A NON-PARAMETRIC STATISTICAL TEST FOR J.R. HICKS’ INDUCED INNOVATION HYPOTHESIS

Ilya Bolotov – Tomáš Evan

Abstract

This paper attempts to construct a simple non-parametric statistical test, a combination of a) Student’s t-test, b) Wald’s F-test and c) calculus for the induced innovation hypothesis published by J. R. Hicks in 1932: “a change in the relative prices of the factors of production is itself a spur to invention, and to invention of a particular kind directed to economizing the use of a factor which has become relatively expensive”. The test is performed on parameters estimated from a dynamicized comparatively static economic model constructed by the authors from Lancaster’s characteristics consumer behaviour theory and the neoclassical production function with a Gorman-style representative consumer and firm. Estimations are performed on a panel dataset, which comprises 154 countries for the years 1980–2015 (5544 rows) with the help of the General method of Moments (GMM). The paper contributes to the overall economic and historical causes of innovations in economies.

Key words: non-parametric tests, induced innovation, GMM

JEL Code: C23, C51, O30

Introduction

Numerous studies in the history of economics, including the ones in the recent years, such as (Savona & Steinmueller, 2013), (Fabre, 2014), or (Milyaeva & Fedorkevich, 2015), have continuously shown that innovations benefit companies, industries and economies in terms of increasing competitiveness, economic growth and development. There is however little consensus on what are the main causes of innovation. Sir J.R. Hicks, (Hicks, 1964, p. 124), has stated that “a change in the relative prices of the factors of production is itself a spur to invention, and to invention of a particular kind – directed to economizing the use of a factor which has become relatively expensive”, which later became the foundation of Hicks’ widely discussed Induced Innovation theory. Not even the critics can deny that the theory has been used and in quite a few cases proven empirically. Attempted from the start and followed by ground building paper of William Fellner “Empirical Support for the Theory of Induced Innovation”, (Fellner, 1971), there have been several fields in which its endorsement and application was found. The line of research confirming empirically the main hypothesis...
included in the theory was originally centred on high wages spurring labour-saving innovation, later agricultural development. More recently the emphasis has shifted towards energy prices and induced innovation in energy-saving technologies. Newell, Jaffe and Stavins, (Newell, Jaffe, & Stavins, 1999), i.a. found the rate of overall innovation independent of energy prices and regulation, however, the direction of innovation was responsive to energy price changes for several products tested by the authors. (Popp, 2002), using patent citations as measure of supply of knowledge found that both energy prices and the quality of existing knowledge have significantly strong positive effects on innovation. Also using patent counts and citation data, (Jang, Du, & al., 2013), confirm that demand and supply factors – including knowledge stocks and crude-oil price – have positive and statistically significant effects on technological biofuel innovations in the United States of America. There is also relatively large number of other correlates to innovation such as inward foreign direct investment (FDI), outward FDI, imports, state guarantees and incentives among many other, as stressed by (Lin & Lin, 2008).

The goal of this paper is to attempt to falsify, in the sense of Popper, Hutchison, Machlup, Lakatos and later Blaug, the Hicks hypothesis by means of a novel statistical test derived from neoclassical and Neokeyesian (orthodox) modelling and corresponding econometric research based on data from official sources in accordance with the instrumentalist principles, formulated by Friedman, Machlup and Musgrave. To do so, we formulate a hypothesis alternative to the Hicks’ theory: “innovation is spurred by an increase in the relative price of one factor of production compensated by a decrease in relative price of another factor of production.” This statement may be used as “sophisticated” falsification of the Induced Innovation Theory\(^1\), for the test of which we employ econometric methodology.

1 **Formulating a non-parametric econometric test**

To derive a non-parametric econometric test for the Hicks’ and alternative hypotheses, we construct an orthodox, i.e. comparative static, model based on Lancaster’s characteristics consumer behaviour theory, (Lancaster, 1966)\(^2\), and on the neoclassical production theory, (Cobb & Douglas, 1928), (Solow, 1956), (Berndt & Christensen, 1973) etc. with Gorman-style

---

\(^{1}\) Popper, Hutchison, Lakatos, and Blaug suggest that “sophisticated” falsification is the only criterion of validity of on any scientific theory, which should be rejected if being systematically refuted.

\(^{2}\) The “characteristics” consumer behaviour theory, published by Lancaster in 1966, is part of ordinal utility analysis in which quantities of goods/services are replaced with a limited set of characteristics (attributes) so that each good/service is a combination from the set. McFadden, Berry and Pakes offered empirical tests for Lancaster’s model.
representative consumer and firm. The following simplifications, in accordance with the
instrumentalist “results justify false assumptions” create the model’s core:

- Two characteristics of goods and services – non-innovative and innovative attributes;
- Four factors of production – capital-intensive (qualified) labour \((N)\), technology
  complementary to labour \((T, T = k \cdot N)\), natural resources \((A)\), and capital \((K)\), a linear
  combination of all inputs, \(K = \xi(N, T, A) = \xi(N, A)^3\);
- End consumer - producer relationships;
- Homogeneity of end consumers and producers;
- Entailed strong(er) competition – monopolistic competition or oligopoly on the supply
  side and perfect or monopsonistic competition on the demand side;
- Absence of extreme solutions and special consumer/company types or relationships as
  specified in (Němečková, 2013), (Machek & Hnilica, 2015) etc.

Let us suppose a model market or mixed economy of any size\(^4\) with at least minimum
access to natural resource and capital\(^5\) where \textit{ceteris paribus} (or \textit{ceteris absentibus}) any good
and / or service \(Y_i\) is a divisible (or at least a mostly divisible) combination of non-innovative
and innovative characteristics, the “old attributes” and the “new attributes”, depending on four
inputs\(^6\): quantities of qualified labour, \(z_N\), technology complementary to labour (with a fixed
\(N/T\) relationship)\(^7\), \(z_T\), natural resources, \(z_A\), and capital, \(z_K\) (hereafter omitted from graphs and
formal representations of the model because of the exact multicollinearity with the other \(z_{k,k \neq K}\)
since \(K = \xi(N, A)\)), as well as on the relative prices of \(z_{k,k \neq RP}\), \(\{z_{RP}\} = \{w_N/w_K, w_N/w_A, w_N/w_T, w_K/w_A, w_K/w_T\}\)\(^8\), where \(w_{k,k \neq RP}\), are the prices of the inputs.

The formal econometric representation of our model for market equilibrium\(^9\) in a
\textit{multiplicative form}, subsequently simplified with regard to constant (relatively abundant)
natural resources, \(z_A\), and the complimentary relationship between technology and qualified

---

\(^3\) In our opinion, such specification is consistent with the neoclassical theory of economics (three factors of
production, \(A, N\) and \(K\)), and reflects the structural changes in economies happening through accumulation of
human capital and technology.

\(^4\) Economy with functioning markets regardless of ownership structure, e.g. U.S., U.K., Germany, France, Japan
or China, Russia, Singapore etc.

\(^5\) Trade and capital restrictions may be present but are not prohibitive in nature, which ensures their relative
abundancy at the country level but not at the global level.

\(^6\) Since productive factor \(A\) is considered to be abundant, \(z_A\) is left out.

\(^7\) A requirement of education to employ certain pieces of technology can serve as an example.

\(^8\) Price of labour is traditionally selected as the most important one.

\(^9\) Due to the simplified nature of our model, based on Gorman-style representative economic agents and \textit{ceteris
paribus} or \textit{ceteris absentibus} assumptions, the terms “general (multimarket) equilibrium” and “market
equilibrium” are employed as synonyms.
The nature of our model, applied to panel data, requires the use of the generalized method of moments (GMM), which has the following level and first differences versions, depending on data’s time series characteristics (unit roots), after a logarithm transformation of equation (1):

\[ E(V - \Gamma \log X | \log X, \log I) = 0, \]

\[ E(\Delta V - \Gamma \Delta \log X | \Delta \log X, \Delta \log I) = 0 \]

(3)

where \( \sum_i \psi_i Y_i^* \) is the equilibrium quantity of innovations in an economy, \( \Pi \) are end prices, \( Z \) are inputs, \( v^i \) and \( e^j \) are utility and expenditure functions, \( \pi^j \) and \( c^j \) are profit and cost functions, and each relative input price \( \left( \frac{w_\beta}{w_\alpha} \right) \) is treated as a one, not two variables. Based on equation (1)\(^{11,12} \), the Hicks’ hypothesis and our alternative hypotheses can be formulated as:

\[ H_0: \forall s, \sum_{s=2}^{s=4} \beta_s > 0, \quad H_1: \forall s, \sum_{s=2}^{s=4} \beta_s \leq 0 \]

(2)

which may be interpreted a custom/made non-parametric statistical test (not based purely on individual parameter values but rather on their combinations)\(^{13} \), a combination of a) Student’s \( t \)-test, b) Wald’s \( F \)-test\(^{14} \), and c) arithmetic comparison, non-dependent on any of the model’s further modifications, such as eventual exogenous variables.

2 Employed methodology

The nature of our model, applied to panel data, requires the use of the generalized method of moments (GMM), which has the following level and first differences versions, depending on data’s time series characteristics (unit roots), after a logarithm transformation of equation (1):

\[ E(V - \Gamma \log X | \log X, \log I) = 0, \]

\[ E(\Delta V - \Gamma \Delta \log X | \Delta \log X, \Delta \log I) = 0 \]

(3)

10 Proof: If \( T = k \cdot N \), then \( T \propto N \) and \( z_T \propto z_N \), since \( T \equiv z_T \) and \( N \equiv z_N \). If \( w_N = \kappa(N) \) and \( w_T = \lambda(T) \), then \( w_T = \lambda(k \cdot N) = \tilde{k} \cdot \tilde{\lambda}(N) \) (for a homothetic function \( \tilde{k} = k \) and \( \tilde{\lambda} = \lambda \)), ergo \( w_T \propto w_N \).

11 In the detailed model specification, the mark “\( \tilde{\lambda} \)” indicating equilibrium values of economic variables is omitted.

12 In this paper, square brackets (crochets) serve the same purpose as parentheses and are used as their alternative to identify and distinguish between parts of equations.

13 The theory of statistics tends to define all non-standard tests as non-parametric which is of course subject to dispute and significantly exceeds the scope of this paper.

14 Wald \( F \)-tests are used to estimate statistical significance of groups of parameters, e.g. regression coefficients, in formal models.
where $V$ is the vector of dependent variables, $\Gamma$ and $X$ are matrices of coefficients and explanatory variables, $I$ is the matrix of eventual exogenous variables. The GMM weights in this paper are calculated, according to the standard rule-of-thumb formula:

$$W = (\text{number of observations}) \ast (\log X, \log I)'(\log X, \log I),$$

$$\tilde{W} = (\text{number of observations}) \ast (\Delta \log X, \Delta \log I)'(\Delta \log X, \Delta \log I)$$

(4)

To choose between the level and first difference versions of the econometric model, we employ the augmented Dickey and Fuller, ADF, test and the Kwiatkowski–Phillips–Schmidt–Shin, KPSS, tests with constant and a time trend. In this paper, we use Choi’s meta-analysis, where the overall (composite) p-value is estimated from the p-values of ADF tests and p-value range estimations of KPSS tests for individual cross-sections.\(^{15}\)

3 Data

Data for all variables were retrieved from official sources, specifically, from the World Intellectual Property Organization (WIPO), Penn World Tables 9.0 (University of Groningen), the International Monetary Fund (IMF), International Labour Organisation (ILO), the World Bank Group (WB), United Nations Development Programme (UNDP), as well as from the United Nations Conference on Trade and Development (UNCTAD). A detailed overview of the panel dataset, which comprises 154 countries for years 1980–2015 (5544 rows and ca. 7.1% missing observations), is provided in Table 1. The following modifications were performed in the panel dataset before performing the GMM estimation: a) missing values were interpolated by arithmetic and geometric means and extrapolated with the help of repeating border values; b) for the calculation of $w_{NK}$ and $w_{NA}$, real wages were logarithmized in order to smooth the effects of monetary policies and inflation in specific years. We also adopted the methodology of the Global Innovation Index (GII), the leading yearbook on innovation in the world, to add several exogenous variables into the dataset.

Tab. 1: Panel dataset overview

<table>
<thead>
<tr>
<th>Description</th>
<th>Unit / Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPP</td>
<td>Intellectual property protection, number of applications</td>
</tr>
</tbody>
</table>

\(^{15}\) There is no single methodology for verifying the presence of unit roots in time series and panel datasets. Hence, it is always advisable to use more than one test. ADF and KPSS have different logic: the $H_0$ hypothesis of the ADF test is the presence of a unit root in the time series, whilst the $H_0$ of the KPSS test is the time series’ stationarity. Choi meta-tests may also not be the only approach for panel unit root testing but we choose them over the 100% pooled Levin-Lin-Chu and Harris-Tzavalis tests.
### Total prices, index

Consumer price index, 2005 = 100

### Natural resources

constant

### Qualified labour, million persons with human capital

Obtained as a product of million persons employed and Penn World Table 9.0. Index of human capital per person, based on years of schooling and returns to education (the $hc$ variable).

### Capital stock

Stock of capital, constant prices of 2011 $US$

### Prices of capital / prices of qualified labour

Real interest rate / real wage index, 2005 = 100

### Prices of qualified labour / prices of commodities

Real wage / International Monetary Fund’s commodity price index, 2005 = 100

### Prices of capital / prices of commodities

Real interest rate / International Monetary Fund’s commodity price index, 2005 = 100

### Exogenous variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>INST</td>
<td>Quality of institutions, index</td>
<td>Transparency international &amp; Others Corruption Perception Index, 0–10</td>
</tr>
<tr>
<td>HCR</td>
<td>Length of schooling and life expectancy, index</td>
<td>Obtained as a residual of ordinary least squares (OLS) regression of human development index into on an improvised gross national index per capita in purchasing power parity. Non-collinear with $zQN$, according to the detrended correlation coefficient.</td>
</tr>
<tr>
<td>INFR</td>
<td>Quality of infrastructure, share</td>
<td>Gross capital formation, current prices, share in GDP</td>
</tr>
<tr>
<td>MBSex</td>
<td>Market and business sophistication indicator - part 1, share</td>
<td>Exports of goods and services, share in GDP</td>
</tr>
</tbody>
</table>
3 Results

Preliminary estimations show heterogeneous non-stationarity of time series inside the above-described panel variables, hence, we recur to their stationarization via log-differences (ld_), which are employed in the model, as shown in Table 2 (all missing observations were omitted in calculation).

Tab. 2: Iterated GMM estimation in gretl

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-Value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>w0</td>
<td>-0.21583</td>
<td>0.0664267</td>
<td>-3.249</td>
<td>0.0012  ***</td>
</tr>
<tr>
<td>w1</td>
<td>-0.18427</td>
<td>0.0679310</td>
<td>-2.713</td>
<td>0.0067  ***</td>
</tr>
<tr>
<td>a0</td>
<td>-0.03130</td>
<td>0.0198508</td>
<td>-1.577</td>
<td>0.1148</td>
</tr>
<tr>
<td>b0</td>
<td>-0.01053</td>
<td>0.0233966</td>
<td>-0.4498</td>
<td>0.6528</td>
</tr>
<tr>
<td>b1</td>
<td>0.69713</td>
<td>0.325927</td>
<td>2.139</td>
<td>0.0324  **</td>
</tr>
<tr>
<td>b2</td>
<td>0.39677</td>
<td>0.159380</td>
<td>2.489</td>
<td>0.0128  **</td>
</tr>
<tr>
<td>b3</td>
<td>0.01991</td>
<td>0.0403584</td>
<td>0.4934</td>
<td>0.6218</td>
</tr>
<tr>
<td>b4</td>
<td>-0.03087</td>
<td>0.0370939</td>
<td>-0.8322</td>
<td>0.4053</td>
</tr>
<tr>
<td>z0</td>
<td>2.0791e-05</td>
<td>0.000816102</td>
<td>0.02548</td>
<td>0.9797</td>
</tr>
<tr>
<td>z1</td>
<td>1.4138e-05</td>
<td>0.000207013</td>
<td>0.06830</td>
<td>0.9455</td>
</tr>
</tbody>
</table>

GMM criterion: Q = 0.00404234 (TQ = 7.87447)
J test: Chi-square(5) = 7.87447 [0.1633]

Source: self-prepared
The only significant parameters of the model, according to the iterated GMM estimation, the quality of which was confirmed by Sargan-Hansen J-test, are $w_0$, $w_1$, $b_1$ and $b_2$. Parameters of relative input prices $b_3$ and $b_4$, crucial for this paper, have been found statistically non-significant at all 1%, 5% and 10% probability levels, which proves the alternative hypothesis $H_1$ and reject $H_0$ in our non parametric test. Parameter $b_5$ and variable $wKA$ were omitted from the model because of the exact collinearity.

**Conclusion**

No significant disagreement exists about the notion that innovations are beneficial to individual companies as well as to economic development of entire economies. There is however little consensus on what is the cause of innovation. The Hicks hypothesis of “induced innovation” is a classic example of a long-standing innovation theory as much as of many inconclusive discussions. In our paper, we have attempted to falsify it by offering a hypothesis alternative to the Hicks’ theory, developed a theoretical model to justify it and constructed and performed an econometric test of the model. Our suggested way of putting the hypothesis to scrutiny by data represents a novel non-parametric (non-standard and not purely parameter value based) statistical test based on combinations of parameter values from an economic model and arithmetic operations, which creates the first (methodological) value added of our research. We agree with Hicks in principle that an increase in the relative price of a factor of production can carry a motivation to replace it through innovation, however, simple increase of price of one (or several) factors might not be enough. More to the point, inflating prices of input(s) can be debilitating as company or industry loses competitiveness and diminishes. Our key point thus is that the increase in relative price of one factor mandate relatively low price level of other factor(s) of production as a compensation of cost for companies for the innovations to take place. Our econometrical results haven’t backed the Hicks’ hypothesis but rather our alternative one, which forms the second (practical) value added of our research.

**Acknowledgment**

This paper was elaborated under the institutional support of the Faculty of International Relations, University of Economics, Prague, project No. IP200040
Annex: Formal derivation of the model

Under the condition of \( \forall \ Y_i \geq 0 \), \( \text{Innovative attributes} = \sum_i \psi_i Y_i, \ Non-innovative \) attributes = \( \sum_i \chi_i Y_i \), where \( i \) is the number of good/service \( Y \), and \( \psi_i \) and \( \chi_i \) are the shares of the two characteristics in each \( Y_i \).

Demand side:

\[ d^j \propto \text{NPC}^j_{d^j} \quad \eta(i) = e^j(i, \Pi, \bar{U}) = \min_{Y_i} \sum p_i Y_i \quad \text{s.t.} \quad \sum \psi_i Y_i \leq y \]

where \( j \) is the end consumer, \( \text{NPC}^j_{d^j} \) is the “novelty/innovation product curve”, from which the individual demand for innovations is derived, and \( p_i \) is the price of of \( Y_i \), is equivalent to:

\[ d^j \propto \text{NPC}^j_{d^j} \quad \eta(i) = e^j(i, \Pi, \bar{U}) = \min_{Y_i} \sum p_i Y_i \quad \text{s.t.} \quad \sum \psi_i Y_i \leq y \]

\[ \eta(i) = e^j(i, \Pi, \bar{U}) = \max_{Y_i} \left( \sum \psi_i Y_i, \sum \chi_i Y_i \right) \]

\[ \text{ergo} \]

\[ d^j_{\text{innovations}} = \sum \psi_i Y_i = \xi \left( e^j \left( i, \Pi, \nu^j(i, \Pi, y) \right) \right) = f^j(\Pi) + g(\Pi) \cdot \bar{U}^j(i, \Pi, y) + h^j \]

and

\[ A D_{\text{innovations}} = \sum d^j_{\text{innovations}} = \sum_e \sum \psi_i Y_i = \sum_e \xi \left( e^j \left( i, \Pi, \nu^j(i, \Pi, y) \right) \right) = f(\Pi) + g(\Pi) \cdot \bar{U}^j(i, \Pi, y) + h \]

\[ \text{ergo, under constant } \nu^j \text{ and } e^j, \]

\[ AD_{\text{innovations}} = \lambda(\Pi | \nu^j, e^j ... \text{ const}) \]

Supply side:

\[ s^j \propto \text{NPC}^j_{s^j} \quad \rho(i) = \pi^j(i, \Pi, IC) = \max_{\pi} \sum p_i Y_i - IC \quad \text{s.t.} \quad \forall \ z_{k,k=RP} \sum w_k z_k \leq IC \]

where \( j \) is the producer, \( \text{NPC}^j_{s^j} \) is the “novelty/innovation product curve” from which the individual supply of innovations is derived, \( k \) is the input, and \( w_k \) is the absolute price of the input \( z_{k,k=RP} \in Z \), is equivalent to:

\[ s^j \propto \text{NPC}^j_{s^j} \quad \phi(i) = c^j(i, Z, \bar{\pi}) = \min_{\sum z_{k,k=RP}} \sum w_k z_k \quad \text{s.t.} \quad \sum p_i Y_i - IC \geq \bar{\pi} \]

\[ \text{ergo} \]

\[ s^j_{\text{innovations}} = \sum \psi_i Y_i = \zeta \left( c^j \left( i, Z, \pi^j(i, \Pi, IC) \right) \right) = \bar{l}^j(Z) + m(Z) \cdot \bar{\nu}^j(i, \Pi, IC) + n^j \]

and
\[ \text{AS}_{\text{innovations}} = \sum_{i} s_{\text{innovations}}^{j} = \sum_{i} \sum_{i} \psi_{i} Y_{i} = \sum_{i} \zeta \left( c^{i} \left( i, Z, \pi^{j}(i, \Pi, IC) \right) \right) = \]

\[ = l(Z) + m(Z) \cdot \pi(i, \Pi, IC) + n \]

\[ \text{ergo, under constant } \pi^{j} \text{ and } c^{i}, \]

\[ \text{AS}_{\text{innovation}} = \mu(Z, \Pi | \pi^{j}, c^{j} \ldots \text{const}), \quad Z = \{w_{N}/w_{K}, w_{N}/w_{A}, w_{N}/w_{T}, w_{K}/w_{A}, w_{K}/w_{T}\} \]

**References**


**Contact**

Ilya Bolotov, Ph.D., MBA  
Dept. of International Business  
University of Economics, Prague  
W. Churchill Sq. 4  
130 67 Prague 3, Czech Republic  
ilya.bolotov@vse.cz
Tomáš Evan, PhDr., Ph.D.
Dept. of Software Engineering
Czech Technical University
Thákurova 9,
160 00 Prague 6, Czech Republic
tomas.evan@fit.cvut.cz