to estimate the extent of shortage and how it varies across space and time. Future research may develop new extensions that estimate multiple shortage indicators or account for fixed effects.
References


Appendix

A1. MN model additional parameters derivations

After estimating the parameters, we can estimate the following additional functions of these parameters as follows:

The probability of a market having excess demand (shortage) is given by

$$\pi^{ED} = \Pr(Q^D > q) = 1 - \Phi\left(\frac{q - X^D \beta^D}{\sigma_{\varepsilon D}}\right)$$

The expected quantity demanded is given by

$$E[Q^D] = E[Q^D | Q^D > Q^S] \Pr(Q^D > Q^S) + E[Q^D | Q^D < Q^S] \Pr(Q^D < Q^S)$$

$$= E[X^D \beta^D + \varepsilon^D | X^D \beta^D + \varepsilon^D > q] \pi^{ED} + q(1 - \pi^{ED})$$

$$= \pi^{ED} \left( X^D \beta^D + \frac{\sigma_{\varepsilon D} \Phi\left(\frac{q - X^D \beta^D}{\sigma_{\varepsilon D}}\right)}{1 - \Phi\left(\frac{q - X^D \beta^D}{\sigma_{\varepsilon D}}\right)} \right) + (1 - \pi^{ED})q$$

We similarly can calculate for expected quantity supplied (using the same equations, substitute $S$ for $D$, and estimate the expected shortage in a market $E[Q^S_{mt} - Q^D_{mt}]$. This we can aggregate up to multiple markets by summing over all of the markets.

A2. ED model additional parameters derivations

Here we derive the elements of Equations 22 and 23.

$$\sigma_A^2 = \gamma_1^2 (\sigma_{\varepsilon D}^2 + \sigma_{\varepsilon S}^2)/p^2 + \sigma_\gamma^2$$

$$E[A|q, X^D, X^S] = \gamma_0 + \gamma_1 (X^D \beta^D - X^S \beta^S)/p$$
\[ \sigma_{Q^D,A} = \text{Cov}(Q^D, A|q, x^D, x^S). \] Using the law of total covariance (conditioning on \( \varepsilon_s \))

\[
\text{Cov}(Q^D, A|q, x^D, x^S) = E[\text{Cov}(Q^D, A|q, x^D, x^S)] + \text{Cov}(E[Q^D|q, x^D, x^S], E[A|q, x^D, x^S]|q, x^D, x^S)
\]

The first term becomes

\[
E[\text{Cov}(Q^D, A|q, x^D, x^S)|Q^D > Q^S] \Pr(Q^D > Q^S) + E[\text{Cov}(Q^D, A|q, x^D, x^S)|Q^D < Q^S] \Pr(Q^D < Q^S)
\]

\[
= E[\text{Cov}(X^D \beta^D + \varepsilon^D, q_0 + \gamma_1 (X^D \beta^D + \varepsilon^D - X^S \beta^S - \varepsilon^S)/p|q, x^D, x^S)|Q^D > Q^S] \pi^\text{ED}
\]

\[
= E[\text{Cov}(q, q_0 + \gamma_1 (X^D \beta^D + \varepsilon^D - X^S \beta^S - \varepsilon^S)/p|q, x^D, x^S)|Q^D < Q^S](1 - \pi^\text{ED})
\]

\[
= \gamma_1^2 \sigma^2 \pi^\text{ED}/p^2
\]

The second term is equal to zero. Substituting in, we have

\[
E[Q^D|q, x^D, x^S, A] =
\]

\[
E[Q^D|q, x^D, x^S] + \frac{\gamma_1^2 \sigma^2 \pi^\text{ED}/p^2}{\gamma_1^2 (\sigma^2 \sigma^2 / p^2 + \sigma^2 \sigma^2 / p^2)} (a - (\gamma_0 + \gamma_1 (X^D \beta^D - X^S \beta^S)/p))
\]

Following similar methodology, we have

\[
E[Q^S|q, x^D, x^S, A] =
\]

\[
E[Q^S|q, x^D, x^S] - \frac{\gamma_1^2 \sigma^2 \pi^\text{ED}/p^2}{\gamma_1^2 (\sigma^2 \sigma^2 / p^2 + \sigma^2 \sigma^2 / p^2)} (a - (\gamma_0 + \gamma_1 (X^D \beta^D - X^S \beta^S)/p))
\]

The expectation for when q is not observed yet is given by

\[
E[Q^D|X^D] = X^D \beta^D \quad E[Q^S|X^S] = X^S \beta^S
\]

If shortage indicators are included, there is additional information:

\[
E[Q^D|X^D, A] = X^D \beta^D + \phi \left( \frac{X^D \beta^D - X^S \beta^S}{\sigma^2 + \sigma^2} \right) \frac{\gamma_1^2 \sigma^2 \pi^\text{ED}/p^2}{\gamma_1^2 (\sigma^2 \sigma^2 / p^2 + \sigma^2 \sigma^2 / p^2)} (a - (\gamma_0 + \gamma_1 (X^D \beta^D - X^S \beta^S)/p))
\]

\[
- (\gamma_0 + \gamma_1 (X^D \beta^D - X^S \beta^S)/p))
\]