M.S. COMPREHENSIVE EXAMINATION
June 2023

You have three hours, after a 20 minute reading period. You do not need to use the whole time period. You need to answer all three questions, which will be weighted equally. 

Watch the time carefully. The logic used to answer each question is important, so be sure to specify your reasoning with full sentences. You may support your answers using diagrams or mathematical derivations where appropriate. If you use graphs, make sure that they are large enough. We are expecting precise and concise answers. Also make sure your writing is legible: if we cannot read it, your answer will be assumed wrong.
I. Equivalent variation and CES utility (5 points)

In ARE 204A, we argued at length that utility being ordinal, not cardinal, differences in indirect utility between two states of the world are meaningless except for their sign. Instead, economists use EV (or CV) to determine the welfare change between the states. In this problem, you will show that if we assume constant-elasticity-of-substitution (CES) preferences and normalize initial prices and the CES utility function appropriately, differences in utility are actually equal to the equivalent variation between the states.

We thus consider two states $i = 0, 1$, each described by a price-income vector $s^i = (p^i, Y^i)$, where $p^i$ is an $L$-vector of prices in state $i$, and $Y^i$ denotes the income of the consumer in state $i$. We further assume that the consumer’s preferences are representable by the following utility function defined over the $L$-vector of goods $x = (x_1, \ldots, x_L)$:

$$U(x) = \left( \sum_{l=1}^{L} \beta_l^{\frac{1}{\sigma}x_l^{1-\sigma}} \right)^{\frac{\sigma}{1-\sigma}}, \quad \text{(I. 1)}$$

where $\beta_l > 0$ for $l = 1, \ldots, L$.

1. Interpret the coefficient $\sigma$ in words. What is its name? What is its expected range?

**Solution:** It is the elasticity of substitution across goods, and it represents the percentage change in the ratio of demands $\frac{x_l}{x_j}$ caused by a 1% change in the ratio of corresponding prices $\frac{p_j}{p_l}$ (note the indices being reversed between the ratio of quantities and the ratio of prices), holding utility constant. Its range is $(0, 1) \cup (1, +\infty)$ as it is a positive number, and utility would not be defined for $\sigma = 1$.

2. Set up the expenditure minimization problem for the preferences represented in Equation (I. 1), denoting by $\bar{u}$ the target utility level.

**Solution:** The expenditure minimization problem for these preferences is

$$\min_{x \geq 0} \sum_{l=1}^{L} p_l x_l \quad \text{subject to} \quad U(x) \geq \bar{u}.$$

3. Using the Lagrange approach, show that the Hicksian (compensated) demand for good $l$ is

$$\bar{x}_l(p, \bar{u}) = \bar{u} \left( \sum_{k=1}^{L} \beta_k p_k^{1-\sigma} \right)^{\frac{\sigma}{1-\sigma}} \frac{\beta_l p_l^{1-\sigma}}{\beta_l p_l^{1-\sigma}}.$$
Solution: See class notes.

4. Deduce the consumer’s expenditure function, and check that it is increasing in prices and in utility.

**Solution:** By plugging the compensated demands into the expenditure \( \sum p_i x_i \), we obtain the expenditure function:

\[
e(p, \bar{u}) = \bar{u} \left( \sum_{i=1}^{L} \beta_i p_i^{1-\sigma} \right)^{-1/\sigma}.
\]

Clearly, this function is increasing in \( p_i \) because \( \beta_i > 0 \), and it is increasing linearly in \( \bar{u} \).

5. Using duality, deduce the consumer’s indirect utility function, and show that it can be written as

\[
v(p, Y) = \frac{Y}{\mathcal{P}(p)},
\]

where \( \mathcal{P}(p) \) is a price index that you will make explicit.

**Solution:** By duality, \( e(p, \bar{u}) = Y \iff \bar{u} = v(p, Y) \). Therefore, to find the function \( v \) we set \( e \) equal to \( Y \) and solve for \( \bar{u} \). We get \( v(p, Y) = \frac{Y}{\left( \sum_{k=1}^{L} \beta_k p_k^{1-\sigma} \right)^{-1/\sigma}} \). Therefore, \( v(p, Y) = \frac{Y}{\mathcal{P}(p)} \), for

\[
\mathcal{P}(p) = \left( \sum_{k=1}^{L} \beta_k p_k^{1-\sigma} \right)^{1/\sigma}.
\]

6. Check that \( \mathcal{P}(p) \) satisfies homogeneity of degree 1 and that it is increasing in \( p_i, i = 1, \ldots, L. \)

**Solution:** The price index is HD-1 because if we multiply all prices by \( \lambda > 0 \), we can extract \( \lambda^{1-\sigma} \) from the summation and therefore \( \lambda \) from the entire expression, so that the index is also multiplied by \( \lambda \). That is, \( \mathcal{P}(\lambda p) = \lambda \mathcal{P}(p) \). The price index is increasing in prices because \( \beta_i > 0 \) for all \( i = 1, \ldots, L \).

We define the equivalent variation between state \( s^0 \) and state \( s^1 \) and denote EV \((s^0, s^1)\), the amount of money that, if added to the consumer’s income under state \( s^0 = (p^0, Y^0) \), would bring her to the utility level reached under state \( s^1 = (p^1, Y^1) \).

7. Using the indirect utility function \( v \), write the equation that must be satisfied by EV.

**Solution:** Given its definition, EV must satisfy \( v(p^0, Y^0 + EV) = v(p^1, Y^1) \).
8. Using your answers to parts 5 and 7, compute $EV\left(s^0, s^1\right)$.

**Solution:** We must have $\frac{Y^{0+EV}}{P(p^0)} = \frac{Y^1}{P(p^1)}$, which implies that

$$EV = P(p^0) \times \frac{Y^1}{P(p^1)} - Y^0.$$ 

9. Now show that if the units of measurement are set such that $p^0 \equiv 1$ (where 1 denotes an $L$-vector of ones), and if $\sum_{i=1}^L \beta_i = 1$, then

$$EV\left(s^0, s^1\right) = v(s^1) - v(s^0).$$

**Solution:** If $p^0 \equiv 1$ and $\sum_{i=1}^L \beta_i = 1$, then $P(p^0) = P(1) = \left(\sum_{k=1}^L \beta_k\right)^{1/1} = 1$. We then have $v(s^0) = \frac{Y^0}{P(p^0)} = Y^0$, and $EV\left(s^0, s^1\right) = \frac{Y^1}{P(p^1)} - Y^0 = v(s^1) - v(s^0)$.

10. With the same normalizations as in part 9, is it the case that $v(s^1) - v(s^0) = CV\left(s^0, s^1\right)$, where $CV\left(s^0, s^1\right)$ denotes the compensating variation between the two states? If not, how would the normalization need to be changed for the change in indirect utility to represent $CV$?

**Solution:** No, because $EV$ and $CV$ are different, even if they have the same sign. $CV$ would be defined by $v\left(p^1, Y^1 - CV\right) = v\left(p^0, Y^0\right)$, that is, $\frac{Y^1 - CV}{P(p^1)} = \frac{Y^0}{P(p^0)}$, or $CV = Y^1 - P(p^1) \times \frac{Y^0}{P(p^1)}$. For this expression to equal $v(s^1) - v(s^0)$, we would need $P(p^1) = 1$, which would require $\sum \beta_i = 1$ but $p^1 = 1$. This means that the prices in the new, not the initial, state, would need to be normalized to one.
II. Analysis of the gasoline market (5 points)

Suppose that you would like to investigate the willingness to pay for convenience of gasoline station location. There are two small towns (named A and B) located along the same highway 10 miles away from each other. Each town has only one gasoline station. Some residents of a town may choose to buy gasoline in the other town if the price of the fuel there is sufficiently cheaper. We conducted a survey of residents indexed by subscript \( i \) of both towns in a particular week and asked where they bought their gas from in that week \( (Y_i = 0 \text{ for town A, and } Y_i = 1 \text{ for town B}) \); their weekly wage \( w_i \); distance in miles from their residence to gas stations in towns A and B \( (d_i^A \text{ and } d_i^B) \); and price for a full tank of gasoline at each gas station, \( p_i^A \) and \( p_i^B \) (assume it is always 10 gallons of gasoline). Suppose that the station in town A provides discounts to clients with low income status, which we also observe \( (D_i = 1 \text{ for people with low-income and 0 otherwise}) \). The station in town B does not provide such a discount.

1. The first model specification that you try to estimate is the following:

\[
P\{Y_i = 1\} = \beta_0 + \beta_1 D_i + \beta_2 (p_i^B - p_i^A) + \beta_3 (d_i^B - d_i^A).
\] (II. 1)

(a) Provide an interpretation for coefficients \( \beta_1, \beta_2 \) and \( \beta_3 \). What are the expected signs of these coefficients? Explain.

**Solution:** \( \beta_1 \) captures effect of qualifying for low-income discount on the choice of the gasoline station in Town B over A. Low-income customers are more likely to choose Town A over B because of the discount, so \( \beta_1 \) is negative.

\( \beta_2 \) captures effect of extra dollar of price difference on probability of choosing station B. This coefficient is negative.

\( \beta_3 \) captures effect of extra mile of driving difference on probability of choosing station B. This coefficient is negative too.

(b) Does the income level affect the willingness to pay for reducing the distance traveled by an extra mile? Would your answer change if you use a Probit model instead? Explain.

**Solution:** In the linear model, the marginal effects do not depend on the regressors. So income level does not change the willingness to pay. In the Probit model, the income level will, in contrast, affect the willingness to pay.

(c) Write down a linear regression model corresponding to Equation (II. 1). Is it possible that the corresponding regression residual is homoscedastic?

**Solution:**

\[
Y_i = \beta_0 + \beta_1 D_i + \beta_2 (p_i^B - p_i^A) + \beta_3 (d_i^B - d_i^A) + U_i
\]

where \( E(U_i|w_i,p_i^B,p_i^A,d_i^B,d_i^A) = 0 \). Residual in the linear probability model is always heteroscedastic.

(d) Suppose that we only have observations for one week, so there is no variation in \( p_i^B \) and \( p_i^A \). Is parameter \( \beta_2 \) identified? Explain.
Solution: In the absence of the variation in the regressor, we cannot identify $\beta_2$. The assumption of no multicollinearity is violated (regressor $p_i^B - p_i^A$ is indistinguishable from the constant term, which causes their perfect collinearity).

(e) Suppose that we know that the size of the low-income discount is such that $\beta_2 = \beta_1$ (i.e. the discount is equal to $1$ per full gasoline tank). Does this solve the identification issue? If so, propose an estimator of the willingness to pay to reduce a trip by one mile.

Solution: If we impose this restriction, the equation (II. 1) becomes

$$P\{Y_i = 1\} = \beta_0 + \beta_1(D_i + p_i^B - p_i^A) + \beta_3(d_i^B - d_i^A).$$

As long as low income status $D_i$ has some variation, $\beta_1$ is identified. The willingness to pay can be computed as ratio of the OLS estimators $\hat{\beta}_3 / \hat{\beta}_1$.

2. Now we are interested in the estimation of the average gasoline demand elasticity at a gasoline station in the entire state, not only in towns A and B. Suppose that we can observe weekly prices $p_{it}$ and sales $q_{it}$ at all gasoline stations $i$ across the state for several weeks $t$ ($i = 1, \ldots, N$; $t = 1, \ldots, T$).

(a) Consider a simple linear regression model

$$\log q_{it} = \mu + \gamma \log p_{it} + \epsilon_{it}.$$ (II. 2)

Under what conditions can we interpret parameter $\gamma$ as an elasticity of demand? Are these conditions realistic? Explain.

Solution: $\gamma$ has interpretation of elasticity of demand only if (i) there are no individual effects $\mu_i$ that are correlated with $\log p_{it}$, i.e. no gasoline station level demand shocks that shift the local price; (ii) gasoline prices are set exogenously by the gasoline station owners (or government) regardless of the demand. These assumptions do not reflect a typical market conditions in most cases.

(b) One can account for heterogeneity of local demand by including individual effects $\mu_i$ instead of the constant. Explain the difference in underlying assumptions required for the FE and RE estimators of this model to be consistent for the elasticity of demand. Which estimator would you prefer? (For this question assume the individual effects fully capture all the demand shocks.)

Solution: RE estimator requires $\mu_i$ to be uncorrelated with $\log p_{it}$, while FE allows for correlation between $\mu_i$ and $\log p_{it}$. As usual for observational data, we prefer FE over RE because of this extra robustness.

(c) After estimating the model with the individual effects using your preferred estimator, you found out that $\hat{\gamma}$ is too close to zero. As you learned in the course, it may be a result of endogenous regressors bias (indeed, prices are endogenously determined in the market equilibrium). You decide to use one of the methods for causal inference that we learned
in the class to address this issue. Suppose that at period \( t = t_1 > 1 \) the authorities in the Northern counties of the state introduced an extra excise tax that resulted in a 1% price increase of their gas price, but the Southern counties did not do it. Suppose that you know the locations of each gasoline station. Propose a diff-in-diff regression design that would allow you to estimate the elasticity of gasoline demand. You need to write down the corresponding panel regression equation. Explain all the variables in the equation.

**Solution:**

\[
\log q_{it} = \mu_i + \gamma D_{it} + \lambda_t + \epsilon_{it}.
\]

Here \( D_{it} \) is the treatment dummy variable that is equal to 1 for the northern counties starting \( t = t_1 \). Otherwise it is equal to 0. \( \lambda_t \) are parallel time trends. \( \epsilon_{it} \) is regression residual that satisfies \( E(\epsilon_{it}|\mu_i, D_{it}, \ldots, D_{iT}, \lambda_1, \ldots, \lambda_T) = 0 \). Since the treatment corresponds to 1% increase in the price driven by the supply side, \( \gamma \) has interpretation of elasticity of demand.

Note that we do not need to include the prices as controls for this design to work. However, one can include them for extra robustness if \( T > 2 \) (otherwise, they would be multi-collinear with the dummy variable).

(d) How can you control for station-specific gradual changes in demand for gasoline? Explain.

**Solution:** One can include station-specific linear trends \( \eta_{it} \) into the model. They would capture demand changes that drive difference in sales that were ongoing before the treatment. That would further reduce the endogeneity bias, if any.

(e) For standard errors you decided to use the cluster-robust standard errors at the gasoline station level. What are the extra robustness properties that these errors have compared to the regular heteroscedasticity-robust errors?

**Solution:** Gasoline sales at each station may be correlated over time. Cluster robust errors account for arbitrary auto-correlation of the sales in addition to the heteroscedasticity.
III. International Graduate Students and Innovation in the US (5 points)

Seemingly small investments in innovation today can have significant effects on economic growth in the long run. Policies that shape and direct innovation incentives can therefore be particularly consequential. Chellaraj, Maskus, and Mattoo (2008) empirically test the effect of international graduate students on innovation—measured by patent applications and grants—in the US, which became a matter of debate in the wake of the attacks of 11 September 2001 as part of a broader discussion of US education visas.

1. Describe the limitations or disadvantages of using patent statistics as a measure of innovation. Explain why so many empirical analyses in economics nonetheless use patents as a proxy for innovation.

**Solution:** Patent statistics are an incomplete measure of innovation because relatively few industries rely heavily on patents and even these industries do not exclusively rely on patents. For these two reasons, much of the innovation that occurs in even high income economies is therefore not captured by patents. Patents also focus on invention rather than innovation per se, which requires use in practice. Patents are used nonetheless because they are readily available and include rich details (e.g., technological classifications) for empirical analysis.

2. The authors use a production function approach to address their research question as described in this passage:

To estimate the contribution of skilled immigrants and foreign graduate students to US innovation, we modify the “national ideas production function” that is widely used in innovation studies (Porter and Stern, 2000; Stern et al., 2002). This may be written in general form as

\[ \dot{\Delta}_t = \delta H_{A,t} A_t^\lambda. \]  

(1)

Thus, “the rate of new ideas produced depends on both the allocation of resources to the R&D sector \((H_{A,t})\), the stock of ideas already in existence \((A_t)\)” and three parameters. Define and interpret the parameters \(\delta\), \(\lambda\), and \(\phi\) as completely as possible in the context of this problem.

**Solution:** \(\delta\) is a general factor-augmenting parameter that affects how efficiently both inputs translate into new ideas. \(\lambda\) indicates the productivity of the \(H\) resources in producing new ideas and is interpreted as an elasticity. \(\phi\) captures the ability of the stock of ideas to generate new ideas. If \(\phi > 0\), the stock of ideas fuels the generation of yet additional ideas (a ‘standing on shoulders’ effect). If \(\phi > 0\), the number of potential ideas may be fixed so that past discovery makes future discovery less rather than more likely.

3. The authors modify this national ideas production function by decomposing \(H_{A,t}\) as follows:
\[ H_{A_t}^t = H_{F_t}^{tF} H_{G_t}^{tG} H_{R_t}^{tR} H_{S_t}^{tS} H_{R_t}^{tR}. \]  

Here, \( H_F \) is the flow (enrollments) of international graduate students, \( H_G \) is the flow (enrollments) of total graduate students, \( H_I \) is the number of skilled immigrants in the country, \( H_R \) is the number of total PhD engineers and scientists, and \( H_S \) is expenditure on R&D. Note that there is some overlap between skilled immigrants and engineers and scientists, but it is not possible with available data to distinguish sharply between these factors.

When the authors later use this modified production function as the basis for an empirical specification, is it a problem that these elements of \( H_A \) are not measured in the same units? Explain why or why not.

**Solution:** It is not a problem because the coefficients in this functional form are interpreted as elasticities, which are unit-less.

4. The econometric specification proposed by the authors is:

\[
\begin{align*}
\text{IPA}_{t+5} &= \alpha_t + \lambda_{F1} \text{FORTGR}_{t} + \lambda_{I1} \text{IMCUM}_{t} + \lambda_{S1} \text{SK}_{t} + \lambda_{R1} \text{RD}_{t} \\
&+ \phi_{t1} \text{TOTPATSTOCK}_{t} + \delta_{B1} \text{BD}_{t} + \theta_{t1} \text{TIME}_{t} + \eta_{t1} \\
\text{IPG}_{t+7} &= \alpha_t + \lambda_{F2} \text{FORTGR}_{t} + \lambda_{I2} \text{IMCUM}_{t} + \lambda_{S2} \text{SK}_{t} + \lambda_{R2} \text{RD}_{t} \\
&+ \phi_{t2} \text{TOTPATSTOCK}_{t} + \delta_{B2} \text{BD}_{t} + \theta_{t2} \text{TIME}_{t} + \eta_{t2}.
\end{align*}
\]

where all the variables except the BD dummy variable and TIME are in logs and the variables are defined in the notes of the correlation matrix shown below as Table 1.

The coefficient of interest in this analysis is the one on the variable \( \text{FORTGR}_t \). What does the coefficient \( \lambda_{F1} \) represent? (i.e., how would you interpret it?)

**Solution:** \( \lambda_{F1} \) represents the elasticity of patent applications with respect to foreign graduate students as a proportion of total graduate students in the US. That is, it is interpreted as the percent change in patent applications for a 1% change in the proportion of foreign graduate students five years earlier.

5. The data used to estimate this regression are aggregate (i.e., across all industries), annual data from the US from 1963 to 2001. The authors assume that \( \delta \) in equation (1) above is a function of time, as represented by the linear TIME variable. If this were your paper, how would you justify this assumption?

**Solution:** This linear time trend captures general changes in the generation of new ideas, in this case represented by patent applications and grants. This could proxy for changes in the way research is conducted or for systematic changes in the efficiency of the research process that are not linked to \( H \) or \( A \). This is a simple way to capture a host of potential
changes across this nearly 40 year period, which included among other things the rise of modern computing.

6. The Bayh-Dohl Act, passed in the US in 1980, allowed universities to apply for and receive patents for inventions produced by their employees (just like companies). The authors use \( \delta_{B1} \) and \( \delta_{B2} \) to denote the coefficients on the associated BD dummy variable because they assume the \( \delta \) parameter in equation (1) is a function of BD (in addition to being a function of linear TIME per the preceding question).

Again, how would you justify this assumption? How does the interpretation of \( \delta_{B1} \) differ from that of \( \lambda_{F1} \)?

**Solution:** Since ‘we’ are interested in the effect of foreign graduate students on patenting, the engagement of university labs in pursuit of patents is important. In this sense, the BD dummy could definitely shape the ‘returns’ to the \( H \) and \( A \) inputs as it denotes a change in the institutional incentives to innovate and use the patent system within universities.

7. Explain why the outcome variable IPA is lagged 5 years relative to all the RHS variables. Given the nature of the data they use, what implicit assumption do the authors make when imposing this 5-year lag?

**Solution:** This lag is meant to adjust for the time required to turn research time and effort into patentable inventions. While 5 years to patent application is somewhat arbitrary, it is loosely based on average patterns in research. The implicit assumption it makes is that this average is relevant across all industries that are represented in the patent statistics. This is clearly not an accurate assumption as some industries require much longer research processes to produce patent applications, but it is necessary because they have no (easy) way of allowing for different lags by industry.

8. Why is a correlation matrix such as the one below useful as part of an econometric exercise? Be specific.

**Solution:** Because multicollinearity is a common problem in regression analysis. Specifically, this problem results in large standard errors because including highly correlated RHS variables results in much of the variation in the variables being ignored by the regressions since it cannot be attributed independently to either of the underlying variables. Ignoring variation in the data can inflate the standard errors, leading to imprecise estimates.

9. The authors present results in Table 2 shown below for four different outcomes: IPA and IPG (as defined above), as well as university patent grants (UIPG) and other patent grants (OIPG). They include the US unemployment rate to control for business cycle effects on innovation.
Assume for a moment that you were one of the esteemed collaborators on this paper. Write a paragraph that interprets and discusses as thoroughly and clearly as you can the results in Table 2 for the variable FORTGR.

**Solution:** Answers should emphasize both statistical significance and economic magnitude of the estimated coefficients, which is surprisingly large (maybe too large). Could compare IPA to IPG, and UIPG to OIPG. Could mention limitations (e.g., sample size).

10. Table 2 reports the Durbin-Watson (DW) statistic, which is like the more general Breusch-Godfrey test. What does the Breusch-Godfrey test check for specifically? Why is such a test relevant in this particular analysis?

**Solution:** These tests check for serial correlation in the residuals of the regression. This is an important test in this case because the authors use exclusively time-series variation to estimate the regression – and serial correlation can introduce confounding spurious relationships.

11. The authors worry that if the key variables in their analysis are non-stationary, their results may be spurious. Explain the nature of this concern in the context of this problem. They find that most of their variables have a unit root. They conduct a cointegration analysis and find that FORTGR has a long-run relationship with patenting (results not reported here). Explain the logic of a cointegration test. Why don’t the authors simply use first-differences to deal with serial correlation?

**Solution:** A cointegration relationship exists when two variables are separately non-stationary, but a linear combination of the two is stationary. In this case, the linear combination is often meaningful as a long-run equilibrium relationship, but only cointegration analysis can pick this up. Standard naive application of first-differencing to remedy the unit root in the two variables makes it impossible to detect this long-run relationship.

12. Based only on what you have learned about the analysis of Chellaraj, Maskus, and Mattoo (2008) through this problem, how reliable do you think the results are as a contribution to the debate on high-skilled immigration and educational visas to the US? Evaluate the strengths and weaknesses of the analysis as you can see them in this problem.

**Solution:** Several possibilities. Among the potential weaknesses worth mentioning are the small sample size, the aggregation across industries with very different patenting/innovation tendencies, and aggregation across states with very different reliance on foreign graduate students and innovation systems. Endogeneity of FORTGR could be an issue, but it is tricky to discuss in much detail.
Table 1. Correlations among Variables

<table>
<thead>
<tr>
<th></th>
<th>IPA</th>
<th>IPG</th>
<th>UIPG</th>
<th>OIPG</th>
<th>FOR*</th>
<th>IM*</th>
<th>SK</th>
<th>RD</th>
<th>URD</th>
<th>ORD</th>
<th>TPS5*</th>
<th>TPS7*</th>
<th>BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPA</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPG</td>
<td>0.83</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UIPG</td>
<td>0.87</td>
<td>0.63</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OIPG</td>
<td>0.81</td>
<td>0.99</td>
<td>0.59</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOR*</td>
<td>0.79</td>
<td>0.69</td>
<td>0.76</td>
<td>0.67</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM*</td>
<td>0.72</td>
<td>0.39</td>
<td>0.88</td>
<td>0.35</td>
<td>0.47</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>0.67</td>
<td>0.79</td>
<td>0.49</td>
<td>0.79</td>
<td>0.34</td>
<td>0.30</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>0.76</td>
<td>0.55</td>
<td>0.84</td>
<td>0.52</td>
<td>0.59</td>
<td>0.71</td>
<td>0.59</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>URD</td>
<td>0.73</td>
<td>0.42</td>
<td>0.94</td>
<td>0.38</td>
<td>0.70</td>
<td>0.80</td>
<td>0.35</td>
<td>0.89</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORD</td>
<td>0.76</td>
<td>0.55</td>
<td>0.83</td>
<td>0.53</td>
<td>0.58</td>
<td>0.70</td>
<td>0.60</td>
<td>0.99</td>
<td>0.88</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPS5*</td>
<td>0.64</td>
<td>0.64</td>
<td>0.37</td>
<td>0.64</td>
<td>0.28</td>
<td>0.33</td>
<td>0.69</td>
<td>0.53</td>
<td>0.24</td>
<td>0.54</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPS7*</td>
<td>0.79</td>
<td>0.82</td>
<td>0.54</td>
<td>0.82</td>
<td>0.53</td>
<td>0.35</td>
<td>0.80</td>
<td>0.64</td>
<td>0.42</td>
<td>0.65</td>
<td>0.93</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>BD</td>
<td>0.22</td>
<td>−0.07</td>
<td>0.59</td>
<td>−0.11</td>
<td>0.39</td>
<td>0.52</td>
<td>−0.17</td>
<td>0.45</td>
<td>0.72</td>
<td>0.44</td>
<td>−0.43</td>
<td>−0.26</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Notes: IPA is patent applications and IPG is patents granted, both as a proportion of labor force. UIPG and OIPG are patents granted to universities and other institutions, respectively, as a proportion of the labor force. FORTGR (FOR* in this table) is foreign graduate students as a proportion of total graduate students. IMCUM (IM* in this table) is the cumulative number of skilled immigrants over a period of six years as a proportion of the labor force. SK is total PhD scientists and engineers as a proportion of labor force. RD is total real R&D expenditures as a proportion of labor force, while URD and ORD refer to university real R&D and other-institution real R&D, each as a proportion of labor force. TOTPATSTOCK is cumulative patents awarded as a proportion of labor force (TPS5 (TPS7) is TOTPATSTOCK lagged 5 (7) years). BD is the dummy variable for the Bayh–Dole Act.
<table>
<thead>
<tr>
<th></th>
<th>IPA (1)</th>
<th>IPA (2)</th>
<th>IPG (3)</th>
<th>IPG (4)</th>
<th>UIPG (5)</th>
<th>UIPG (6)</th>
<th>OIPG (7)</th>
<th>OIPG (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>3.819</td>
<td>−1.363</td>
<td>4.797</td>
<td>−3.459</td>
<td>0.798</td>
<td>−2.543</td>
<td>4.523</td>
<td>−3.585</td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(−0.80)</td>
<td>(1.84)*</td>
<td>(−1.34)</td>
<td>(0.67)</td>
<td>(−3.6)**</td>
<td>(1.77)*</td>
<td>(−1.40)</td>
</tr>
<tr>
<td>FORTGR</td>
<td>0.764</td>
<td>1.004</td>
<td>0.572</td>
<td>0.895</td>
<td>0.672</td>
<td>0.506</td>
<td>0.554</td>
<td>0.884</td>
</tr>
<tr>
<td>IMCUM</td>
<td>0.114</td>
<td>0.040</td>
<td>0.048</td>
<td>−0.086</td>
<td>0.134</td>
<td>0.055</td>
<td>0.048</td>
<td>−0.085</td>
</tr>
<tr>
<td></td>
<td>(2.01)**</td>
<td>(0.46)</td>
<td>(0.75)</td>
<td>(−0.79)</td>
<td>(1.93)*</td>
<td>(0.79)</td>
<td>(0.76)</td>
<td>(−0.79)</td>
</tr>
<tr>
<td>SK</td>
<td>1.222</td>
<td>1.820</td>
<td>1.286</td>
<td>1.805</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.32)**</td>
<td>(6.04)**</td>
<td>(3.00)**</td>
<td>(6.06)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>0.197</td>
<td>0.489</td>
<td>0.091</td>
<td>0.532</td>
<td>−0.052</td>
<td>0.240</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.10)</td>
<td>(2.95)**</td>
<td>(0.48)</td>
<td>(2.57)**</td>
<td>(−0.18)</td>
<td>(0.88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTPATSTOCK</td>
<td>0.360</td>
<td>1.020</td>
<td>0.039</td>
<td>0.988</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.25)</td>
<td>(4.00)**</td>
<td>(0.26)</td>
<td>(3.19)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPATSTOCK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.331</td>
<td>0.794</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.68)*</td>
<td>(7.98)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPATSTOCK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.016</td>
<td>0.965</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(3.12)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BD</td>
<td>−0.041</td>
<td>0.116</td>
<td>0.018</td>
<td>0.245</td>
<td>0.090</td>
<td>0.234</td>
<td>0.016</td>
<td>0.241</td>
</tr>
<tr>
<td></td>
<td>(−0.53)</td>
<td>(2.75)**</td>
<td>(0.38)</td>
<td>(2.86)**</td>
<td>(1.21)</td>
<td>(2.14)**</td>
<td>(0.34)</td>
<td>(2.85)**</td>
</tr>
<tr>
<td>TIME</td>
<td>−0.008</td>
<td>−0.013</td>
<td>−0.010</td>
<td>−0.015</td>
<td>0.040</td>
<td>0.006</td>
<td>−0.010</td>
<td>−0.015</td>
</tr>
<tr>
<td></td>
<td>(−1.67)*</td>
<td>(−2.47)**</td>
<td>(−1.64)</td>
<td>(−1.69)*</td>
<td>(1.77)*</td>
<td>(0.33)</td>
<td>(−1.74)*</td>
<td>(−1.71)*</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>0.007</td>
<td>−0.022</td>
<td>−0.014</td>
<td>−0.040</td>
<td>0.056</td>
<td>0.060</td>
<td>−0.016</td>
<td>−0.044</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(−0.28)</td>
<td>(−0.19)</td>
<td>(−0.39)</td>
<td>(0.38)</td>
<td>(0.34)</td>
<td>(−0.24)</td>
<td>(−0.42)</td>
</tr>
<tr>
<td>Observations</td>
<td>39</td>
<td>39</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.94</td>
<td>0.90</td>
<td>0.87</td>
<td>0.78</td>
<td>0.99</td>
<td>0.99</td>
<td>0.87</td>
<td>0.77</td>
</tr>
<tr>
<td>DW</td>
<td>1.06</td>
<td>1.01</td>
<td>2.33</td>
<td>1.35</td>
<td>2.19</td>
<td>1.94</td>
<td>2.34</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Notes: Variables in the IPA equations are lagged five years (sample 1963–2001), while those in the IPG, UIPG, and OIPG equations are lagged seven years (sample 1965–2001). Figures in parentheses are t-statistics based on Newey–West estimated standard errors and are marked as significantly different from zero at the 1% (***) , 5% (** ) and 10% (*) levels.
References