JOURNAL OF APPLIED ECONOMICS AND BUSINESS

VOL. 8, ISSUE 3 – SEPTEMBER, 2020

Education and Novel Technology Research Association

Journal of Applied Economics and Business

VOL. 8, ISSUE 3 – SEPTEMBER, 2020

The Journal of Applied Economics and Business (JAEB – ISSN: 1857-8721) is an international peer-reviewed, open-access academic journal that publishes original research articles. It provides a forum for knowledge dissemination on broad spectrum of issues related to applied economics and business. The journal pays particular attention on contributions of high-quality and empirically oriented manuscripts supported by various quantitative and qualitative research methodologies. Among theoretical and applicative contributions, it favors those relevant to a broad international audience. Purely descriptive manuscripts, which do not contribute to journal's aims and objectives are not considered suitable.

JAEB provides a space for academics, researchers and professionals to share latest ideas. It fosters exchange of attitudes and approaches towards range of important economic and business topics. Articles published in the journal are clearly relevant to applied economics and business theory and practice and identify both a compelling practical issue and a strong theoretical framework for addressing it.

The journal provides immediate open-access to its content on the principle that makes research freely available to public thus supporting global exchange of knowledge.

JAEB is abstracted and indexed in: DOAJ, EZB, ZDB, Open J-Gate, Google Scholar, JournalITOCs, New Jour and UlrichsWeb.

Publisher

Education and Novel Technology Research Association

Web: www.aebjournal.org

E-mail: editorial@aebjournal.org

support@aebjournal.org

publisher@aebjournal.org

Content

Journal of Applied Economics and Business

OPEN ACCESS ARTICLE

AN EMPIRICAL ASSESSMENT OF THE SYSTEMATIC RELATIONSHIP BETWEEN INTEREST RATE AND EXPECTED INFLATION IN TANZANIA

Nguling'wa Philip Balele

Bank of Tanzania, Tanzania

Abstract

The theorization of the Fischer effect considers expected inflation as the driver of short-term interest rates. However, empirical data does not show a clear support of this theory implying that there are other drivers of interest rates beyond inflation expectations. This paper considers existence of a systematic relationship between the short-term interest rate and expected inflation, specifically investigating whether expected inflation has impact on short term interest rates, other factors held constant. In achieving this objective, econometric methods are applied on Tanzanian data obtained from the Bank of Tanzania. Using OLS and correlation coefficient analyses, the findings indicate existence of a systematic positive relationship between the nominal short-term interest rate and expected inflation, albeit weak with short term interest rates responding to expected inflation. These findings imply that the short-term interest rate, particularly the overnight interbank cash market rate is a potential policy rate for the Bank of Tanzania monetary policy. The findings provide empirical evidence that the intermittent unsystematic relationship between the short-term nominal interest rate and expected inflation are temporary phenomena along the path of nominal interest rate. This justifies the central bank's transition from the monetary targeting framework to the interest rate-based framework. The study distinguishes itself from the existing literature by initially taking account of the causality between expected inflation and short-term nominal interest rate before estimation of the effect of expected inflation on short term interest rate.

Keywords: Fisher hypothesis; Fisher equation; Bank of Tanzania.

INTRODUCTION

Liberalization of interest rate in 1991 among other financial sector reforms, constitutes the 1990s package of financial sector reforms in Tanzania. The interest rate reforms entailed a switch from a controlled interest rate policy that had existed until 1991 to the prevailing market-determined interest rate policy. During the controlled economic regime, interest rates were determined administratively by the government. The financial market was not supportive to economic growth because of its small size, comprised of a few government-owned commercial banks, two pension funds, one insurance company, and one hire purchase company. Given that interest rates were exogenously determined, there was no systematic relationship among interest rate, inflation, and money supply growth.

With the current market-determined interest rate policy, real interest rates respond to movements in expected inflation. This implies that inflation is among the drivers of the real cost of funds. In market economies, movements in inflation and interest rates imply an opportunity for central banks to influence macroeconomic variables. Lending rates are determined by several factors including the cost of funds, risk profile of customers, profit margins, Treasury bill rates, and quality of collaterals.

Since liberalization of interest rates in 1990s Tanzania, interest rates have been determined by market forces and hence have moved in similar patterns with inflation albeit with lags. During this period, interest rates on average have responded to movements in inflation (Figure 1).

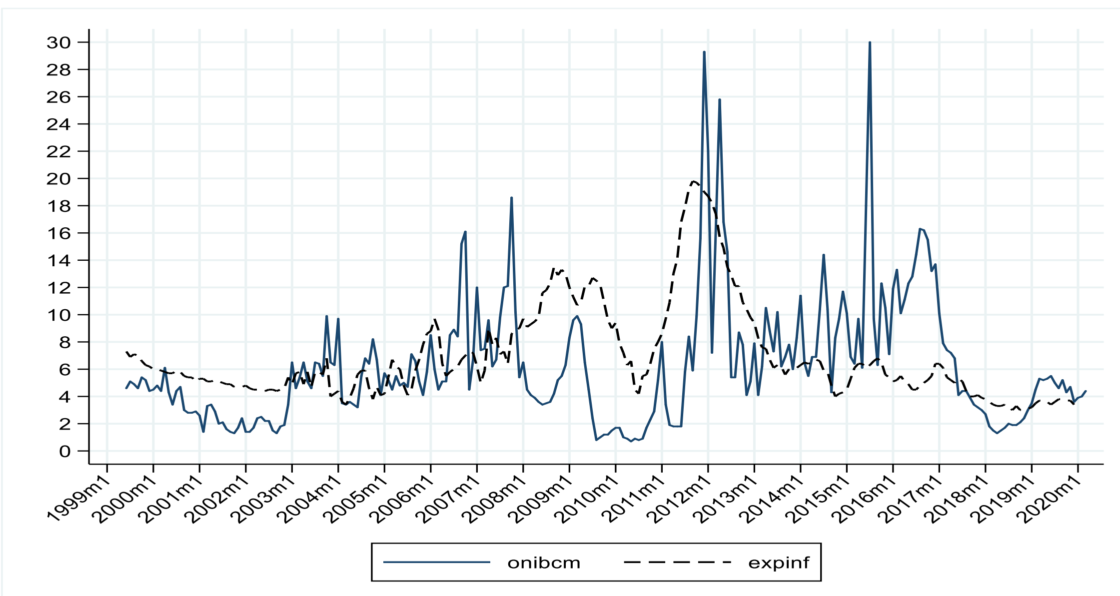


FIG 1. INTEREST RATE AND EXPECTED INFLATION DYNAMICS

For the past two decades, expected inflation and short-term interest rates, particularly the overnight interbank cash market rate have moved dynamically with expected inflation leading or predicting the nominal short-term interest rate. Comparing the movements, interest rate depicted volatilities with marked spikes at some intervals, particularly around 2011 and around 2015. As the closest measure of cost of funds, the overnight interbank cash market rate portrays volatilities, reflecting market liquidity dynamics in the economy.

We test whether there is a systematic relationship between the short-term nominal interest rate and expected inflation. Specifically, we test the hypotheses:

- Expected increase in inflation leads to increase in short-term nominal interest rate.
- There is a positive partial relationship between expected inflation and the short-term nominal interest rate.
- Expected inflation and the short-term nominal interest rate have a long-run relationship.

In terms of contributions, this study provides empirical evidence on whether the intermittent unsystematic relationship between short-term nominal interest rate and inflation temporary phenomena along the path of nominal interest rates. The study also sheds some light on whether the Bank of Tanzania is ready to modernize its monetary policy framework, the process which constitutes switching from monetary targeting framework to the price-based framework, in which central bank policy rate is the instrument of monetary policy.

LITERATURE REVIEW

Theoretical literature review

The starting point in this study is the Fisher hypothesis (1930), which suggests a one-for-one adjustment of the nominal interest rate to changes in expected inflation. In a simple mathematical form, this can be written as

$$r_t = i_t - \pi^e \quad (1)$$

where r_t is the real interest rate in the current period, i_t is the nominal interest rate in the current period and π^e is expected inflation in period t+k. The one-for-one movement between the nominal interest rate and expected inflation is a stricter condition such that it is not supported by empirical data because it implies that the real interest rate is fixed. This is not feasible because real interest rate is also influenced by other factors such as investment returns and agents' preference. The one-for-one condition does not hold; instead, a less than the one-for-one relationship between expected inflation and interest rate is not only theoretically relevant but also supported by empirical evidence according to literature (Fisher, 1930; Mishkin, 1992). This outcome allows comparison between the Fisher hypothesis with the theory of rational expectations and efficient market hypothesis given that the two imply that nominal interest rate movements reflect movements in expected inflation. While the weak form of the Fisher hypothesis corresponds to the weak form of efficient market hypothesis, the strict form of the Fisher hypothesis corresponds to the strong form of the efficient market hypothesis. However, the strict form of the Fisher hypothesis and strong form of efficient market hypothesis are rarely supported by empirical evidence in the literature. This requires relaxing the Fisher (1930) condition of one-for-one relation between the nominal interest rate and inflation rate by imposing a condition that expected inflation have impact on short-term nominal interest rate. That is, nominal interest rate movements reflect expected inflation, that is, $\frac{\delta i_t}{\delta \pi^e} = \frac{\delta i_t}{\delta \pi_t} > 0$.

The above condition can be broken down into two parts: First, the impact of expected inflation on interest rate is positive but less than unit (i.e. $0 < \frac{\delta i_t}{\delta \pi^e} < 1$) (e.g. Fisher, 1930; Mundel, 1963; Tobin, 1965). Second, the impact of expected inflation on interest rate is greater than unit (i.e. $\frac{\delta i_t}{\delta \pi^e} > 1$), consistent with the Darby-Feldstein effect (Darby, 1975; Feldstein, 1976) who suggest a modified Fisher equation that includes a tax parameter, which is also corroborated by Crowder and Hoffman (1996). However, Carmichael and Stebbing (1983) and Choudhry (1997) find a negative relationship between expected inflation and interest rate or the inverted Fisher effect.

The Fisher effect suggests that movements in expected inflation are associated with movements in nominal interest rate with some lags (Figure 1), consistent with the weak form of efficient market hypothesis. Specifically, returns on assets reflect developments in the market, including inflation developments. The Fisher hypothesis implies that central banks must keep inflation low to ensure low nominal interest rates. Keeping inflation low and stable has been one of the goals of many central banks under the price stability and financial stability objectives.

According to the monetary model of exchange rate determination, policy-driven rises in the nominal interest rates decrease real money balances, leading to decrease in inflation and appreciation of currency (Balele, 2013; Chinn, 1999; Edwards, 1983). This implies that the increase in the interest rate and inflation achieved by cutting money supply lead to currency appreciation. For some periods in Tanzania, the pattern among interest rate, inflation, money supply, and exchange rate has moved inconsistently. For instance, towards the end of 2011, inflation was rising, driven by food prices, but there was no justified significant growth in broad money supply. At the same time, the three-month Treasury bill rate was also rising. This puzzle can be associated with the inverted Fisher effect as coined by Carmichael and Stebbing (1983). Contrary to expectations, short-term interest rates from time to time tend to decouple from the fundamentals, particularly delinking from trends in money supply, inflation, and exchange rate. This is also contrary to the liquidity preference theory.

Despite extensive investigation of the Fisher hypothesis (1930), the debate is still inconclusive given that the findings are characterized by mixed views (Shabir, 2010). While some studies support the Fisher hypothesis of monetary policy neutrality, others reject it. Nevertheless, both options contain important economic implications in terms of monetary policy and macroeconomics (Mishkin & Simon, 1995; Shabir, 2010). For instance, evidence in support of the Fisher hypothesis implies that interest rate is endogenous such that a central bank cannot effectively achieve monetary policy goals using interest rate as an instrument of monetary policy. This is consistent with efficient market hypothesis prediction which argues that economic agents take advantage of market information, both private and public because such information is already reflected in the market (Fama, 1975). That means movements in short-term interest rates reflect inflation expectations. This argument is also consistent with the Real Business Cycle hypothesis which assumes that at any point in time along the business cycle, the economy optimizes such that policy interventions may lead to sub-optimal outcome level. On the other hand, the dearth evidence in support of the Fisher hypothesis implies that interest rate is exogenous to central banks such that it can be effectively used to achieve monetary policy goals.

The common empirical finding among the Fisher hypothesis studies is the less than one-for-one relationship between expected inflation and nominal interest rate instead of a one-for-one relation predicted by Fisher (1930). The literature raises several reasons in explaining both the less than one-for-one outcome and the complete breakdown of the Fisher hypothesis as described in the empirical literature review section. Given the objectives of this study, the empirical literature focuses on the strength (i.e. coefficient size) of the relationship between the nominal interest rate and expected inflation, particularly the evidence of the one-for-one and the less than one-for-one outcomes.

Empirical literature

Mundel (1963) and Tobin (1965) describe the less than one-for-one Fisher's (1930) outcome in terms of wealth effect, that is, the Mundel-Tobin effect, arguing that wealth imposes some lags on the relationship between expected inflation and nominal interest rate. Specifically, expected inflation decrease the real interest rate by lowering real money balances, and the consequent decrease in wealth stimulates savings in form of equities and bonds (Mundel, 1963). Further, inflation pressures tend to increase the demand and hence the stock of real assets held in the economy, leading to a decrease in asset returns (Tobin, 1965). In view of this, increase in expected inflation is associated with a less than proportionate increase in interest rate. This is referred to as the Mundel-Tobin effect.

The Darby-Feldstein effect (Darby, 1975; Feldstein, 1976) provides another argument for the less than one-for-one Fisher effect in which the tax structure influences the Fisher relationship. To insulate the borrowers and lenders expected real payments, nominal interest rates need to increase by more than one-for-one with the increase in expected inflation to ensure positive real returns (Darby, 1975). Specifically, for each basis point increase in expected inflation, the nominal interest rate should increase by $\frac{1}{1-\tau}$ basis points, where τ is the tax parameter which ranges between zero and unit. This assumption requires modification of the ordinary Fisher relationship in equation (1) by incorporating a tax parameter τ, yielding:

$$(1-\tau)i_t = r_t + \pi^e \qquad (2)$$

where $0 < \tau < 1$

Taking partial derivative of the nominal interest rate with respect to expected inflation and considering the requirement $0 < \tau < 1$, it gives

$$\frac{\delta i_t}{\delta \pi^e} = \frac{1}{1-\tau} > 1$$

which implies that movements in the pre-tax interest rate are driven by fiscal (i.e. tax) decisions embedded in inflation expectations. This implies that the higher the tax rate, the larger the change in the nominal interest rate as a result of the expected higher inflation rate.

Using the US three-month Treasury bill rate as a proxy for the market interest rate during 1953q1-1978q4, Carmichael and Stebbing (1983) note the complete breakdown of the Fisher hypothesis in terms of an inverted Fisher effect. Their analysis assumes substitutability between money and financial assets and that interest rate on financial assets is invariable over time such that real interest rate and expected inflation move in inverse direction. This can be written as

$$r_t = \bar{\imath} - \pi^e \qquad (3)$$

r_t and π^e are as defined in equation (1), $\bar{\imath}$ is the nominal interest rate, invariant to financial assets.

Equation (3) implies an inverse relationship between the real interest rate and expected inflation (i.e. $\frac{\delta i_t}{\delta \pi^e} = -1$). However, this specification is not supported by empirical evidence (e.g. Choudhry, 1997; Moazzami, 1991; Woodward, 1992), instead, the real interest rate is not fixed, rather it is endogenously determined by return on real investment and economic agents' time preference.

Using correlation coefficient, Fisher (1930) examines the relationship between nominal interest rate and expected inflation using US and UK annual data for the period 1890-1927 and 1820-1924, respectively. Their findings show a less than one-for-one relationship between nominal interest rate and expected inflation in the short-run, that is, correlation coefficients of 0.86 and 0.98 for the US and UK, respectively. These findings imply that inflation expectations are slightly less instantaneously reflected in the nominal interest rates in the US compared to the UK. According to Fisher, the one-for-one outcome is on account of money illusion in which money holders enjoy holding money irrespective of the opportunity cost involved in terms of foregone interest income.

Mishkin and Simon (1995) examine the Fisher effect using Australian data and find evidence in support of the long-run Fisher hypothesis but rejects the Fisher hypothesis in the short-run. This implies that, while short-run movements in nominal interest rates reflect movements in monetary policy, the long-run movements in interest rates are driven by inflationary expectations. This outcome is more appealing since it is consistent with short-term interest rate rather than the long-term interest rate as a central bank choice tool of monetary policy.

Using Japanese and US data in the 1990s, Mishkin (1992) provides two important empirical applications of the Fisher hypothesis. The first empirical application is the Japanese case in the 1990s when prolonged recessions were accompanied by deflation when the six-month Treasury bill rate turned slightly negative. The second interesting empirical application is based on the US data which shows rising and falling interest rate (i.e. three-month Treasury bill rate) during the business cycle expansions and recessions, respectively.

ESTIMATION METHODOLOGY

Variables and data sources

Monthly data on short-term interest rates[1], and inflation were obtained from the Bank of Tanzania. The data covers the period from June 1999 to March 2020.

Choosing the representative short-term interest rate

The correlation coefficient test results suggest statistically significant relationship between the overnight interbank cash market rate and expected inflation. However, the results indicate no statistically significant relationship between expected inflation and the remaining short-term interest rates at all conventional significance levels. This finding

[1] These are overnight interbank cash market rate, repo rate, three-month Treasury bills rate, and Lombard rate.

serves as an independent test for the Fisher hypothesis, widely referred to in the literature. Table 1 reports the correlation coefficients between expected inflation and a set of short-term interest rates.

TABLE 1. CORRELATION COEFFICIENTS BETWEEN INFLATION AND SHORT-TERM INTEREST RATES

	exp_inf	$nibcm$	$repo$	$tmtb$	$lombard$
exp_inf	1.0000				
$onibcm$	**0.2738**	1.0000			
	(0.0001)				
$repo$	0.1040	**0.7139**	1.0000		
	(0.1480)	(0.0000)			
$tmtb$	0.0958	**0.5494**	**0.4696**	1.0000	
	(0.1828)	(0.0000)	(0.0000)		
$lombard$	0.0850	**0.9314**	**0.8084**	**0.5841**	1.0000
	(0.2795)	(0.0000)	(0.0000)	(0.0000)	

Notes: In brackets are p-values and bold implies significant at 5% level and below.
exp_inf is expected inflation rate, onibcm is overnight interbank cash market rate
repo is repo rate, tmtb is the three-month Treasury bill rate, and lombard is the Lombard rate.

Based on the statistically significant correlation between nominal interest rate and expected inflation, the overnight interbank cash market rate is considered to represent short-term interest rates. The correlation between expected inflation and the interbank cash market rate is less than unit (i.e. 0.27). This is consistent with Fisher's (1930) finding of less than one-for-one relationship (Table 1). However, the small size correlation indicates a weak relationship between expected inflation and short-term nominal interest rate. Formal econometric tests are conducted in section 4.0.

The model

This study tests the Fisher hypothesis using econometric tests in addition to the correlation coefficients reported in Table 1. Since there is no data on expected inflation, we assume perfect foresight to allow using future values of inflation as proxy for expected inflation.

The Fisher effect is mainly on the relationship between the nominal interest rate and expected inflation as specified below.

$$1 + i_t = (1 + E_t\{r_{t+k}|\Omega_t\})(1 + E_t\{\pi_{t+k}|\Omega_t\}) \tag{4}$$

where i_t is the nominal interest rate known at time 't,' r_{t+k} is the real return on assets expected by economic agents expect in period 't+k' given expected inflation in period 't+k,' and π_{t+k} is the expected rate of inflation in time 't+k.' For small values of nominal interest rate and expected inflation, equation (4) simplifies to

$$i_t = r_t + \pi_{t+k} \quad (5)$$

If agents' expectations are correct (i.e. perfect foresight), the Fisher equation (5) imposes an implicit assumption that the real interest rate is fixed because the nominal interest rate is driven by expected inflation and hence follows the evolution of expected inflation, as predicted in Fisher hypothesis (1930). Estimation and analysis of the results of equation (5) serve more than one purpose in the literature. While it is used to test for market efficiency (Fama, 1975), it is used in testing the relationship between nominal interest rate and expected inflation under the Fisher hypothesis. The two hypotheses are similar since the strong form of the efficient market hypothesis corresponds to Fisher's (1930) finding of a one-for-one relationship between nominal interest rate and expected inflation.

The limitations of the strong efficient market hypothesis and strict Fisher hypothesis make them weakly testable, especially in developing countries where financial markets are not well developed. To estimate the Fisher hypothesis econometrically, an error term (i.e. $\in_t$) and a constant term (i.e. γ) are tagged in equation (5). We assume the constant term γ to be an estimate of the real interest rate as assumed in the Fisher hypothesis. The same interpretation of the constant term γ applies to the Taylor rule estimates (Clarida et al., 2000; McCallum, 2002).

$$i_t = \gamma + \beta\pi_{t+k} + \mu_t \quad (6)$$

The strict form of the Fisher hypothesis is when the null hypothesis is $\beta = 1$ and the residuals (μ_t) are stationary. If the null hypothesis holds, then full adjustments in the nominal interest rate are determined by expected inflation, which is the one-for-one outcome. The more appealing analysis of the Fisher hypothesis is testing its weak form such that the null hypothesis is positive but less than unit and the residuals are stationary. If this holds, then partial adjustments in the nominal interest rate are said to depend on expected inflation while the remaining adjustments come from other sources such as productivity of capital and investor's time preference. This motivates testing for the nominal interest rate adjustment by imposing the lagged nominal interest rate on the right-hand side of equation (5), which then becomes

$$i_t = \alpha + \rho i_{t-1} + \lambda\pi_{t+k} + \varepsilon_t \quad (7)$$

The central bank is said to conduct monetary policy adjustment if the coefficient on the lagged coefficient (i.e. ρ) lies between zero and unit (i.e. $0 < \rho < 1$).

EMPIRICAL RESULTS

Introduction

In testing the Fisher hypothesis, we conduct a Granger causality Wald test and estimate the model by ordinary least squares (OLS). We include a lagged dependent variable to examine whether monetary policy involves adjustment.

Unit roots test

Before conducting the causality test and OLS estimation, we examine the time series characteristics of the data by conducting unit roots test. We employ appropriate lag length values obtained by testing for the optimal number of lags using the Final Prediction Error and the Akaike Information Criteria (Appendix). Visual inspection of Figure 1 suggests that expected inflation involves drift(s) and therefore a drift is added in the Augmented Dickey Fuller test to capture such behavior. The results are reported in Table 2.

TABLE 2. SUMMARY OF UNIT ROOTS TEST RESULTS

	Augmented Dickey-Fuller (ADF) test		
Variable	t-stat	Cr-value (5% level)	p-value
exp_inf	-2.765*	-2.884	0.0634
$onibcm$	-3.071**	-2.884	0.0288
$repo$	-5.258***	-2.884	0.0000
$tmtb$	-3.039**	-2.884	0.0373
$lombard$	-3.076**	-2.886	0.0284

Note: MacKinnon (1996) one-sided p-values
*** significant at 1%, ** significant at **, and * significant 10%.

The unit-roots test results show that all variables are stationary at the conventional significance levels. The unit-roots test findings for all the nominal interest rate categories are consistent with the intuition that interest rates have a zero lower bound which implies that investors are not willing to accept negative nominal returns. Given this, monetary authorities tend to ensure positive and stable interest rates as part of macroeconomic stabilization measures. The same is also consistent with the Taylor rule principle which suggests that a central bank will hike the policy rate by more than unit in response to a unit increase in inflation (Balele, 2013). Stationarity of the variables imply absence of the long run but a dynamic relationship between expected inflation and short-term interest rate.

Granger causality Wald test

The literature on the Fisher effect assumes one-way causality from expected inflation to nominal interest rate. However, there is no consensus on causality between nominal interest rate and expected inflation. For instance, while studies on the Fisher effect investigate the effect of inflation on interest rates (e.g. Fisher, 1930; Darby, 1975; Feldstein, 1996), other studies (e.g. Barsky & Delong, 1991) assess the effect of nominal interest rate on inflation.

A source of uncertainty on causality between nominal interest rate and expected inflation is that most of the Taylor rule and Taylor-type rule studies are inclined to application of the generalized method of moments as a remedy to the endogeneity problem between interest rate and expected inflation (e.g. Balele, 2013; Rotich et al., 2007; Sanchez-Fun, 2002). This uncertainty motivates testing for causality between expected inflation and interest rate to ascertain the direction of movement between the two variables. Granger causality Wald test results are reported in Table 3.

TABLE 3. GRANGER CAUSALITY WALD TEST

Equation	Excluded	chi2	df	Prob>chi2
$onibcm$	exp _inf	5.4234	2	0.066
$onibcm$	All	5.4234	2	0.066
exp_inf	0nibcm	0.3541	2	0.838
exp_inf	All	0.3541	2	0.838

Based on the statistically significant p-values of the Granger causality Wald test results, we reject the null hypothesis that expected inflation does not cause interest rate but cannot reject the null hypothesis that interest rate does not cause expected inflation. This implies a one-way causality from expected inflation to interest rate. This further implies that expected inflation predicts short-term interest rates which is consistent with theory, particularly the Fisher hypothesis.

OLS estimates

TABLE 4. OLS REGRESSION RESULTS

Coefficient	Estimate	p-value
α	0.95**	0.043
ρ	0.72***	0.000
λ	0.11*	0.052
F(2, 242)	156.20	0.000
Adjusted R-squared	0.56	
No. observations	245	

Note: In brackets are p-values

***stationary at 1%; **stationary at 5%; and *stationary at 10% level.

We estimate the modified Fisher equation (7) using OLS by including the lagged dependent variable to examine whether monetary policy involves partial adjustment. Table 4 reports the OLS regression results.

Discussion of the findings

According to the F-statistic, the overall model is statistically significant at the 1% level. The coefficient of the adjustment term is 0.72 and statistically significant at the 1% level with the expected positive sign. This implies that monetary policy does adjustments to smooth short-term fluctuations in nominal interest rate. Inclusion of the lagged dependent variable guarantees absence of serial correlation in the model. The constant term which is interpreted as the real interest rate is 0.95 and is statistically significant at 5% level. The coefficient of the expected inflation is 0.11 and is marginally statistically significant at 10% level with the expected positive sign. This suggests that the response of interest rate to a 1% increase in expected inflation is 0.11%. This finding is consistent with Fisher's (1930) finding of less than the one-for-one relationship between interest rate and expected inflation. In terms of the Taylor rule principle, the observed less than one-for-one relationship between expected inflation and interest rate can be interpreted as monetary policy accommodation.

The one-way causality from expected inflation to short-term interest rate justifies testing for the Fisher hypothesis. This is the novel aspect of this study as it distinguishes it from the prevailing literature, where most of the studies test the Fisher hypothesis directly without empirical evidence on causality between the two variables. These findings, in addition to the correlation coefficient and OLS regression results are consistent with each other and hence provide evidence that the observed finding on the relationship between expected inflation and short-term nominal interest rate is robust.

CONCLUSION AND POLICY IMPLICATIONS

The study examined the seemingly unsystematic relationship between inflation and nominal interest rate by testing the Fisher hypothesis. The Fisher hypothesis predicts a one-for-one relationship between nominal interest rate and expected inflation. The reviewed literature provides empirical evidence of less than one-for-one relationship between expected inflation and interest rate which is evidence of the weak form of the Fisher hypothesis versus the Fisher's (1930) prediction of one-for-one relationship between expected inflation and interest rate. Specifically, the study examined the one-for-one relationship between inflation and nominal interest rate. The main hypothesis is that the weak form of the Fisher hypothesis (i.e. less than one-for-one) holds despite intermittent unsystematic relationship between inflation and nominal interest rate. The study went beyond the empirical evidence in the literature by testing causality between expected inflation and short-term nominal interest rate to ascertain causality between the two variables. Finally, a modified Fisher equation which incorporates an adjustment term as estimated to investigate whether monetary policy involves interest rate adjustment.

The correlation coefficient and OLS estimates show small coefficients (less than unit), implying not only less than a one-for-one relationship but also a weak response of short-term interest rate to expected inflation. This is consistent with the observed weak relationship between expected inflation and short-term nominal interest rate from correlation coefficient estimates. The finding of stationary short-term nominal interest rate and expected inflation rate provides evidence of a dynamic relationship between short-term interest rate and expected inflation but no long run relationship between the two variables. Causality test results show evidence of a one-way causality from expected inflation to short-term nominal interest rate. This allows to conclude that expected inflation predicts short-term interest rate. The economic intuition behind this outcome is that inflation expectations are reflected in the cost of funds.

The statistically significant adjustment term provides evidence for monetary policy adjustment which are intended by the central bank to smooth short-term fluctuations in interest rates. Interest rate smoothing reduces uncertainty to borrowers/investors especially in lending, which in turn motivates economic growth by creating a favorable investment climate. According to the Taylor rule principle, the Bank of Tanzania has been conducting monetary policy accommodation as indicated by not only the less than one-for-one but also a weak relationship between expected inflation and interest rate. This implies that the Bank of Tanzania reactions to inflationary shocks have been less aggressive to fully offset the increase in expected inflationary pressures. These findings corroborate the empirical evidence in the literature.

In terms of policy implications, the observed weak Fisher hypothesis implies that short-term interest rate movements do not perfectly reflect inflation expectations. These findings provide evidence that the overnight interbank cash market rate is a potential operating target under the interest rate-based monetary policy framework. The less than one-for-one finding provides justification and comfort for the Bank of Tanzania to switch from the prevailing monetary targeting regime to the interest rate-based regime as a means of solving the imminent challenge of ineffective monetary policy due to unstable money demand. Instability in money demand is the result of the substantial progress in financial inclusion particularly mobile money services in the country. However, there should be continued measures to strengthen the monetary policy transmission through the interest rate channel.

The finding that inflation predicts short-term interest rates implies that inflation stabilization is important in stabilizing interest rates. Stable interest rates imply confidence on investment decisions and hence a good environment for investors. All these contribute to economic growth. Therefore, the Bank of Tanzania has a crucial role on economic growth and macroeconomic stability through stable interest rates in achieving its primary objective of price stability.

REFERENCES

Balele, N. P. (2013). Monetary Policy in Tanzania: The Monetary Sector, Conduct of Monetary Policy and Exchange Rate Dynamics. Saarbrucken, Deutschland: Lambert Academic Publishing.

Barsky, R. B., & Delong, J. B. (1991). Forecasting Pre-World War I Inflation: The Fisher effect and the Gold Standard. Quarterly Journal of Economics, 106(3), 815-836.

Carmichael, J., & Stebbing, W. P. (1983). Fisher's Paradox and the Theory of Interest. American Economic Review, 73, 619-630.

Chinn, M. D. (1999). A monetary model of the South African Rand. African Finance Journal, 69-91.

Choudhry, A. (1997). Cointegration Analysis of the Inverted Fisher Effect: Evidence from Belgium, France and Germany. Applied Economics Letters, 4, 257-260.

Clarida, R., Gali, J., & Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. Quarterly Journal of Economics, 15, 147-180.

Crowder, W. J., & Hoffman, N. D. L. (1996). The Long-run relationship between nominal interest rates and inflation: The Fisher equation revisited. Journal of Money, Credit and Banking, 28(1), 102-118.

Darby, M. R. (1975). The financial and tax effects of monetary policy on interest rates. Economic Inquiry, 13(2), 266-276.

Edwards, S. (1983). Floating exchange rates in less developed countries: A monetary analysis of Peruvian experience. Journal of Money, Credit and Banking, 18, 73-81.

Fama, E. F.(1975). Short-term Interest rates as predictors of inflation. American Economic Review, 65(3), 269-282.

Feldstein, M. (1976). Inflation, income tax and the rate of interest: A theoretical Analysis. American Economic Review, 66, 809-820.

Fisher, I. (1930). The Theory of Interest. New York: Macmillan.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. Journal of Applied Econometrics, 11, 601-618.

McCallum, B. T. (2002). The use of policy rules in monetary policy analysis, Shadow Open Market Committee.

Mishkin, F. S. (1992). Is the Fisher Effect or Real? A Reexamination of the Relationship between Inflation and Interest Rates. Journal of Monetary Economics, 230, 195-215.

Mishkin, F. S., & Simon, J. (1995). An empirical examination of the Fisher effect in Australia. Economic Record, 71(214), 217-229.

Moazzami, B. (1991). The Fisher equation controversy re-examined. Applied Financial Economics, 1(3), 129-133.

Mundel, R. (1963). Inflation and real interest. Journal of Political Economy, 71, 280-283.

Rotich, H., Kathanje, M., & Maana, I. (2007). Monetary policy reaction function for Kenya. Annual African Econometric Society Conference.

Sanchez-Fung, J. R. (2002). Estimating a monetary policy reaction function for the Dominican Republic. Discussion Paper 02/01.

Shabir, A. (2010). The long-run Fisher effect in developing economies. Studies in Economics and Finance, 27(4).

Tobin, J. (1965). Money and economic growth. Econometrica, 33, 671-684.

Woodward, G. T. (1992). Evidence of the Fisher effect from UK Indexed bonds. Review of Economics and Statistics, 74, 315-320.

APPENDIX

LAG SECTION-ORDER CRITERIA

	Information Criteria		No. lags
Variable	AIC	FPE	(conclusion)
exp_inf	3	3	3
$onibcm$	4	4	4
$repo$	1	1	1
$lombard$	4	4	4

Journal of Applied Economics and Business

OPEN ACCESS
ARTICLE

HEALTH SERVICES COSTS, HOUSEHOLD INCOME AND HEALTH EXPENDITURE IN SOUTH AFRICA

Elvis Munyaradzi Ganyaupfu

Maxima Research Analytics, South Africa

Abstract

The aim of this study was to estimate the effects of the changes in health services costs and household income on health services expenditure by households in South Africa, and assess if disturbances follow a linear autoregressive moving average structure. Using time-series data for the sample period 1976-2017 and controlling for households' disposable income, the ARIMAX method was used for estimation using Stata software. Results show that a rise in health services costs had a statistically significant negative effect on health expenditure by households, whereas household disposable income had a significant positive effect. Based on the sizes of estimates, health expenditure's strong positive reaction to increases in household disposable income and low responsiveness to increases in health services costs provide evidence that households indeed regard health as a necessity as opposed to being a luxury.

Keywords: Health services costs; Household; Consumer; Disposable income.

INTRODUCTION

Health services costs and households' disposable incomes have critical implications for health of populations. Thus, technical discussions around whether health can be regarded as a luxury or necessity have key implications for government's role in the financing of health care (Costa-Font et al., 2012; Tsai, 2014; Ganyaupfu, 2014). Based on the orthodox microeconomic theory of consumer behavior, consumer choices on health care utilization are modeled based on two key approaches (Jack, 1999). The first approach considers health as key part of a set of several commodities over which consumers have defined preferences. Since demand for health care is derived demand in the sense that health care utilization is valued subject to the degree to which it improves health, consumer preferences are for health rather than health care per se. The second approach uses an inter-temporal model of consumption decisions and considers health as a stock variable within the human capital investment framework (Jack, 1999).

Consistent with Grossman (1972), the intertemporal model states that household consumption expenditure on health care occurs not because of the value ascribed to health per se, but rather because of the gains in the stock of health due to health care utilisation. Based on household disposable income levels, consumer preferences for health and health care are made subject to the budget constraint in which consumption choices are bundled into a composite consumption good (c) given by the utility function $u(c, h)$, where h is the level of health, and not the amount of health care consumed (Jack, 1999). When an individual gets sick, additional units of health care (θ) are needed such that θ varies with the level of sickness. Thus, an individual consumer with a health status (θ) and disposable income level (y) faces the budget constraint $c + \theta h \leq y$.

Ceteris paribus, individuals who suffer increased illnesses usually increase their demand for health care services. If the price elasticity of demand for health care services lies between 0 and 1, a proportionate increase in cost of health care can lead to a less than proportionate drop in the preferred amount of health. When the price elasticity of demand for health reveals evidence of health as a normal good, it implies that for an individual with a health status characterized by $\theta = 1$, each additional unit of health would require one additional unit of health care. Household income and price elasticities of demand for health care therefore remain central in determining the allocation of resources devoted for health care services. If household demand for health care appears to be responsive to variations in costs of health services, some user fees should be imposed to induce rationality in consumption of health care services. Similarly, if variations in disposable incomes show a significant direct effect on demand for health care, some form of special subsidization of specific health care services would be desirable (Ganyaupfu, 2019).

LITERATURE AND THEORETICAL FRAMEWORK

Several previous empirical studies that used macro-level and household-level data to analyze the impacts of variations in incomes on health care expenditure present mixed results, where certain studies consider health as a necessity while some regard heath as a luxury. Some studies which include Okunade et al., (2004), van Elk et al., (2009) and de Mello-Sampayo & de Sousa-Vale (2014) regard health as a necessity. From an empirical standpoint, studies which confirm health as a necessity include Culyer (1989), Blomqvist & Carter (1997), Di Matteo (2003), Baltagi & Moscone (2010), Moscone & Tosetti (2010), Duarte (2012), Tsai (2014), Caporale et al., (2015) and Khan et al., (2016). Based on income elasticities greater than one, research studies that conversely find health as a luxury include Kleiman (1974), Newhouse (1977), Gbesemete & Gerdtham (1992), Roberts (1999), Okunade & Murthy (2002), Gerdtham & Lothgren (2002), Freeman (2003), Tsai (2014) and Zeng et al., (2018).

Households are consumers of health services and non-medical goods (Koijen, Philipson & Uhlig (2016), whose preferences over health and consumption follow the utility function:

$$U = E\left[\sum_{t=0}^{\infty}\theta^{t}\frac{\left(c_t^{\lambda}h_t^{1-\lambda}\right)^{1-\eta}}{1-\eta}\right] \tag{1}$$

where c_t represents consumption at time period t, h_t denotes health at time period t, $\lambda\in(0,1)$ measures the trade-off between health and consumption, $0<\frac{1}{\eta}<1$ signifies the inter-temporal elasticity of substitution, and $\theta\in(0,1)$ represents the time discount factor.

For each given household consumer, the level of health at time period t is produced according to the Cobb-Douglas production function specified in equation (2):

$$h_t = \underbrace{\underline{h}\sigma^{t}}_{\text{exogenous health}} + \underbrace{x_t}_{\text{health due to medical treatments}} \tag{2}$$

where $\underline{h}\sigma^{t}$ denotes the household's base level of health, and x_t represents a health input in terms of medical treatments which increase the level of health beyond the base level. The upper bound of health $\left(\bar{h}\sigma^{t}\right)$ gets realized only when medical treatments adopted reach an optimum level.

The manner in which health care service markets operate is determined by the economic choices, decisions and actions taken by household consumers of health services, health care firms, and the government. Given the dualistic nature of health care markets, prices of medical treatments paid by consumers in the private health care services sector are largely driven by the nature and degree of competition in the sector, subject to some form of government intervention. Three major forms in which government intervenes in the health market include (i) regulation of prices which private health care providers charge to health care consumers, (ii) proportional subsidization of some medical treatments to ensure that consumers pay for a fraction $(1-\pi)$ of the market price or cost of medical treatments such that $0<\pi<1$, and (iii) proportional subsidization of research and development undertaken by firms to ensure that such firms privately pay for a fraction $1-\gamma$ of research and development costs such that $0<\gamma<1$ (Koijen et al., 2016).

Given that consumers enjoy subsidies from government for purchase of medical treatments and pay taxes (τ), they maximize utility (u) specified in equation (1) by choosing an optimal bundle of consumption goods (c_t) and medical care (x_t) subject to equation (2), and sequence of the budget constraint given by equation (3):

$$c_t + (1-\pi)\int_0^1 p_{jt}\, x_{jt}\, d_j + \tau_t = \vartheta\sigma \tag{3}$$

where p_{jt} denotes the prices (costs) for medical treatments of type j at time period t, x_{tj} denotes medical treatment of type j at time t, d_j signifies the amount of investment in research and development of type j-knowledge, τ_t denotes lump sum taxes paid by households, ϑ represents households labor income, and σ represents the increase in labor productivity.

The maximization of utility based on parameters specified in equation (2) are subject to consumer preferences for health care utilization, which are made based on several factors which include the decision of whether to seek healthcare service or not and type of facility to visit. Since decisions on types of medical treatments adopted are based on recommendations of medical specialists, final demand for healthcare services is modelled subject to the budget constraint with regards to variations in disposable incomes and prices of the desired health care services. The household aggregate demand for health care services can be derived from the inter-temporal optimization problem defined by the function:

$$\max_{x_t} \frac{\left(c_t^{\lambda}\, h_t^{1-\lambda}\right)^{1-\eta}}{1-\eta} \tag{4}$$

subject to equation (2) and the household budget constraint (equation 3).

METHODOLOGY

Data

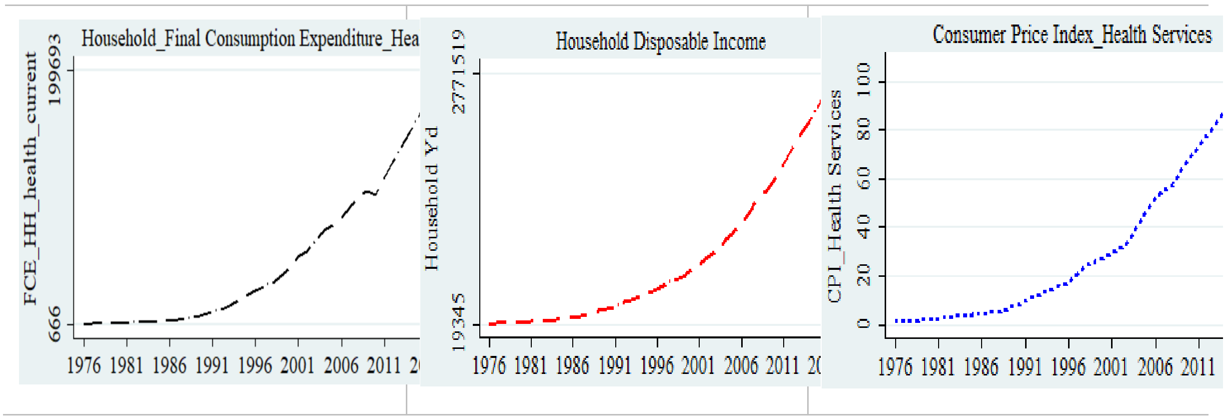


FIG 1. GRAPHICAL EXPOSITION OF DATA SERIES TRENDS

Source: South African Reserve Bank

Annual time-series data for household final consumption expenditure on health at current prices, index of consumer prices of health services (proxy of health services costs) and disposable income of households for the period 1976-2017 was obtained from the South African Reserve Bank (2020) macroeconomic statistical historical online database. Trend plots of the variable used in the econometric estimation are depicted in Figure 1.

Stationarity Tests

Unit root tests were performed against three assumptions that a series should have a constant mean over a given time horizon such that a transitory deviation of the mean reverts to the long-run mean $E(x_t) = E(x_{t-k}) = \mu$, the variance must be constant and finite $E(x_t - \mu)^2 - E(x_{t-k} - \mu) = \sigma^2$, and covariance must depend on the chosen lag length $E(x_t - \mu)(x_{t-k} - \mu) = E(x_{t-h} - \mu)(x_{t-h-k} - \mu) = \sigma_k$.

Since the Augmented Dickey-Fuller (ADF) method performs satisfactorily even when the sample size is fairly small (Dickey & Fuller, 1979), univariate unit root tests were conducted using the ADF criterion which uses an AR(p) process defined by $X_t = \pi + \gamma_1 X_{t-1} + \gamma_2 X_{t-2} + \ldots + \gamma_p X_{t-p} + \varepsilon_t$. The stationarity tests on an AR (p) process thus modelled the regression based on the function:

$$\Delta X_t = \pi + \beta X_{t-1} - \sum_{i=1}^{p-1} \alpha_i \Delta X_{t-i} + \varepsilon_t \quad (5)$$

where ε_t is a pure white noise error term, $\Delta X_{t-i} = X_{i-1} - X_{i-2}$ and p is the class of autoregression. Given the trends exhibited in the series (Figure 1), stationarity tests were performed in levels and first differences using the trend in regression model at 1%, 5% and 10% significance levels.

Estimation Method

The ARIMAX dynamic regression method was used to estimate the impacts of the variations in the index of consumer prices of health services (proxy for health services costs) and household disposable incomes on household final consumption expenditure on health services. The ARMA component was used to determine whether time-dependent disturbances significantly followed a linear autoregressive moving average process, estimated parameters specified in the model:

$$y_t = z_t \varphi + \mu_t \quad (6)$$

$$\mu_t = \varrho \mu_{t-1} + \omega \varepsilon_{t-1} + \varepsilon_t \quad (7)$$

where z_t is a matrix of the AR and MA terms, and φ is a vector of coefficients for z_t parameters.

Operationally, equation (6) is the structural equation, while equation (7) denotes the disturbance ARMA (p, q); where ϱ represents the first order autocorrelation parameter, ω denotes the first order moving average parameter and $\varepsilon_t \approx \text{IID}, N(0,\sigma^2)$ implying that ε_t represents a white noise disturbance. Combining equations (6) and (7), the ARMA (p, q) in the disturbance process was modeled based on the function specified in equation (8):

$$y_t = z\varphi + \rho_1\left(y_{t-1} - z_{t-1}\varphi\right) + \omega_1\varepsilon_{t-1} + \varepsilon_t \quad (8)$$

Inclusion of the stationary independent variables, namely index of consumer prices on health care services (x_1) and household disposable income (x_2), resulted in formulation of an ARIMAX (transfer function) model defined by equation (9):

$$y_t = v(\beta)X_t + \alpha + z\varphi + \mu_t \quad (9)$$

where y_t is the dependent variable (household final consumption expenditure on health), X_t is a matrix of the exogenous variables (the index of consumer prices of health services, and disposable income of households), β is a vector of coefficients of exogenous variables, $v(\beta)X_t$ denotes the transfer function (impulse response function) that permits exogenous variables to influence the dependent variable via a distributed lag, α is a constant term, z_t is a matrix of the AR and MA terms, and φ is a vector of coefficients for z_t parameters and $\mu_t = \varrho\mu_{t-1} + \omega\varepsilon_{t-1} + \varepsilon_t$ denotes the stochastic disturbance of the structural equation that is autonomous of exogenous variables.

RESULTS AND DISCUSSION

Stationarity Tests Results

Stationarity tests statistics (Table 1) of variables at first difference with a trend term in equations rejected the null hypothesis of presence of a unit root in al variables at 5% significance level.

TABLE 1. ADF STATIONARITY TESTS, MODEL: TREND TERM†

Variable	Critical Value			Test Statistic
	α = 1%	α = 5%	α = 10%	τ_t
Household final cons exp_health	-4.242	-3.540	-3.204	2.120
log(Household final cons exp_health)	-4.242	-3.540	-3.204	-0.101
dlog(Household final cons exp_health)	-4.242	-3.540	-3.204	-4.210**
Household disposable income	-4.242	-3.540	-3.204	3.461
log(Household disposable income)	-4.242	-3.540	-3.204	-0.105
dlog(Household disposable income)	-4.251	-3.544	-3.206	-6.365***
Consumer price index_health services	-4.242	-3.540	-3.204	0.478
d(Consumer price index_health services)	-4.251	-3.544	-3.206	-4.203**

[***] (**) * represent significance at [1%], (5%) and 10% levels, respectively

τ_t denotes ADF test statistics with a trend term (t) included in the equation

† the ideal lag order equal to 1 was chosen based on the optimal lag order selection criteria (Appendix 1)

Regression Estimates

TABLE 2. ARIMAX ESTIMATES

Sample: 1977 – 2017						
				No. of obs	=	41
				Wald ch2(4)	=	88.28
Log pseudolikelihood = 71.335				Prob > ch2	=	0.000

dlog(hh_final cons exp_health)	Coeff.	Semirobust Std Err	z	P >|z|	[95% Conf. Interval]	
d(CPI_health services)	-0.010	0.004	-2.76	0.006	-0.017	-0.003
dlog(hh_disposable income)	0.450	0.203	2.22	0.026	0.053	0.847
_cons	0.109	0.021	5.14	0.000	0.067	0.150
ARMA						
ar(L1)	0.752	0.223	3.37	0.001	0.314	1.189
ma(L1)	-0.513	0.210	-2.44	0.015	-0.924	-0.101
/sigma	0.042	0.005	8.02	0.000	0.032	0.053

The ARIMAX regression results (Table 2) reveal that both health service costs (index of consumer prices of health services) and households' disposable income had statistically significant impacts on households' final consumption expenditure on health services over the period 1977 to 2017. The significant (t-statistic = -2.76; $p < 0.05$) cost of health services elasticity (coefficient = -0.01) far less than one confirm that household expenditure on health services was insensitive to changes in costs of health care services over the period 1977 to 2017. In addition, the significant (t-statistic = 2.22; $p < 0.05$) households' disposable income elasticity of health expenditure less than one (coefficient = 0.450) confirm that households indeed regarded health services as a necessity over the period under review. The log differenced series of households' health expenditure shown by AR (1) was highly correlated at a level of 0.752, while the innovations shown by MA (1) had a significant (z-statistic = -2.44; $p < 0.05$) negative impact (coefficient = -0.513) in the succeeding period, with a standard deviation of the white noise disturbance of 0.042.

Diagnostic Tests of ARIMAX Estimates

The eigenvalues of autoregressive (AR) and moving average (MA) parameters were examined to determine whether they satisfied the stability condition and invertibility condition; respectively.

TABLE 5. STABILITY AND INVERTIBILITY OF ARIMAX ESTIMATES

Parameters	Eigenvalue	Modulus
AR parameters	0.752	0.752
MA parameters	0.513	0.513

The eigenvalues for AR parameters and MA parameters reveal that ARIMAX estimates satisfy the stability and invertibility conditions; respectively. The analytic plot for the inverse roots of ARMA polynomials (Appendix 2) with AR roots and MA roots that lie inside the unit circle confirm that the ARIMAX estimates indeed satisfy the stability and invertibility conditions.

CONCLUDING REMARKS

Enjoyment of the highest standard of health remains as one of the fundamental rights of every human being, paramount to fulfilment of other basic human rights. The relatively substantial contribution to total health care spending in South Africa largely comes from the private sector in form direct out-of-pocket payments, voluntary payments and medical schemes. Results in this study reveal that households regard health as a necessity, while they are inelastic to changes in costs of health services. In order to reduce the burden of bearing high costs for accessing health services and medical treatments, government should consider subsidization of health services and also prioritize financing of health investment programs towards attainment of the goal of universal health coverage.

Given the empirical evidence that health is regarded as a necessity in the country, government should continuously develop and scale-up implementation of health-access initiatives and health improvement interventions that prevent certain illnesses, or provide for early treatment to avoid high downstream health care costs associated with illnesses and subsequent complications. Such interventions can reduce health-care burdens on households and government, thus saving public funds for capital investments on other quality-life enhancing streams, economic growth and human development streams such as infrastructure and education.

APPENDIX

APPENDIX 1

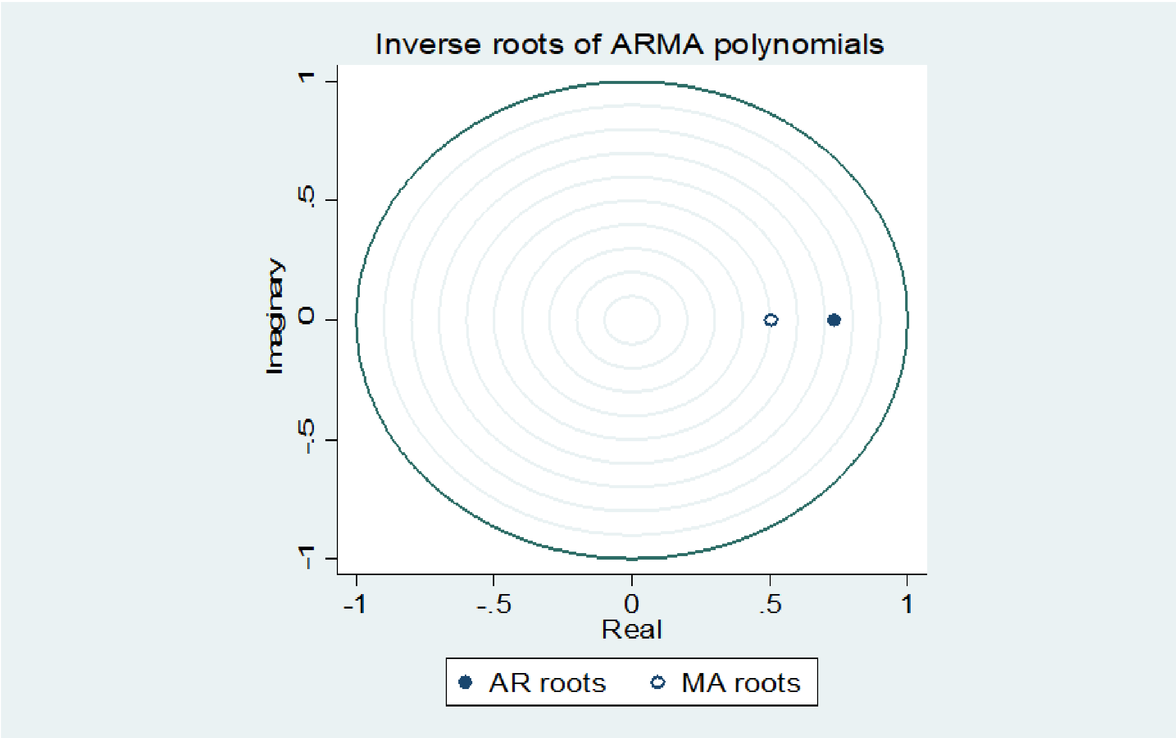
VAR LAG ORDER SELECTION

Lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-1084.19	-			8.1e+19	54.3596	54.4054	54.4863
1	-859.296	449.79*	9	0.000	1.7e+15*	43.5648*	43.748*	44.0714*
2	-852.446	1.402	9	0.133	1.9e+15	43.6723	43.9929	44.5589

* Indicates the optimal lag order selected by the respective criterion.

APPENDIX 2

ANALYTIC PLOT FOR THE INVERSE ROOTS OF ARMA POLYNOMIALS



REFERENCES

Baltagi, B. H. & Moscone, F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. Discussion Paper 4851, IZA.

Blomqvist, G., & Carter, R. A. L. (1997). Is health care really a luxury? Journal of Health Economics, 19, 207-229.

Caporale, G. M., Cunado, J., Gil-Alana, L. A., & Gupta, R. (2015). The Relationship between Healthcare Expenditure and Disposable Personal Income in the United

States: A Fractional Integration and Cointegration Analysis. Working Paper: 2015-32. Department of Economics, University of Pretoria.

Costa-Font, J., McGuire, A., & Stanley, T. (2012). Publication Selection in Health Policy Research: The Winner's Curse Hypothesis. Working Paper, No: 25/2012. LSE Health. The London School of Economics and Political Science.

Culyer, A. J. (1989). Cost containment in Europe. Health Care Financing Review, 21-32.

De Mello-Sampayo, F., & de Sousa-Vale, S. (2014). Financing Health Care Expenditure in the OECD Countries: Evidence from a Heterogeneous, Cross-Sectionally Dependent Panel. Panoeconomicus, 61, 207-225.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association, 74, 437-431.

Di Matteo, L. (2003). The income elasticity of health care spending: A comparison of parametric and nonparametric approaches. The European Journal of Health Economics, 4, 20-29.

Duarte, F. (2012). Price elasticity of expenditure across health care services. Journal of Health Economics, 31, 824-841.

Freeman, D.G. (2003). Is health care a necessity or a luxury? Pooled estimates of income elasticity from US state-level data. Health Economics, 8, 485-496.

Ganyaupfu, E. M. (2014). Estimating the Relative Impacts of Health and Education on Economic Development in Southern Africa. Asian Journal of Economic Modelling, 2(2), 85-92.

Ganyaupfu, E. M. (2019). Health Care Financing and Expenditure Dynamics in Southern Africa. Global Journal of Economics and Finance, 3(2), 1-11.

Gbesemete, K.P., & Gerdtham, U.G. (1992). Determinants of healthcare expenditure in Africa. A cross-sectional study. World Development, 20, 303-308.

Gerdtham, U.G., & Lothgren, D.G. (2002). New panel results on cointegration of international health expenditure and GDP. Applied Economics, 34, 1679-1686.

Grossman, M. (1972). On the concept of health capital and the demand for health. Journal of Political Economy, 80, 223-255.

Jack, W. (1999). Principles of Health Economics for Developing Countries. WBI Development Studies. World Bank Institute.

Khan, H.N., Razali, R.B., & Shafie, A.B. (2016). Modelling Determinants of Health Expenditures in Malaysia: Evidence from Time Series Analysis. Frontiers in Pharmacology, 7, 1-7.

Kleiman, E. (1974). The determinants of national outlay on health. In Perlman M, editor. The Economics of Health and Medical Care. London: Macmillan.

Koijen, R. S. J., Philipson, T. J. & Uhlig, H. (2016). Financial Health Economics. Econometrica, 84, 195-242.

Moscone, F., & Tosetti, E. (2010). Health expenditure and income in the United States. Health Economics. 19, 1385-1403.

Newhouse, J. P. (1977). Medical care expenditure: a cross-sectional survey. Journal of Human Resources, 12, 115-125.

Okunade, A. A., Karakus, M. C., & Okeke, C. (2004). Determinants of Health Expenditure Growth of the OECD Countries: Jackknife Resampling Plan Estimates. Health Care Management Science, 7, 173-183.

Okunade, A. A., & Murthy, V. N. R. (2002). Technology as a 'major driver' of health care costs: a cointegration analysis of the newhouse conjecture. Journal of Health Economics, 21, 147-159.

Roberts, J. (1999). Sensitivity of elasticity estimates for OECD health care spending: analysis of a dynamic heterogeneous data field. Health Economics, 8, 459-472.

South African Reserve Bank. (2020). Historic Macroeconomic Timeseries Information.

Tsai, Y. (2014). Is Health Care an Individual Necessity? Evidence from Social Security Notch. Centers for Disease Control and Prevention.

Van Elk, R., Mot, E., & Franses, P. H. (2009). Modelling health care expenditures. Discussion Paper 121, CPB.

Zeng, Z., Qi, Y, & Ge, M. (2018). Is health care a necessity or a luxury? Evidence from urban China. Applied Economics Letters, 25, 1204-1207.

Journal of Applied Economics and Business

OPEN ACCESS ARTICLE

SCIENTIFIC ABSORPTIVE CAPACITY: THE ROLE OF CORPORATE SCIENTISTS AS MICRO-FOUNDATIONS OF INNOVATION CAPABILITY

Ramin Vandaie

Abstract

Purpose. The purpose of this study is to explain the connection between firms' investment in basic scientific research and their innovation performance. Basic research and its outcomes generally belong to the public domain and are accessible to all interested parties. Yet, private firms' engagement in it, as we argue, helps create absorptive capacity needed to access and benefit from cutting edge knowledge created outside their organizational boundaries.

Design/methodology/approach. This is a conceptual paper dedicated to the development of a contingent theoretical framework to link firm investments in basic scientific research to their ability for form and benefit from external knowledge access collaborations. The developed theoretical model draws from extant empirical literature on the role of basic research in firm innovation including anecdotal observations that hint at managerial intentions behind their encouragement of corporate scientists to engage in basic research, particularly in collaboration with external scientists.

Findings. The model suggests two sets of mechanisms as part of a contingent theory of the role of scientific absorptive capacity in driving external knowledge sourcing of firms. Whereas the first set of contingencies is based on the nature of the firm's scientific research activities, the second set focuses on other internal firm processes that may suppress or substitute the effect of the resultant absorptive capacity.

Research limitations/implications. As a conceptual paper, this study is limited in that it relies on insights from prior literature and the verification of its proposed relationships would require future empirical investigations. Yet, the theoretical insights developed here could pave the way for more fruitful future inquiries into the nature of basic scientific research in the private sector and enhance our understanding of how science fits into firms' overall innovation strategy.

Originality/value. As one of the first studies to elaborate on the concept of scientific absorptive capacity, this paper not only contributes to the burgeoning literature on the process of absorptive capacity development in firms, it also offers insights for practitioners regarding the effective management of the research and development process inside their firms.

Keywords: Scientific research; Knowledge accessing; Absorptive capacity; Firm innovation.

INTRODUCTION

Research has long speculated about the how basic scientific research contributes to firm innovation (Gambardella, 1992; Cockburn & Henderson, 1998; Zucker et al., 1998; Durand et al., 2008). Science has been generally identified in this research as a driver of a significant subset of innovations in various industries (Mansfield, 1991 and 1997). Mansfield (1991), for instance, found that about one-tenth of all product and process innovations introduced between 1975-1985 in a wide range of industries including pharmaceuticals, chemicals, metals, oil, instruments, and information processing could not have been developed in the absence of recent basic scientific research. A common theme among these various lines of research is that firms do not consider science as a mere external entity existing in the public domain (Stern, 2004; Huang & Murray, 2009). To the contrary, firms, especially in technology-driven industries, have been shown to actively pursue science by making significant investments in its creation and development (Gambardella, 1992; Gittelman & Kogut, 2003).

The observation that for-profit firms make significant investments in basic scientific research (Rosenberg, 1990), has presented a puzzle to scholars given the public good nature of its outcomes that generate no tangible rents for the firm in the traditional sense. Relatedly, this literature has also noted that significant heterogeneity exists among firms in their ability to benefit from knowledge created outside their boundaries (Cohen & Levinthal, 1990; Gambardella, 1992). In particular, while a great deal of external knowledge, including public science, appears to be equally and inexpensively accessible to all firms, only a subset of firms manage to harness its benefits in their own internal innovation processes. A joint consideration of these two observations has offered a solution to the puzzle of private firm's engagement in public science: private firms invest in basic science in order to develop the 'absorptive capacity' to understand and utilize external knowledge (Cohen & Levinthal, 1990; Rosenberg, 1990; Gambardella, 1992). Absorptive capacity allows firms to monitor the flow of knowledge outside their boundaries, internalize relevant knowledge, and exploit it in their internal innovation processes (Rosenberg, 1990).

However, while the general benefits of investing in basic science for facilitating firms' access to external knowledge are discussed at length in this literature, there is surprisingly little discussion on the challenges faced by firms in developing and deploying the type of absorptive capacity that results from basic science investments, also known as scientific absorptive capacity. Particularly, on the development side, the question of how to effectively allocate such internal investments toward maximizing the resultant absorptive capacity remains unanswered. That is, we do not know which types of research projects tend to have higher contributions to the firm's absorptive capacity and hence, merit higher shares of the firm's basic research budget. On the deployment side, literature also seems to have assumed that once developed, absorptive capacity is automatically incorporated into the firm's innovation strategy to its fullest potential. As such, we know little about internal mechanisms that might potentially interfere with the deployment of absorptive capacity and suppress its role in enabling the firm's access to external knowledge.

We set out to close these gaps in our understanding of how firms develop and deploy absorptive capacity. On the one hand, we consider the nature of the research projects that represent the ingredients of the firm's scientific absorptive capacity to determine the characteristics that set the projects apart in terms of their eventual contribution. On the other hand, we consider a set of internal firm mechanisms that could potentially eclipse the role of scientific absorptive capacity and suppress its effect in facilitating external knowledge access. In particular, we formulate a contingent theory of the effect of basic science projects, and their resultant absorptive capacity, on the formation of new knowledge sourcing collaborations (e.g. R&D alliances), encompassing two sets of contingencies that correspond to our study's dual contributions outlined above. The first set of contingencies – *isolation of firm scientists*, *arm's length co-authorships*, and *organizational dispersion of scientific projects* – probes the nature of the research projects that contribute to the firm's absorptive capacity. The second set – *patenting productivity of the firm and R&D alliance experience* – reflects internal firm capabilities that arguably hold more tangible associations with the process of new collaboration formation relative to scientific absorptive capacity and hence, overshadow its fundamental role within the process. Finally, we consider the role of scientific absorptive capacity at the time of forming new collaborations on the long-term benefits of those collaborations for the firm's innovative output.

A THEORY OF SCIENTIFIC ABSORPTIVE CAPACITY

METHODOLOGY

Given the conceptual nature of this study, we employed an extensive literature review methodology to identify all key studies in various streams of literature that have dealt with the topic of basic scientific research in the private sector. After several rounds of refinement, we ended up with a collection of scholarly work that directly addressed the role of scientific research in firm innovation and the potential contingencies surrounding its impact. The insights from these studies were distilled into a novel theoretical framework that proposes a contingent relationship between firm investments in basic scientific research and the ability to form external collaborations to benefit from knowledge residing outside firm boundaries. Figure 1 summarizes our proposed theoretical framework which is explained in detail in the following sections.

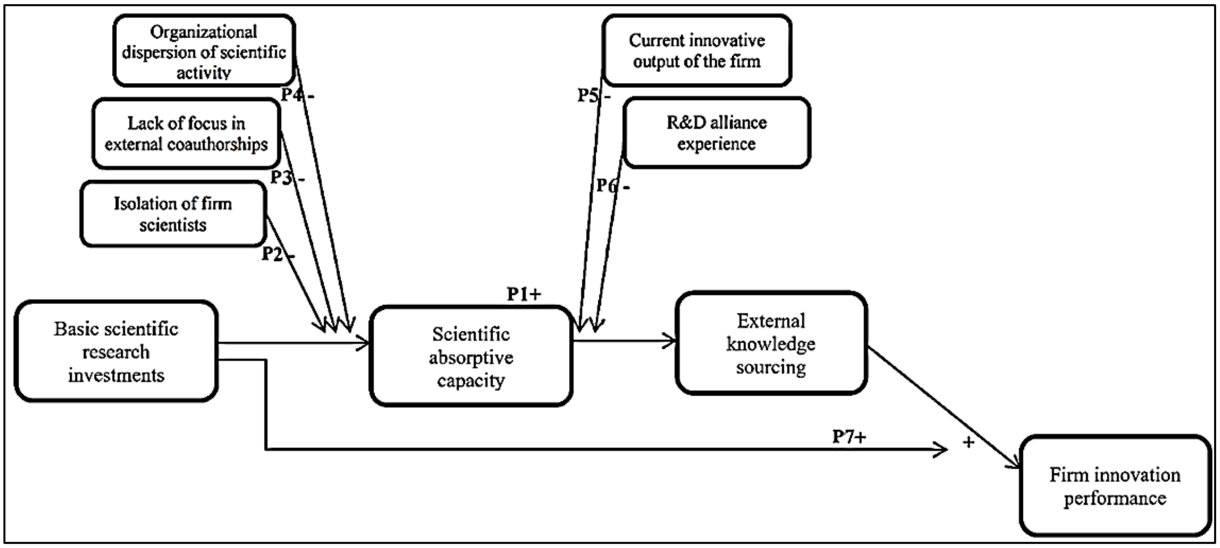


FIG 1. THE PROPOSED THEORETICAL FRAMEWORK OF THE STUDY

SCIENTIFIC ABSORPTIVE CAPACITY AND KNOWLEDGE SOURCING

A long tradition of research (e.g. Rosenberg, 1990; Gambardella, 1992; Zucker et al., 2002) has established that engaging in in-house basic scientific research allows firms to stay abreast of and tap into external flows of scientific and technological knowledge to boost innovation within the firm. Such capacity based on in-house scientific research to benefit from external knowledge is, in essence, an element of absorptive capacity. Cohen and Levinthal's (1990) ground-breaking article introduced the notion of absorptive capacity to strategy and organizational research as a firm's ability to evaluate, assimilate, and apply external information and knowledge in its innovation processes. According to Cohen and Levinthal (1990), absorptive capacity provides a ready explanation for firms' investment in basic research despite the fact that its findings inevitably spill out into the public domain in the form of scientific publications. Specifically, they suggested that firms' true goal in conducting basic research is not the mere results of the research itself but also the creation and maintenance of the ability to tap into knowledge developed outside the firm. In other words, in-house basic research can be thought of as "…broadening the firm's knowledge base to create critical overlap with new knowledge and providing it with a deeper understanding that is useful for exploiting new technical developments that build on rapidly advancing science and technology." (Cohen & Levinthal, 1990: 148).

Since the notion of absorptive capacity has grown to include several dimensions of a firm's internal capabilities, to maintain our focus on the role of basic science in firm's innovation strategies we adopt the term 'scientific absorptive capacity' (Schmidt, 2005; Roach, 2009) to refer to the capacity provided based on conducting in-house scientific research. Our baseline proposition captures the relationship between scientific absorptive capacity and external innovative knowledge sourcing by the firm. Specifically, we argue that a firm's scientific absorptive capacity reflected in its in-house science productivity impacts its willingness and ability to form future knowledge access collaborations. Particularly, scientific absorptive capacity enables the firm to assess a larger group of potential collaborating partners, leading to the formation and execution of a greater number of such collaborations (Arora & Gambardella, 1994). Scientific absorptive capacity also alleviates partners' appropriability concerns that inevitably overshadow any joint knowledge creation and access (Gulati & Singh, 1998; Oxley and Sampson, 2004). That is, the deeper knowledge of firm scientists about the knowledge areas involved in the collaboration will enable the firm to identify and prevent partner opportunism should it arise during the course of the collaborative effort (Gulati & Singh, 1998). Building on this line of reasoning, we propose that scientific absorptive capacity will impact the firm's extent of future external knowledge accessing collaborations. Therefore,

> P1: Scientific absorptive capacity will positively impact the extent of the firm's external knowledge access collaborations.

CONTINGENCIES BASED ON CHARACTERISTICS OF SCIENTIFIC PROJECTS

Isolation of firm scientists

Our first set of contingencies is based on the features of the scientific projects that underlie the firm's scientific absorptive capacity. We begin by considering the contingency of the isolation of firm scientists in their projects from the broader scientific community. The public good nature of basic science implies that scientific advances result from the collective efforts of the broader scientific community. In other words, the locus of major developments in basic science is most likely to fall in the public domain consisting of all public and private institutions that engage in basic scientific research. Collaboration of firm scientists with external coauthors represents the main mechanism through which the firm's internal research program stays connected to scientific advances in the field. Research has shown that connectedness to the external scientific community not only enhances firms' ability to recognize, evaluate, and utilize scientific developments, it also increases research productivity inside the firm (Cockburn & Henderson, 1998).

The majority of external collaborations by firm scientists are with academic researchers who act as contact points helping corporate researchers to keep their knowledge updated with scientific advancements in academic research labs and other publicly-funded institutions that are dedicated to basic research. These collaborations are crucial to firm scientists as they can gain direct exposure to new knowledge at the boundaries of their respective fields. Particularly, knowledge of ground-breaking scientific discoveries is often tacit and its links to current codified knowledge base tends to be less clear. Hence, by collaborating with the star scientists, corporate researchers learn enough to be able to transfer that knowledge for internal use inside the firm.

To effectively benefit from external scientific knowledge, it is not enough for firms to hire top researchers and inspire them to actively engage in basic research. It is also important for the corporate researchers to actively engage in external collaborations with those in the academia and the public sector (Cockburn & Henderson, 1998). Zucker et al., (2002), for instance, argued that academic collaborations can be utilized by corporate researchers to capture and internalize the complex knowledge components that form the foundations for novel discoveries in university labs. The effectiveness of such collaborations are higher when they connect firm scientists to 'star' scientists most of whom work at top universities. In earlier research, Zucker et al., (1998) found that both start-up firms and existing firms entering a new technology area tend to collocate with the top star scientists of the field.

Overall, connectedness to the broader scientific community through external co-authorships implies that firm scientists are involved in the generation of cutting-edge knowledge within the scientific discipline and have first-hand exposure to its underlying components. In contrast, when firm scientists are isolated from the broader scientific community, the knowledge represented in their research output is likely to be distant from recent advances in the field, and hence, result in the diluting of the effect of the firm's scientific absorptive capacity in driving external knowledge access efforts of the firm. Therefore,

> P2: The positive effect of scientific absorptive capacity on the extent of the firm's external knowledge access collaborations is weaker when the degree of scientist isolation is high.

Lack of focus in external collaborations

By further examining the nature of the external collaborations of firm scientists, the next contingency complements the preceding arguments regarding the importance of external connectedness for firm scientists. Specifically, Proposition 3 distinguishes between external collaborations that truly represent hands-on involvement of firm scientists in joint research projects with colleagues outside the firm, and those resembling arm's length relationships where firm scientists only receive credit for their general attachment to the project without performing any serious collaborative work. The basic argument here is that publications where a large number of scientists representing a myriad of public and private organizations are listed as coauthors most often represent instances of the latter type of external collaborations (Cronin, 2001).

Numerous studies in various disciplines have reported a growing trend in the number of authors on a typical research publication – e.g. Cronin (2001) in biomedicine, Slafer (2005) in crop science, Gibelman & Gelman (2000) in social work, and Englebrecht et al., (2008) in accounting research. The shift from solo-authored to multiple-authored publications has been generally regarded as a justified reaction to historical changes in the nature of research and publishing such as the shifting patterns of research funding, increased specialization, demands for higher quality and precision of scientific inquiry, the growing professionalism in academia, the need to train apprentices, and the demand for cross-fertilizing across disciplines (Katz & Martin, 1997). However, systematic analyses of scientific publications and their patterns of co-authorship have also revealed a creeping trend toward overcrowded lists of coauthors and the organizations and institutions represented by them on a growing subset of coauthored publications. Gelman and Gibelman (1999), for instance, reported that among the papers published in 4,000 journals indexed by the Institute for Scientific Information, the number of papers with 50 or more coauthors grew from 49 in 1981 to 407 in 1994. Such overcrowded lists of coauthors and institutions have been generally associated with denigrating the scientific research process without offering any real added value per every additional listing beyond what constitutes a convincing size of the actual team of researchers and their supporting institutions (Woods et al., 2010).

Thus, we argue that the presence of arm's length-type coauthored publications in the firm scientists' output will likely be a sign of the lack of engagement of firm scientists in rigorous joint efforts with the leading minds behind major discoveries in the discipline. Therefore,

> P3: The positive effect of scientific absorptive capacity on the extent of the firm's external knowledge access collaborations is weaker when firm scientists engage in many external co-authorships of the arm's length type.

Organizational dispersion of scientific activity

As a third contingency, we examine the implications of research activities that are dispersed across the corporate structure and are carried out by scientists employed at the subsidiaries. Particularly, we argue that basic research projects that are not concentrated in specific organizational units and are scattered across various organizational divisions tend to generate a type of knowledge base that requires significant mobilization costs in order to become beneficial for the firm's internal innovation process (Teece, 1977). In other words, the dispersion of basic research projects is likely to introduce added integration costs to the process of mobilizing and employing the firm's scientific absorptive capacity.

We know from past research that integrating the knowledge created by distributed research efforts across different corporate divisions presents significant challenges. In one such study, Singh (2008) explored how the geographic dispersion of a firm's R&D activities impacts the quality of its innovative output, and found that such distribution lowers the value of the resulting innovations. The higher transfer costs of knowledge created by distributed research efforts is partly driven by a phenomenon known as knowledge 'stickiness' (Kogut & Zander, 1993; Szulanski, 1996). The concept of knowledge stickiness builds on the notion that, besides transfer agents and transfer media, transfer costs are also driven by the characteristics of the knowledge itself. Szulanskie (1996) found that contrary to the previously-accepted wisdom, the major impediments to the transfer of knowledge are not rooted in motivational factors, but are due to the characteristics such as the causal ambiguity of the knowledge itself. Organizations can deal with knowledge stickiness by investing in proper organizational structures and cultures that inspire employees to engage in tacit knowledge transfer across the entire organizations rather than simply focusing inside their work teams (Osterloh & Frey, 2000). Organizations would also need to speed knowledge transfer by attempting to codify their tacit knowledge as much as possible and in doing so, risk imitation of their innovative ideas by competitors (Zander & Kogut, 1995). Thus, we expect that the impact of the resulting scientific absorptive capacity will be diluted by the integration and deployment costs arising from the dispersion of scientific projects across the organizational structure. Therefore,

> P4: The positive effect of scientific absorptive capacity on the extent of the firm's external knowledge access collaborations is weaker when the organizational dispersion of scientific projects is high.

CONTINGENCIES BASED ON INTERNAL FIRM PROCESSES

Current innovative output of the firm

Our second set of contingencies is based on internal firm processes that suppress the effect of scientific absorptive capacity by reducing the firm's sensitivity to its availability in its decisions regarding the formation of new knowledge access collaborations. The underlying logic here is that various internal capabilities may be considered as drivers of the same organizational outcome even though the nature of their impact as well as their long-term consequences may be inherently different. By overlooking such differences, managers are prone to presumptions regarding their equifinality and substitutability. Prior literature has shown that firms do indeed substitute certain capabilities for each other. Makadok (2001), for instance, developed a model that predicted that the two rent creation mechanisms of resource picking and capability building, while complementary in some circumstances, often tend to substitute each other. Relatedly, Rothaermel and Hess (2007) examined the individual-, firm-, and network-level antecedents to firm innovation and found that these antecedents sometimes compensate for each other's effects on a firm's innovative output.

In Proposition 5 we examine the contingency of the firm's patenting productivity as an important internal capability and argue that firms with a high patenting productivity will be less sensitive to the availability of scientific absorptive capacity in their external knowledge access efforts. This is primarily due to the upstream position of scientific absorptive capacity in the innovation value chain as opposed to the downstream position of the firm's patenting productivity. That is, whereas scientific publications contain knowledge pertaining to potential future innovations that are still in their nascent stage, patents represent applied inventions with more tangible connections to the firm's current innovations and new product development processes. In fact, science is so far upstream relative to patented inventions that its role is often portrayed as a map in technological search and patenting activity. Fleming and Sorenson (2004), in particular, argued that scientific knowledge functions as a map by providing a means of predicting untried combinations of technological components, identifying fruitless directions before the inventors attempt them, and motivating inventors in the face of repeated failure to continue the search.

Hence, the downstream position of patents in the innovation value chain implies that firms are more likely to assess the vigor of their internal innovation activities as well as their capacity to undertake more collaborations based on their patenting productivity than their scientific absorptive capacity. As such, we expect that under situations of high patenting productivity firms will become increasingly less sensitive to the availability of scientific absorptive capacity. Therefore,

> P5: The positive effect of scientific absorptive capacity on the extent of the firm's external knowledge access collaborations is weaker when the firm has a high patenting productivity.

R&D alliance experience

Next, we consider the contingency of the firm's R&D alliance experience as a particularly relevant internal firm capability and argue that more experienced firms are likely to be less sensitive to scientific absorptive capacity in their decisions to form new collaborations. Research has shown that, over repeated collaborations, organizations learn the nuances involved in managing alliances and deploy such learning in future alliances to ensure their success despite potential shortcomings that may strain less experienced partners (Anand & Khanna, 2000; Kale & Singh, 2007; Gulati et al., 2009). Experience firms, on the other hand, have a good grip on the challenges of collaborative knowledge creation and transfer which allows them to better navigate the challenges that often cripple less experienced partners (Simonin, 1999; Larsson et al., 1998). One such major challenge is the possibility of spiraling into a 'learning race' that typically characterizes knowledge-focused alliances where each partner feels the pressure to learn as much as possible while minimizing the leakage of its proprietary knowledge in the partnership (Khanna et al., 1998; Hamel, 1991). Therefore, the ability to maintain the balance between learning new knowledge and protecting the firm's proprietary knowledge represents a valuable capability to every alliance partner (Kale et al., 2000). Alliance experience, especially if captured and internalized through a dedicated alliance function (Kale et al., 2002), contributes to the firm's capability to overcome the challenges of learning and knowledge transfer in R&D alliances.

R&D alliance experience accrues to the firm as it engages in repeated R&D collaborations with various partners and continues to store and retrieve the inferences drawn from every instance of partnership. The progress along the alliance learning curve increasingly enables the firm to identify the effective interfirm routines for managing the process of knowledge creation and transfer within the alliance (Lane & Lubatkin, 1998). The growing alliance management capability, we argue, will likely boost the firm's confidence regarding its ability to face the challenges in potential future knowledge access collaborations even in the absence of sufficient scientific absorptive capacity. Therefore,

> P6: The positive effect of scientific absorptive capacity on the extent of the firm's external knowledge access collaborations is weaker when the firm's R&D alliance experience is high.

CONSEQUENCES FOR FIRM INNOVATIVE OUTPUT

Finally, we turn to the consequences for the innovative output of the firm of building external knowledge access collaborations on a strong foundation of absorptive capacity. Our basic argument here is that such a foundation not only increases the benefits of such collaborations for the focal firms, it also minimizes their potential hazards including the internalization of technology components that are incompatible with the firm's existing knowledge and technology base. Avoiding such incompatible components minimizes the challenges to the firm's internal innovation processes and prevents the subsequent weakening of the firm's innovative capabilities. The arguments in this section build on a modular representation of firms' internal innovation systems that particularly suits the firms in industries with a fast pace of technological change (Sanchez & Mahoney, 1996; Pil and Cohen, 2006). Modular innovation systems encompass various loosely-coupled components of knowledge and technology whereby the loose coupling between components reduces the costs and difficulty of adaptation and increases the firm's speed and flexibility in responding to rapid technological changes (Ethiraj & Levinthal, 2004). The flexibility of the modular architecture is due to the fact that innovative products can be developed by substituting different modular components into the product architecture without the need to redesign other components. In other words, the 'mixing and matching' capacity of the modular system allows firms to develop a potentially large number of innovations by recombining new or existing components of knowledge and technology (Henderson & Clarck, 1990; Sanchez & Mahoney, 1996).

The ability to recombine technological knowledge in novel ways to explore potential innovations that alter and advance the firm's current technological trajectories largely determines the innovative performance of a firm (Kogut & Zander, 1992; Tzabbar, 2009). Dividing organizational competence into component and architectural, Henderson and Cockburn (1994) argued that architectural competence, composed of the organizational control systems and the dominant values, allows a firm to exploit its component competence by integrating them in new and flexible ways. Henderson and Cockburn (1994) also suggested that in the context of the pharmaceutical industry the ability to access external knowledge and the ability to flexibly recombine and integrate knowledge across the various disciplinary and therapeutic class boundaries represent two forms of architectural competence that determine the innovative performance of the firm.

Successful exploitation of externally-sourced components of knowledge and technology requires a certain level of complementarity and compatibility to exist between those components and the firm's existing social and technological structures (Teece, 1986). That is, in order for the process of mixing and matching of components in a modular innovation system to work efficiently, every new knowledge and technology component must have a certain level of compatibility with the existing structure of the system. Compatibility implies that the new and existing components can be connected using a common interface without needing any extraordinary translation and interpretation efforts. Since firms typically search for technological solutions that fall within the boundaries of their existing knowledge base (March and Simon, 1958), knowledge components that substantially deviate from this existing base significantly challenge the firm and its members as they attempt to comprehend and recombine them in crafting future innovations.

External knowledge access collaborations typically expose firms to unfamiliar technology landscapes without a map to guide the firm's search for compatible components (March, 1991; Fleming & Sorenson, 2003). If, due to a low level of absorptive capacity, the firm also lacks a clear understanding of the nature of knowledge and technology components involved in the joint R&D projects, the search for new knowledge within the framework of such collaborations will be a blind search on a rugged technological landscape. Such a blind search is likely to lead the firm to try to internalize any components that might appear to be relevant to the firm's internal knowledge base and fail to verify their actual compatibility. Incompatible components that are introduced into the firm's internal innovation system through external knowledge access collaborations present significant challenges to the firm in its attempts to incorporate them into ongoing recombination efforts and match them with existing knowledge and technology components. Particularly, attempting and discarding a multitude of potential configurations and combinations is likely to introduce a creeping element of inefficiency into the firm's innovation processes leaving a negative impact on the firm's innovative output. Such a drop in the innovative output is likely to result from the gradual alterations of the firm's search routines to accommodate the inclusion of incompatible external components in new configurations.

Moreover, every new component of knowledge tends to deform and expand the search space of the firm's innovation system by suggesting new competing hypotheses and presenting previously unknown discovery paths (Orsenigo et al., 2001; Fleming & Sorenson, 2004). However, when new components are incompatible, such deformation of the search space is likely to throw the firm's innovation processes off their current functioning paths requiring extra effort over time to restructure the search space and restore the efficiency of the innovation processes. As time and other resources are invested toward reformation and restoration of the firm's innovation system, the inefficiency induced by incompatible knowledge components is likely to build up leading to a drop in the system's output over time.

Conversely, building external knowledge access collaborations on a strong foundation of absorptive capacity enables the firm to more effectively screen for compatible components of knowledge and technology while navigating the novel technology landscape. Absorptive capacity guides the firm's search on the less familiar technology landscapes that dominate the inter-firm search space and allows the firm to pick the most compatible from among the various components that might initially appear as attractive additions to the firm's existing technology base (Fleming & Sorenson, 2004). Therefore, only those components will be internalized that imply a minimal impact on the efficiency of the firm's recombination efforts to introduce new innovations. Also, such compatible components will likely only deform the firm's innovation search space in ways that minimally impact its ongoing innovation processes. Given the gradual nature of the subsequent inefficiencies introduced into the firm's innovation system in external knowledge access collaborations not supported with sufficient absorptive capacity, we expect that the benefits of absorptive capacity in offsetting those inefficiencies to also emerge in the long run. Therefore,

> P7: Science productivity at the time of forming new external knowledge access collaborations will enhance the benefits of those collaborations for the firm's long-term innovative output.

DISCUSSION AND CONCLUSION

This study joins a lively stream of research that aims to explain how science relates to firms' innovation strategies (Gambardella, 1992; Zucker et al., 1998; Gittelman & Kogut, 2003; Tzabbar, 2009). Firms invest in basic research to develop absorptive capacity that allows them to recognize, access, and exploit knowledge and technology developed outside their boundaries (Cohen & Levinthal, 1990; Gambardella, 1992; Roach, 2009). Particularly, absorptive capacity functions as a driver of the firm's engagement in external knowledge access collaborations (Arora & Gambardella, 1994). We theorized about a number of key contingencies surrounding the link between firms' science productivity and their extent of engagement in external knowledge access collaborations (Arora & Gambardella, 1994). The first set of moderators– i.e. isolation of firm scientists, arm's length co-authorships, and organizational dispersion of scientific projects – corresponded to factors that dilute the effect of scientific absorptive capacity. The second set – i.e. the firm's patenting productivity and R&D alliance experience – corresponded to internal firm capabilities that lowered the firm's sensitivity to the availability of absorptive capacity. We also proposed that science productivity at the time of forming new external knowledge access collaborations will enhance the long-term benefits of those collaborations for the firm's innovative output.

Our study holds important implications for the knowledge management literature, particularly the stream of research focused on transfer and absorption of external knowledge from professional networks as well as the broader knowledge community (Oliva, 2014; Matricano et al., 2014; Del Giudice & Maggioni, 2014). The process of knowledge management encompasses several elements from knowledge definition and identification to knowledge usage and retentions (Oliva, 2014). The knowledge management processes most relevant to our work are those pertaining to accessing, acquiring, and using of externally-sourced knowledge in the firm's internal innovation process. We know from prior research that knowledge management indeed influences firm competitiveness and performance through different organizational mechanisms (Andreeva & Kianto, 2012), including via its role in firm innovation as innovation is highly dependent on the availability and accessibility of advanced and cutting-edge knowledge (du Plessis, 2007). The ultimate connection from investments in basic scientific research to the firm's innovative performance maps onto what is known as the realized absorptive capacity. Specifically, the contingent role of basic research investments in the development of scientific absorptive capacity that in turn enables the firm to form future knowledge access collaborations can be thought of as an underlying mechanism for the development of a dynamic capability to acquire and assimilate cutting-edge knowledge from external sources (Teece et al., 1997). However, such cutting-edge knowledge can only benefit the firm's internal innovations if it is transformed and exploited effectively by deploying the firm's advanced dynamic capabilities developed earlier through proper investments in underlying basic research to align the firm's knowledge repertoire with the external knowledge landscape of the industry. Moreover, the contingencies proposed in our model can be thought of as particular instances of knowledge management barriers in the special context of managing the transfer of basic scientific knowledge (Oliva, 2014). For instance, isolation of firm scientists and lack of focus in external co-authorships would most resemble human barriers, while organizational dispersion of research efforts, R&D alliance experience, and current innovative output of the firm would correspond to organizational barriers.

The argument that isolation from the external scientific community and the organizational dispersion of scientific projects both lower the firm's availability of absorptive capacity holds implications for the process of managing and organizing basic research in firms. These implications particularly concern firms with well-developed hierarchical structures where the firm's in-house basic research program is likely to be managed as an independent function. Specifically, by encouraging external collaborations and consolidating basic research in the organization, the executives in charge of the firm's basic research program may be able to boost the impact of science as a foundational link in the firm's innovation value chain and ensure that basic research maintains its status as a map guiding the firm's search for new innovations (Fleming & Sorenson, 2004). We also argued that the firm's patenting productivity, experience with R&D alliances, and size are all likely to suppress the effect of scientific absorptive capacity as a driver of new R&D alliance formations. This suggests that executives in charge of basic research programs in large, innovative, and experienced firms may have to face higher barriers to recognition regarding the contribution of their program and its outcomes to the firm's innovation strategy.

The theoretical framework developed in this paper also contributes by painting a multi-layered picture of firms' innovative capabilities (Teece et al., 1997; Eisenhardt & Martin, 2000). Prior research has mainly focused either on the internal or the external dimension of what constitutes a firm's innovative capability. Subramaniam and Youndt (2005), for instance, examined the influence of intellectual capital in a firm on its innovative capabilities. Hagedoorn and Duysters (2002), on the other hand, explored the differences in firms' preferences toward external sources of innovative capabilities such as strategic alliances or mergers and acquisitions. This study highlights the need to re-conceptualize innovative capabilities as multilayered constructs with internal components (e.g. science productivity) and external elements (e.g. R&D alliances) that are closely-knit and nearly impossible to dissect and analyze independently. In fact, innovative capabilities may be better understood as residing not only in individual internal and external components, but also in the architecture connecting these components (Henderson & Cockburn, 1994).

Future research could extend our theory by adopting a two-sided perspective on the process of forming knowledge collaborations to account for differences in partner capabilities and motivations in the formation of the collaborations. Future research could also examine the nature of external knowledge access collaborations in more details to determine if the alliances that are not based on sufficient absorptive capacity involve a systematically different set of partners that might have been attracted to the focal firm for such reasons as favorable terms of partnership. Research could also probe how differences in partner type and quality impact the formation and outcomes of knowledge accessing collaborations. Lastly, it is worth examining if scientific collaborations with researchers employed at former organizational collaboration partners impacts the effectiveness of the resulting absorptive capacity in any specific manner including its effect on the trajectory of the future collaborations formed by the focal firm.

REFERENCES

Anand, B. N., & Khanna, T. (2000). Do firms learn to create value? The case of alliances. Strategic Management Journal, 21(3), 295-315.

Andreeva, T., & Kianto, A. (2012). Does knowledge management really matter? Linking KM practices, competitiveness and economic performance. Journal of Knowledge Management, 16(4), 617-636.

Arora, A., & Gambardella, A. (1994). Evaluating technological information and utilizing it: Scientific knowledge, technological capability, and external linkages in biotechnology. Journal of Economic Behavior & Organization, 24(1), 91-114.

Cardinal, L. B. (2001). Technological innovation in the pharmaceutical industry: The use of organizational control in managing research and development. Organization Science, 12(1), 19-36.

Cockburn, I. M., & Henderson, R. M. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. The Journal of Industrial Economics, 46(2), 157-182.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), 128-152.

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? Journal of the American Society for Information Science and Technology, 52(7), 558-569.

Del Giudice, M., & Maggioni, V. (2014). Managerial practices and operative directions of knowledge management within inter-firm networks: a global view. Journal of Knowledge Management, 18(5), 841-846.

Dunne, D., & Dougherty, D. (2006). Learning for Innovation in Science-Based Industries: The Case of Pharmaceutical Drug Discovery. In DRUID Summer Conference Proceedings, 18-20.

Durand, R., Bruyaka, O., & Mangematin, V. (2008). Do science and money go together? The case of the French biotech industry. Strategic Management Journal, 29(12), 1281-1299.

Du Plessis, M. (2007). The role of knowledge management in innovation. Journal of knowledge management, 11(4), 20-29.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? Strategic Management Journal, 21(10-11), 1105-1121.

Emerson, R. M. (1962). Power-dependence relations. American Sociological Review, 27(1), 31-41.

Englebrecht, T. D., Hanke, S. A., & Kuang, Y. (2008). An assessment of patterns of co-authorship for academic accountants within premier journals: Evidence from 1979–2004. Advances in Accounting, 24(2), 172-181.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. Strategic Management Journal, 25(8-9), 909-928.

Gelman, S. R., & Gibelman, M. (1999). A quest for citations? An analysis of and commentary on the trend toward multiple authorship. Journal of Social Work Education, 35(2), 203-213.

Gibelman, M., & Gelman, S. R. (2000). The trend toward multiple authorship in social work: A British and US comparison. Social Work Education, 19(4), 349-358.

Gittelman, M., & Kogut, B. (2003). Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science, 49(4), 366-382.

Gulati, R., & Singh, H. (1998). The architecture of cooperation: Managing coordination costs and appropriation concerns in strategic alliances. Administrative Science Quarterly, 43(4), 781-814.

Gulati, R., Lavie, D., & Singh, H. (2009). The nature of partnering experience and the gains from alliances. Strategic Management Journal, 30(11), 1213-1233.

Hagedoorn, J., & Duysters, G. (2002). External sources of innovative capabilities: the preferences for strategic alliances or mergers and acquisitions. Journal of Management Studies, 39(2), 167-188.

Hamel, G. (1991). Competition for competence and interpartner learning within international strategic alliances. Strategic Management Journal, 12(S1), 83-103.

Henderson, R., & Cockburn, I. (1994). Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal, 15(S1), 63-84.

Huang, K. G., & Murray, F .E. (2009). Does patent strategy shape the long-run supply of public knowledge? Evidence from human genetics. Academy of Management Journal, 52(6), 1193-1221.

Inkpen, A. C., & Beamish, P. W. (1997). Knowledge, bargaining power, and the instability of international joint ventures. Academy of Management Review, 22(1), 177-202.

Kale, P., Dyer, J. H., & Singh, H. (2002). Alliance capability, stock market response, and long-term alliance success: the role of the alliance function. Strategic Management Journal, 23(8), 747-767.

Kale, P., Singh, H., & Perlmutter, H. (2000). Learning and protection of proprietary assets in strategic alliances: Building relational capital. Strategic Management Journal, 21(3), 217-237.

Kale, P., & Singh, H. (2007). Building firm capabilities through learning: the role of the alliance learning process in alliance capability and firm-level alliance success. Strategic Management Journal, 28(10), 981-1000.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? Research Policy, 26(1), 1-18.

Khanna, T., Gulati, R., & Nohria, N. (1998). The dynamics of learning alliances: Competition, cooperation, and relative scope. Strategic Management Journal, 19(3), 193-210.

Kogut, B., & Zander, U. (1993). Knowledge of the firm and the evolutionary theory of the multinational corporation. Journal of International Business Studies, 24(4), 625-645.

Lane, P. J., & Lubatkin, M. (1998). Relative absorptive capacity and interorganizational learning. Strategic Management Journal, 19(5), 461-477.

Makadok, R. (2001). Toward a synthesis of the resource-based and dynamic-capability views of rent creation. Strategic Management Journal, 22(5), 387-401.

Mansfield, E. (1991). Academic research and industrial innovation. Research Policy, 20(1), 1-12.

Mansfield, E. (1998). Academic research and industrial innovation: An update of empirical findings. Research Policy, 26(7-8), 773-776.

Matricano, D., Candelo, E., Sorrentino, M., & Martínez-Martínez, A. (2019). Absorbing inbound knowledge within open innovation processes. The case of Fiat Chrysler Automobiles. Journal of Knowledge Management, 23(4), 786-807.

Oliva, F. L. (2014). Knowledge management barriers, practices and maturity model. Journal of Knowledge Management, 18(6), 1053-1074.

Orsenigo, L., Pammolli, F., & Riccaboni, M. (2001). Technological change and network dynamics: Lessons from the pharmaceutical industry. Research Policy, 30(3), 485-508.

Osterloh, M., & Frey, B. S. (2000). Motivation, knowledge transfer, and organizational forms. Organization Science, 11(5), 538-550.

Oxley, J. E., & Sampson, R. C. (2004). The scope and governance of international R&D alliances. Strategic Management Journal, 25(8-9), 723-749.

Park, S. H., Chen, R., & Gallagher, S. (2002). Firm resources as moderators of the relationship between market growth and strategic alliances in semiconductor start-ups. Academy of Management Journal, 45(3), 527-545.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. Administrative Science Quarterly, 41(1), 116-145.

Roach, M. (2009). When do firms use public research? The determinants of knowledge flows from universities and government labs to industrial R&D. In DRUID Summer Conference Proceedings.

Rothaermel, F. T., & Hess, A.M. (2007). Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects. Organization Science, 18(6), 898-921.

Schmidt, T. (2005). What determines absorptive capacity. In DRUID summer conference Proceedings.

Serajuddin, A. T. (1998). Education in pharmaceutical sciences needs a brand new direction to meet the challenges of drug research and development. Pharmaceutical Research, 15(1), 8.

Singh, J. (2008). Distributed R&D, cross-regional knowledge integration and quality of innovative output. Research Policy, 37(1), 77-96.

Slafer, G. A. (2005). Multiple authorship of crop science papers: are there too many co-authors? Field Crops Research, 2(94), 272-276.

Stern, S. (2004). Do scientists pay to be scientists? Management Science, 50(6), 835-853.

Subramaniam, M., & Youndt, M. A. (2005). The influence of intellectual capital on the types of innovative capabilities. Academy of Management Journal, 48(3), 450-463.

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. Strategic Management Journal, 17(S2), 27-43.

Teece, D. J. (1977). Technology transfer by multinational firms: The resource cost of transferring technological know-how. Economic Journal, 87(346), 242-261.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18(7), 509-533.

Tzabbar, D. (2009). When does scientist recruitment affect technological repositioning? Academy of Management Journal, 52(5), 873-896.

Woods, R. H., Youn, H., & Johanson, M. M. (2010). Single vs. co-authored and multi-authored research articles: Evaluating the views and opinions of ICHRIE scholars. International CHRIE Conference-Refereed Track 11.

Zander, U., & Kogut, B. (1995). Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. Organization Science, 6(1), 76-92.

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1994). Intellectual capital and the birth of US biotechnology enterprises (No. w4653). National Bureau of Economic Research.

Zucker, L. G., Darby, M. R., & Armstrong, J. S. (2002). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. Management Science, 48(1), 138-153.