Aalborg Universitet

AALBORG UNIVERSITY

DENMARK

The reason why profitable firms do not necessarily grow

Holm, Jacob Rubæk; Andersen, Esben Sloth; Metcalfe, J. Stanley

The reason why profitable firms do not necessarily grow

Confounded economic selection and how to measure it

Jacob R. Holm (corresponding)

Aalborg University

jrh@business.aau.dk

Esben S. Andersen

Aalborg University

J. Stanley Metcalfe

University of Manchester

Draft, 2014-06-30

1. Introduction

Aggregate outcomes such as technological change and economic growth are results of microevolutionary processes of novelty creation and selection. Much research focus on novelty creation such as innovation and entrepreneurship. This paper extends the analysis of economic selection by formalizing tools for empirical analysis of coevolution of multiple characteristics in a population, and by applying these tools in a simple simulation exercise.

Economic selection – the increasing predominance of superior routines through the propensity of business units with superior performance to increase in relative size – is generally studied empirically or formally modelled in a simplified manner where it is assumed to be directional and depending on a single performance characteristic. In Andersen and Holm (2014) we explored analytically and simulated more complex cases in which selection is not necessarily directional while the possibility of confounding selection processes working in opposite directions (e.g. at the firm and industry level) was studied empirically in Holm (2014). However, implicitly underlying both of these studies is the assumption that selection is upon a single characteristic. This is a common assumption, especially in the guise that selection among firms is assumed to be based on productivity of profitability, but it is often too simplistic. Research has not always been able to identify this simple selection process empirically and it has hence been suggested that selection might work on some combination of characteristics such as financial performance combined with the propensity to re-invest profits (Coad 2007; Bottazzi et al. 2010; Coad and Teruel 2013). It does not take a very complicated formal model of competition for firms' heterogeneous preferences for expansion to seriously confound the relationship between productivity and growth (Metcalfe 2010).

With inspiration from Rice (2004) we show that it is possible and practically feasible to quantify economic selection in empirical studies even when simultaneous selection on multiple characteristics of business units is confounding the relationship between covarying characteristics and fitness. The confounding effects of investment behaviour and financial performance may be disentangled with this methodology. But it also has other uses such as the study of simultaneous selection in factor and output markets (Metcalfe 1997; Baldwin and Gu 2006; Metcalfe and Ramlogan 2006). Even in a straightforward model of competition where firms compete by undercutting each other's prices the simple assumption that labour markets are less than perfect and that growing firms hence must offer a relatively high wage rate to attract employees, and thus have higher unit costs, entails that the relationship between profitability and growth becomes confounded (Metcalfe 1997). As soon as profits are imperfectly correlated with investment decisions we cannot ensure that the most competitive firm in terms of unit costs is the fittest firm, in the evolutionary sense of having the fastest rate of growth among its population of rivals.

The discussion below is, in essence, a warning against the perils of simple models of selection in which firms are allowed to vary in only one dimension, usually unit cost or its inverse total factor productivity, and have that selective characteristic tested in only one market, typically the product market for the firm. As a pedagogic device this is perfectly reasonable and much can be demonstrated about the attributes of evolutionary processes and the dynamics of creative destruction, particularly the idea that evolution depends on correlation of characteristics with aspects of firm performance, growth price setting and so on. However, real world economic selection processes are not typically of this simple kind and many an empirical puzzle can only be illuminated if a more general approach is followed. In particular, firms differ in the quality of the goods that they sell and an essential element of Schumpeterian competition is defined by product innovation and not just process innovation. Moreover, firms compete not only for customers but

for employees and for access to capital so that labour and capital markets deeply condition the rate and direction of evolution across the population of firms that we call an industry. A far richer account of economic evolution depends on taking these, and related dimensions of the competitive process seriously and our task here is to set out some of the grounding principles of a more general evolutionary economics in which selection across bundles of many characteristics takes place within a system of interdependent markets for goods and factors of production.

This draft paper proceeds as follows. Section 2 presents a simple analytical framework for handling the confoundedness produced by multivariate selection and multiple types of selection. Section 3 presents models developed within this framework as well as simulation results. Section 4 includes a few conclusions.

2. An analytical framework

More than thirty years ago, Nelson and Winter (1982, 29) wrote that "the intellectual coherence and power of thinking about Schumpeterian competition have been quite low, as one would expect in the absence of a well-articulated theoretical structure to guide and connect research." The analytical situation has in the meantime improved radically, especially with respect to the analysis of the selection and evolution of single characteristics. In contrast, the study of selection and evolution of multiple and interdependent characteristics still seems to lack analytical guidance and integration. Since the much-needed analytical framework does not seem to emerge spontaneously within evolutionary economics, we suggest the shortcut of cautiously importing and redesigning analytical tools from evolutionary biology. The need for caution should be obvious for any evolutionary economist who recognises the heavy dependence of the biological toolbox on specific modes of transmission of genetic traits across generations. Nevertheless, many evolutionary biologists help to bridge the gap by working with generic evolutionary tools and/or at a high level of abstraction.

2.1. Approaching multivariate evolution

Metcalfe (1994, 329) moved from R. A. Fisher's specific theorem of genetics-based natural selection to the general "Fisher Principle" in order to make the work of the great statistician and evolutionary biologist relevant for evolutionary economics. The Fisher Principle states that "in the context of a population of diverse behaviours across which selection is taking place in a constant environment, the rate of change of mean behaviour is a function of the degree of variety in behaviour across the population." Under such circumstances the gradually evolving mean behaviour becomes increasingly informed about and adapted to the environment of the population. Evolutionary economists have formalised and applied this principle in the study of simple selection and evolution in simple environments. However, the conditions of Fisher's Principle are seldom fulfilled. First, the stability or lawful patterning of the environment of an economic population obviously cannot always be taken for granted. Second, selection can work on a number of more or less conflicting behavioural characteristics. For example, a fluctuating environment may repeatedly shift the characteristics that are focussed upon by selection. Furthermore, the input markets and the output markets can emphasise conflicting characteristics of the population. Third, the importance of multiple and shifting characteristics means that it is often not obvious which characteristics of behaviour that has to be recreated when old variance has been used up by the selection process.

The move from Fisher's Principle toward an extended and more operational toolbox for theoretical and applied evolutionary economics involves a large research agenda. The turbulent environment and its

shifting focus on different characteristics of behaviour have already to some extent been confronted by innovation studies. Furthermore, industrial dynamics has studied the systematic change of selective focus between different behavioural characteristics during the industry life cycle. However, we are still missing general principles and statistical methods for handling selection and evolution of multiple and potentially conflicting characteristics of behaviour. The lack of analytical tools seems to have slowed down the move from the well-understood univariate analysis to the general analysis of multivariate selection and evolution. In turn, the lack of multivariate analysis has decreased the analytical clarity and power of evolutionary economic studies that try to extend Fisher's Principle in other directions.

To move from univariate to multivariate selection has already been made within evolutionary biology. The statistical procedures for theorising and data analysis can be traced back to Fisher (1930), but a very helpful jump forward was made by the Chicago School, a group of Chicago biologists working within quantitative genetics in the late 1970s and early 1980s (Lande and Arnold 1983; Conner and Hartl 2004). The Chicago approach to phenotypical selection and evolution is based on the statistical analysis of the fundamental requirements for any evolutionary process: variance of the characteristics of the population, covariance between characteristics and the reproduction of members, and the intertemporal inertia of the characteristics. By focusing on these requirements for phenotypical evolution rather than on the direct study of genetic evolution, this approach has been very successful for studying "natural selection in the wild" (Endler 1986; Brodie et al. 1995; Kingsolver et al. 2001; Kingsolver and Pfennig 2007). With some caution and modification, the approach can also be used for the analysis of economic selection and evolution. This use has been eased by reformulations and developments by e.g. Rice (2004) of the Chicago School approach in relation to the very general analytical framework of R. A. Fisher and George R. Price. Since we have already developed the basic analytical framework elsewhere (Andersen 2004; Metcalfe and Ramlogan 2006; Andersen and Holm 2014; Holm 2014), we in the following move quickly from Price's Equation to the Chicago novelties with respect to evolutionary economics.

George Price (1970; 1995) worked at a deeper level than the Chicago School. He thought in terms of a population that is studied at two subsequent points of time, t and t'. He assumed that any member of the t'-population can be connected to a member of the t-population. This made it possible for him to define absolute fitness for each t-member as the number by which multiply its t-size to get its representation in the t'-population. Then Price defined evolution as the change of the population mean of a characteristic between the two points of time. He also defined selection as the part of evolution that can be explained by the covariance between the characteristic values of the members of the t-population and their fitness. The residual of the evolutionary change of mean characteristic is explained fully by mean intra-member change evaluated in the t'-population. Thus Price's Equation – or Price's Identity – can be written as

$$\text{Total evolutionary change} = \text{Selection effect} + \text{Intramember effect}. \qquad (1)$$

The developers and users of the Chicago approach agree with Price, but they normally focus on concrete problems of artificial selection and natural selection in the wild. In these connections, the problems of handling selection on multiple characteristics are obvious. For example, when breeders are performing artificial selection, they recognise that by selecting on a single characteristic they are often coselecting unwanted characteristics. The Chicago approach handles this and similar problems by thinking of total evolutionary change as a vector that consists of the changes in a number of different characteristics (e.g. Lande and Arnold 1983). In the context of artificial selection, each element of the vector can e.g. be a selection differential, i.e. the difference between the mean value of the patents chosen for breeding and

the mean value of all potential parents in the population (most of which have be slaughtered or harvested). However, this selection differential is the combined result of direct selection on the studied characteristic and the indirect effects on that characteristic of (artificial) selection working on other characteristics.

To confront this difficulty the Chicago School has provided two new tools (under the assumption of multivariate normal distribution; otherwise things get complex). The first tool is the vector of selection gradients, i.e. the direct effects of selection on the different characteristics. While a selection differential includes both the direct and the indirect selection on a characteristic, a selection gradient is the partial regression of relative fitness on a characteristic. Thus the selection gradient ignores indirect selection due to other analysed characteristics and measures only the direct selection of the characteristic. The second tool to handle multiple characteristics is the matrix of phenotypic covariances between characteristics. This matrix reflects the fact that different characteristics may be interdependent. For example, we have the case in which members of the t-population that have high values of one characteristic also tend to have high (or low) values of coupled characteristics. This means that when selection acts directly on one characteristic, it also influences the population mean of more or less closely coupled characteristics. The elements of the phenotypic covariance matrix can be zero, positive, or negative.

By combining the two new tools, we can understand the strange ways in which the process of selection on coupled characteristics might work. For instance, the change of the mean of the focal first characteristic is potentially influenced by all the studied characteristics. The selection effect in equation (1) now consists of one direct effect and multiple indirect effects. The direct effect is found by multiplying the first element of the covariance matrix by the first element of the vector of selection gradients. Thus we are multiplying a covariance with a partial regression coefficient. However, since we are here dealing with the covariance of the first characteristic with itself, we are actually multiplying the variance of the first characteristic by the efficiency of direct selection on that characteristic. The indirect effects might involve important covariances (and thus correlations). For example, the first indirect effect on the change of the mean of the first characteristic is found by multiplying the covariance between characteristics 1 and 2 by the selection gradient of characteristic 2. Surprising things can result from this multiplication since the covariance might be negative and the selection gradients of characteristics 1 and 2 might have opposite signs. Thus this indirect selection of characteristic 1 might remove or invert a positive direct selection on characteristic 1. However, although the effects of such couplings of characteristics have been analysed intensively by evolutionary biology, there are still discussion about the frequency of this phenomenon in nature (Agrawal and Stinchcombe 2009).

The Chicago School has also promoted the analytical distinction between different types of selection on the characteristics of the vector of evolutionary changes of characteristics. Although we have already dealt with this contribution (Andersen and Holm 2014), it is worth repeating that it is important to add other types of selection to the type of selection that is implied by Fisher's Theorem and Fisher's Principle. While most thinking of selection with evolutionary economics has been dominated by the – positive or negative – directional selection that results in a change of the mean of a characteristic, it is easy to define and search for two other types of types of selection that can take place with a constant mean of the population. In the Chicago approach all three types of selection are easily captured by means of multiple regression. The dependent variable is relative fitness. The linear independent variable is the value of the characteristic, which relates to the mean of the population. The nonlinear independent variable is the squared distance of the characteristic from the mean, which relates to variance. Given these definitions, the model estimates the coefficients of the independent variables, b_1 and b_2. We observe pure directional selection if b_1 is

different from zero and b_2 is zero. We have pure stabilising selection if b_1 is zero and b_2 is negative. In contrast, pure diversifying selection means that b_1 is zero and b_2 is positive. Both types of "variance selection" seem important for the overall evolutionary process. Diversifying selection can help to create new populations. Stabilising selection (with the special case of purifying selection) helps to uphold complex systems of characteristics on which gradual adaptation depends.

2.2. A simple framework for analysing multivariate selection

It is time to convert the short presentation of the works of Fisher and Price as well as of the Chicago School approach into a simple framework for analysing multivariate selection and evolution. The core trick is to operate in the short run. In contrast to long-term analysis we can to a larger degree concentrate on selection and thus minimise the importance of the some of the huge differences between economics and biology. We also think that the simple short-term framework helps to think clearly about concepts and measurements of selection. Finally, in the Nordic countries and some other countries immensely rich data on each firm and each citizen has become available for relatively short-term scientific analysis.

Our basic analyses apply two subsequent population censuses. Since we emphasise selection, we call them the pre-selection census and the post-selection census. When needed, we distinguish by adding a prime to variables that relate to the post-selection census. The two censuses provide the basis for calculating statistics on fitness and characteristics as well as the relationships between them. This procedure can be presented in three basic steps (Andersen and Holm 2014), but in the present paper we include three additional steps.

1. In each of the two censuses we for each member measure its population share. Then we calculate the (relative) fitness of each member as the ratio of its population shares after and before selection. Thus the population has mean fitness $\bar{w} = 1$.

2. The censuses provide information on a focal characteristic 1 of the members of a population. In each census we for each member measure the characteristic value of z_1, and we calculate the member-level change of z_1 between the two censuses. We then calculate the weighted means of z_1 in each of the two censuses, $\bar{z}_1$ and $\bar{z}_1'$, as well as the change of the mean between censuses, $\Delta\bar{z}_1$. We also with respect to the post-population shares calculate the weighted mean of the member-level change of the characteristic, $E(w\Delta z_1)$.

3. We use the member-level information to calculate the covariance between fitness and characteristic 1, $Cov(w, z_1)$. This covariance is equal to the product of the total regression of fitness on the characteristic and the variance of the characteristic, $\beta_{\omega,z_1}^{\text{Total}}Var(z_1)$.

4. We extend the censuses beyond the focal characteristic to cover a set of new characteristics labelled from 2 to n. We do so by repeating steps (2) and (3) for each additional characteristic. One of the results is that we are provided with a vector of changes of mean characteristics,

$$\Delta\bar{\mathbf{z}} = \begin{bmatrix} \Delta\bar{z}_1 \\ \cdots \\ \Delta\bar{z}_n \end{bmatrix}.$$

   We want to explain this vector – with special emphasis on $\Delta\bar{z}_1$. We have much of the information for this analysis, but steps (5) and (6) provide us with crucial tools.

5. For the pre-selection census we check whether the characteristics are correlated by calculating the "phenotypic" covariance matrix,

$$\mathbf{P} = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \cdots & \cdots & \cdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} = \begin{bmatrix} Cov(z_1, z_1) & \cdots & Cov(z_1, z_n) \\ \cdots & \cdots & \cdots \\ Cov(z_n, z_1) & \cdots & Cov(z_n, z_n) \end{bmatrix},$$

where the diagonal represents variances, since e.g. $Cov(z_1, z_1) = Var(z_1)$. The rest of the symmetric matrix is filled with covariances, where e.g. $Cov(z_1, z_n) = Cov(z_n, z_1)$. We can also calculate $\mathbf{P}'$, that is, the phenotypic covariance matrix for the post-selection population. By comparing it with $\mathbf{P}$ we can check the stability over time of the matrix, but this comparison is not made in the present paper.

6. We end the census-related work by calculating the vector of partial regressions of fitness on each of the characteristics,

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{w,z_1} \\ \cdots \\ \beta_{w,z_n} \end{bmatrix}.$$

We have an interesting case if, for example, β_{w,z_1} is different from $\beta_{w,z_1}^{\text{Total}}$, which was provided by step (3).

This six-step procedure prepares real analysis. Much can be learned from the first three steps, For example, we can turn to simple applications of Price's Equation (1) for analysing the relative importance of the selection effect and the intramember effect with respect to individual characteristics

$$\Delta \bar{z} = Cov(w, z) + E(w \Delta z). \tag{2}$$

By rewriting Price's Equation (2), we obtain

$$\Delta \bar{z} = Cov(w, z) + Cov(w, \Delta z) + E(\Delta z), \tag{3}$$

which is quite interesting since it is exactly identical to the technique introduced in 1998 by Foster and colleagues. The last term is the "within effect", i.e. the change in $\bar{z}$ in the absence of any selection, while the middle term is the "cross-level covariance-type effect". It is covariance between relative fitness and the change in characteristic. As with the original Price's equation, equation (3) has the important property that it "eats its own tail" meaning that it can be applied in multilevel studies of evolutionary processes (Holm 2014). However, in the remainder of this paper we will focus on $Cov(w, z)$ as we will generally assume that $\Delta z = 0$ and thus $Cov(w, \Delta z) = 0$.

In any case, the main task of our six-step procedure is to promote the analysis of multivariate selection. We can use the $\mathbf{P}$ matrix and the vector of selection gradients for the characteristics to predict the responses to selection pressures that act simultaneously on these characteristics. In condensed matrix format the equation is

$$\Delta \bar{\mathbf{z}} = \mathbf{P} \boldsymbol{\beta}. \tag{4}$$

Let us, for example, expand the mean change of the first characteristic:

$$\Delta \bar{z}_1 = Var(z_1) \beta_{w,z_1} + Cov(z_1, z_2) \beta_{w,z_2} + \cdots + Cov(z_1, z_n) \beta_{w,z_n}.$$

Here the evolutionary response to selection of the first characteristic has n components. First, there is the direct response to selection, $Var(z_1) \beta_{w,z_1}$, which consists of the change in the mean of characteristic 1 due to selection acting directly on characteristic 1. The other $n-1$ contributions are the indirect responses to selection due to correlation (and this covariance). For example, $Cov(z_1, z_2) \beta_{w,z_2}$ represents the indirect change in the mean of characteristic 1 due to its covariance with the second characteristic.

It is, of course, possible that the change in the mean of characteristic 1 is totally or largely due to the direct selection on characteristic 1. Another possibility is that several correlated responses to selection dominate over the direct response (see table 1). In the extreme case, direct selection tend to produce high levels of the first characteristic while its mean becomes smaller due to negative covariances or to low values of other characteristics. These and other possibilities are presented in table 1. Here we distinguish between five types of bivariate selection (cf. Conner and Hartl 2004, 223). If we ignore the case in which bivariate selection reduces to two univariate selections, we can classify the outcomes in terms of the signs of the selection gradients and the correlations of characteristics. For example negative correlation of the characteristics and gradients with opposite signs leads to negative augmentation of direct selection. However, when correlations are still negative while gradients have the same sign, we are facing what may be called a correlation constraint on the evolution of the characteristics.

Table 1: Effects on evolutionary change of signs of selection gradients and correlations of characteristics

Type of selection on two characteristics	Definition	Stylised example
Univariate selection	No correlation of characteristics	$\begin{bmatrix} \Delta\bar{z}_1 \\ \Delta\bar{z}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Positive augmentation	Positive correlation of characteristics + gradients with same sign	$\begin{bmatrix} \Delta\bar{z}_1 \\ \Delta\bar{z}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Negative augmentation	Negative correlation of characteristics + gradients with opposite sign	$\begin{bmatrix} \Delta\bar{z}_1 \\ \Delta\bar{z}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
Gradient constraint	Positive correlation of characteristics + gradients with opposite sign	$\begin{bmatrix} \Delta\bar{z}_1 \\ \Delta\bar{z}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
Correlation constraint	Negative correlation of characteristics + gradients with same sign	$\begin{bmatrix} \Delta\bar{z}_1 \\ \Delta\bar{z}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

2.3. Path analysis of multivariate selection

As emphasised by Frank (2013), the use of multivariate analysis has two major purposes. The approach to multiple regression analysis that is based on classical statistical theory focuses on correlations and variances and rejects causal interpretation (since correlation is not causation). In contrast, the approach to multiple regression analysis that is based on path analysis makes statistical models that are based on hypotheses about causes. We clearly prefer the latter approach since it can support theoretical analyses of adaptation produced by selection.

Let us take the study the evolution of three interdependent characteristics: z_Y, z_L, z_K, which might be interpreted as firm characteristics that, respectively, relates to the output market, the capital market, and the labour market. If the population of firms that have these three characteristics is statistically well-behaved, we can mechanically use the Chicago approach. Since $\Delta\bar{z} = \mathbf{P}\boldsymbol{\beta}$, we have

$$\begin{bmatrix} \Delta\bar{z}_Y \\ \Delta\bar{z}_K \\ \Delta\bar{z}_L \end{bmatrix} = \begin{bmatrix} Var(z_Y) & Cov(z_Y, z_L) & Cov(z_Y, z_K) \\ Cov(z_L, z_Y) & Var(z_L) & Cov(z_L, z_K) \\ Cov(z_K, z_Y) & Cov(z_K, z_L) & Var(z_K) \end{bmatrix}\begin{bmatrix} \beta_{w,z_Y} \\ \beta_{w,z_L} \\ \beta_{w,z_K} \end{bmatrix} \quad (5)$$

The same result can be obtained by a more elaborate but potentially more enlightening way by path analysis. Thus Rice (2004) suggests the use of path analysis to deduce decomposition equations for multivariate selection problems. Path analysis is a particularly useful tool when seeking to describe and understand the origins of selection differentials as it allows complicated problems to be illustrated in a lucid figure. This is trivial for the above simple versions Price's equation, but they can be important when multivariate selection is being studied (since the relational structure quickly becomes complicated).

A path diagram contains all variables of interest and their relations. Relations are described by a straight line with an arrow at one end if it represents a hypothesis about a partial regression coefficient and by a curved line with arrows at both ends if it is a covariance. Source variables are those that have no incoming straight lines, so they can only be connected by covariance relations. Other variables are called downstream variables.

The covariance between any two variables can be computed as the sum of all different paths leading from one variable to the other. The paths are constructed using the following rules. First, you can move either backwards or forwards along a straight line but not both. Second, you cannot pass through the same point more than once. Third, you cannot pass through more than one curved line.

The value of each path is computed as the product of the following effects: a covariance for the curved line (if any), a variance for each source variable with no curved line and a regression coefficient (potentially partial) for each straight line. As a simple example consider the covariance term of equation (2). The associated path diagram is then

$$z_i \xrightarrow{\hspace{4cm}} w$$
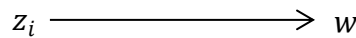
Here we obviously have only one path. The covariance term of equation (2) can then be described as the variance of the source variable times the slope coefficient from a linear regression of w on z : $Cov(w, z\ ) = Var(z\ )\beta_{w,z}$, where $\beta_{w,z}$ is the slope coefficient from the simple regression $w_i = \beta_0 + \beta_{w,z}z_i + error$

Path analysis becomes interesting when there is more than one source variable. Suppose that we are studying the evolution of three characteristics of equation (4): $z_Y, z_K, z_L,$. A path diagram for this case is illustrated in figure 1.
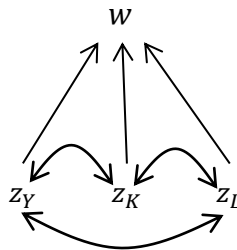

Figure 1: The path diagram behind equation (5)

$Cov(w, z_Y)$ must now be computed as the sum of three paths: The direct path from the source variable z_Y to w and the two paths passing through each of the other source variables on the way to w:

$$Cov(w, z_Y) = Var(z_Y)\beta_{w,z_Y} + Cov(z_Y, z_L)\beta_{w,z_L} + Cov(z_Y, z_K)\beta_{w,z_K}$$

9

where the betas are slope coefficients from the multiple regression $w_i = \beta_0 + \beta_{w,z_Y} z_{Y,i} + \beta_{w,z_L} z_{L,i} + \beta_{w,z_K} z_{K,i} + error$. The mean characteristic then evolves in the population according to

$$\Delta \bar{z}_Y = Var(z_Y)\beta_{w,z_Y} + Cov(z_Y, z_L)\beta_{w,z_L} + Cov(z_Y, z_K)\beta_{w,z_K}$$

For a given vector of characteristics identified as determinants of relative growth of firms (i.e. fitness) the associated path diagram generally consist of the characteristics as source variables and fitness as the only downstream variable. Based on figure 1 it can be deduced that the selection differential for the coevolution of the three characteristics can be written in matrix form as the product of the covariance matrix of the characteristics and the vector of partial regression coefficients of fitness on the characteristics. Total evolution will in general be this selection differential plus the usual intramember effects, cf. equation (2). The resulting equation simply repeats equation (5):

$$\begin{bmatrix} \Delta \bar{z}_Y \\ \Delta \bar{z}_L \\ \Delta \bar{z}_K \end{bmatrix} = \begin{bmatrix} Var(z_Y) & Cov(z_Y, z_L) & Cov(z_Y, z_K) \\ Cov(z_L, z_Y) & Var(z_L) & Cov(z_L, z_K) \\ Cov(z_K, z_Y) & Cov(z_K, z_L) & Var(z_K) \end{bmatrix} \begin{bmatrix} \beta_{w,z_Y} \\ \beta_{w,z_L} \\ \beta_{w,z_K} \end{bmatrix}. \tag{6}$$

Inspecting the elements of the decomposition allows us to explain the coevolutionary pressures creating unexpected results in an empirical or simulation application, given the amount of fuel (variation in characteristics) and efficiency (partial regression coefficients) of the process.

The matrix form expression for the coevolution of multiple characteristics in a population, as in equation (5) and (6), and the approach based on deriving separate weights for various elements in the process in Metcalfe (1997; 2010) provide two different perspectives on the processes. In the matrix form, the evolution of a characteristic will depend on the strength of selection in each market weighted by the covariance of the characteristic in question and the characteristic being selected upon in the relevant market. For example, the evolution of the characteristic that is selected upon in the labour market, z_L, depends on the strength of selection in the capital market only to the extent that z_L covaries with the characteristic that is selected upon in the capital market, z_K (assuming that z_L and z_K are not identical).

3. Modelling and simulation results

3.1. Modelling strategy

In this section we will demonstrate selection may be confounded so that seemingly fit firms do not grow. The modelling strategy has two steps: a data generating algorithm and a deterministic selection process. The data generating algorithm is a set of rules for defining the pre-selection population. It is at this step that we may introduce correlation among the traits of business units. The selection process determines how the pre-selection population evolves into the post-selection population. At this step we may implement different selection functions based on different assumptions regarding the characteristics of business units. Finally the evolution from pre-selection to post-selection population is described by the identity in equation 6. This step is merely a measurement step and we have no influence on the results. The strategy is illustrated on figure 2.
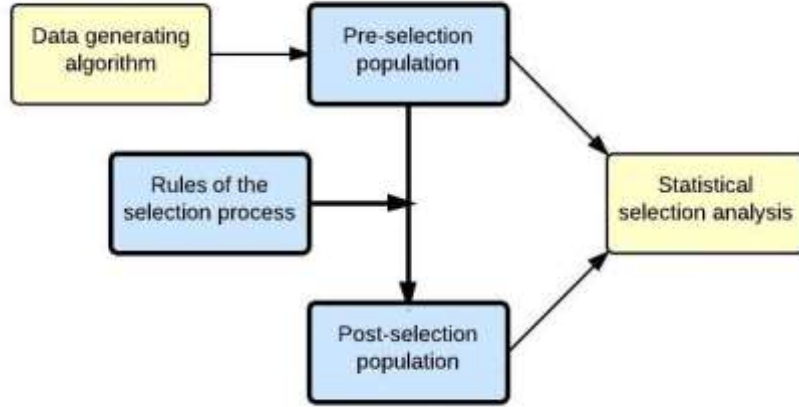
Figure 2: The modelling strategy

When evaluating whether selection is confounded we compare the values of the selection gradient with the observed evolutionary change. Selection is argued to be confounded when these have opposite signs. That is, $\Delta \bar{z}_Y > 0$ while $\beta_{w,z_y} < 0$ selection on z_Y is confounded in the sense that business units with relatively high values of z_Y have decreasing population shares, while at the same time the mean of z_Y is increasing in the population. Each simulation will be repeated 100 times and the results plotted in a $\Delta \bar{z}$ by $\beta_{w,z}$ space.

The pre-selection population will consist of 100 business units. Each business unit is characterised by a vector of three characteristics: $z_i = (z_{Yi} \quad z_{Li} \quad z_{Ki})$. In a more general model business units might enter and exit, and the characteristics of a business unit would change over time through adaptation and innovation but these complications are not included here as they are inconsequential for the research question, as explained above. Thus the evolution of the characteristics in the population $\Delta \bar{z}$, is described fully by $\Delta \bar{z} = P\beta$. Any change in the means of the characteristics must come from the relative growth or decline of business units.

The data generation algorithm proceeds as follows: The three characteristics are all drawn from standard normal distributions with correlation matrix ρ. All business units have equal population shares in the pre-selection population: $s_i = \dfrac{1}{100} \forall i$

For the rules of the selection process we follow the general lines also applied in Andersen and Holm (2014). This means that we specify a deterministic function for absolute fitness. We then transform the outcome into relative fitness and allow the population to evolve according to equation 7.

$$s_i{}' = s_i w_i \tag{7}$$

Relative fitness is defined as absolute fitness divided by population fitness: $w_i = \dfrac{W_i}{E(W)}$ and absolute fitness is determined by the relation:

$$W_i = (1 + C)^{F(z_i)} \tag{8}$$

The general fitness function specified in equation 8 is an exponential function. The parameter C determines the pace of evolution in the sense that a business unit will grow by $C * 100$ percent as many times as

11

specified by the exponent. The specification of the exponent determines the type of selection. In the current paper we have chosen a specification where there is stabilising selection on z_Y with optimum at $z_Y = 0$ and positive directional selection on z_L and z_K. The exponent is determined by:

$$F(z_i) = -|z_{Yi}| + (z_{Li} - \overline{z_{Li}}) + (z_{Ki} - \overline{z_{Ki}}) \tag{9}$$

The rules of the selection process are uniform in all of the simulations presented in the current paper. C is fixed at 0.5. This sets a relatively high pace for evolution but allows us to disregard the possibility of multiple time periods between pre- and post-selection populations. The data generation algorithm is also the same in all simulations except for the value of ρ_{YL}: the correlation between the standard normal distributions from which z_Y and z_L are drawn.

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & \rho_{YL} & 0 \\ \rho_{YL} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

In all simulation there will be directional selection on z_K which is independent of the other characteristics. There will also be directional selection on z_L but this characteristic will to varying degrees be correlated with a third characteristic, z_Y, upon which there is stabilising selection.

3.2. Baseline simulation ($\boldsymbol{\rho_{YL} = 0}$)

In the first simulation the three traits of the business units are uncorrelated; $\rho_{YL} = 0$ in equation 10 above. The pre-selection population is described by the vector of mean characteristics $\bar{\mathbf{z}}$. After having being subject to the deterministic selection process described in equations 7 to 9 the post-selection population is created, and it is described by $\bar{\mathbf{z}}'$. The evolution of the population is described by the change in mean characteristics, $\Delta\bar{\mathbf{z}}$. This evolution is then decomposed into the variance-covariance matrix of the characteristics and the selection gradient as described in equation 5. In a typical baseline simulation the results look as presented in figure 3. The result is as should be expected. $\Delta\bar{z}_Y$ has decreased slightly as the initial population mean of z_Y, $\bar{z}_Y$, was higher than the assumed optimum at zero. The remaining two characteristics have increased in accordance with the assumed directional selection process. The final product on the right is the decomposition $\boldsymbol{P\beta}$. The elements of the selection gradient reflect the assumed selection rule and correspond to the observed evolution in means: z_L and z_K have positive selection coefficients while z_Y has a coefficient close to zero.

$$\begin{bmatrix} \Delta\bar{z}_Y \\ \Delta\bar{z}_L \\ \Delta\bar{z}_K \end{bmatrix} = \begin{bmatrix} 0.232 \\ 0.629 \\ 0.285 \end{bmatrix} - \begin{bmatrix} 0.239 \\ 0.145 \\ -0.108 \end{bmatrix} = \begin{bmatrix} -0.007 \\ 0.485 \\ 0.393 \end{bmatrix} = \begin{bmatrix} 1.174 & 0.042 & 0.190 \\ 0.042 & 1.108 & 0.051 \\ 0.190 & 0.051 & 0.929 \end{bmatrix} \begin{bmatrix} -0.089 \\ 0.421 \\ 0.418 \end{bmatrix}$$

Figure 3. A typical result from a baseline simulation

However the stochastic nature of the data generation process entails that sampling variation can create covariance in $\boldsymbol{P}$ even when correlation is explicitly excluded from $\boldsymbol{\rho}$. This means that the baseline simulation may also result in the output in figure 4 where selection on z_Y is confounded: z_Y is growing in the population while the higher the value of z_Y for a single business unit the lower its relative fitness. It is clear from figure 4 that the result is created by covariance between z_Y and both z_L and z_K: $P_{YL} = 0.143, P_{YK} = 0.266$.

$$\begin{bmatrix} \Delta \bar{z}_Y \\ \Delta \bar{z}_L \\ \Delta \bar{z}_K \end{bmatrix} = \begin{bmatrix} 0.006 \\ 0.425 \\ 0.743 \end{bmatrix} - \begin{bmatrix} -0.121 \\ -0.021 \\ 0.073 \end{bmatrix} = \begin{bmatrix} 0.127 \\ 0.446 \\ 0.670 \end{bmatrix} = \begin{bmatrix} 1.036 & 0.143 & 0.266 \\ 0.143 & 1.046 & 0.035 \\ 0.266 & 0.035 & 1.375 \end{bmatrix} \begin{bmatrix} -0.061 \\ 0.418 \\ 0.488 \end{bmatrix}$$

Figure 4. A relatively rare result from a baseline simulation

Figure 5 plots the results from figures 3 and 4 along with 98 additional simulations with the baseline specification. The values of $\Delta \bar{z}_Y$ and β_{w,z_Y} cluster around zero while of $\Delta \bar{z}_L$, $\Delta \bar{z}_K$, β_{w,z_L} and β_{w,z_K} all are consistently positive.
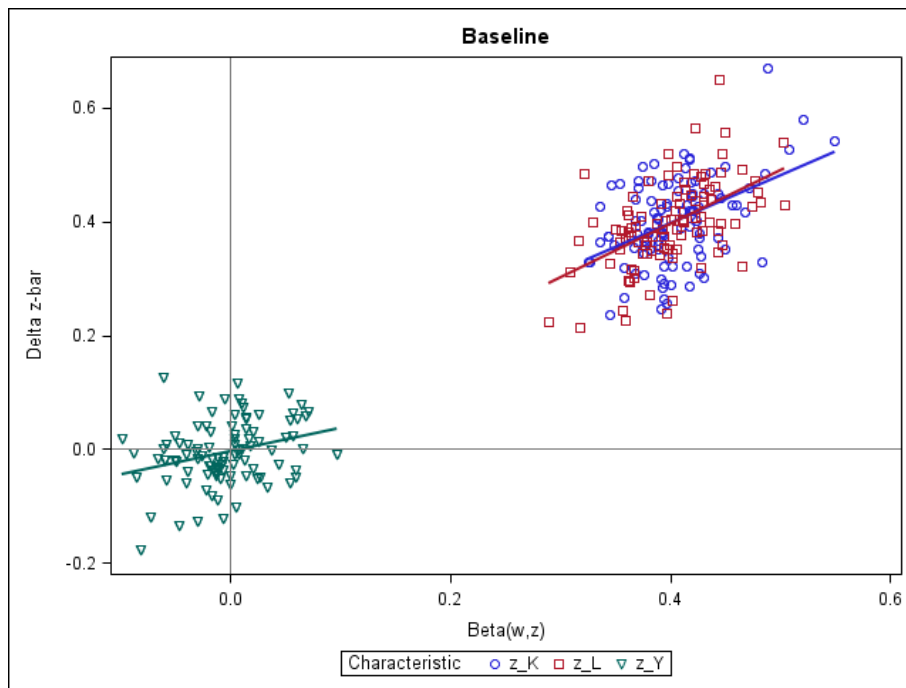

Figure 5: 100 baseline simulations

Figure 5 also includes three regression lines; one for each characteristic. All three have a positive slope indicating that there is a positive relationship between the gradient element, even for z_Y. Even though there is stabilising selection on z_Y we still observe that the change in mean characteristic takes the same sign as the selection gradient.

3.3. Simulations with correlation ($\rho_{YL} \neq 0$)

In this section we will present the results from assuming that $\rho_{YL} \neq 0$. Specifically, we will let the correlation approach unite in a stepwise manner which can be illustrated in a series of simulations. Adding correlation between z_Y, upon which there is stabilising selection, and z_L, upon which there is directional selection is expected to lead to confounded selection, as the selection mechanism will be pushing $\bar{z}_Y$ towards zero and $\bar{z}_L$ towards ever higher values, while at the level of the business unit they are positively correlated.

The results from specifying $\rho_{YL} = 0.1$ to $\rho_{YL} = 0.9$ in increments of 0.1 are shown in figure 6. This yields a total of 9 different parameterisations. Compared to figure 5 (where $\rho_{YL} = 0$) it does not make a big difference if instead $\rho_{YL} = 0.1$. However, as the correlation increases the result for the evolution of $\bar{z}_Y$ moves into the top left part of each plot. In this region selection is confounded: $\Delta \bar{z}_Y$ is positive while β_{w,z_Y} is negative meaning that the average of z_Y is increasing in the population but business units with high values of z_Y have relatively low growth. The reason why this can happen is that the z_Y is correlated with z_L upon which there is positive directional selection.
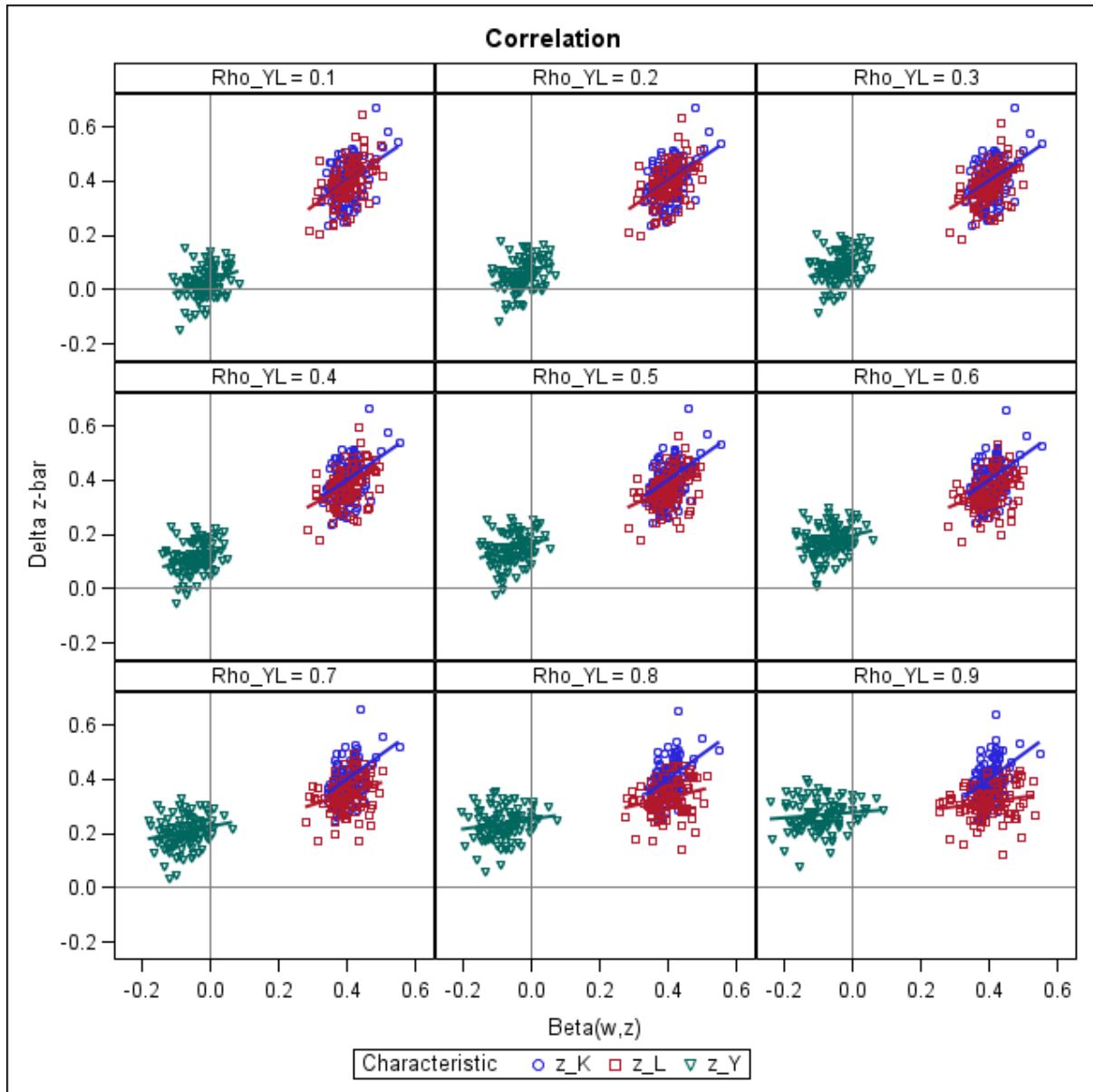

Figure 6: 9 different parameterisations for ρ_{YL}

It is also worth noticing what happens to the evolution of $\bar{z}_L$. Selection on this characteristic is direction and in all simulations both $\Delta \bar{z}_L$ and β_{w,z_L} are positive. However the regression line in the plots illustrating

14

the relationship between $\Delta\bar{z}_L$ and β_{w,z_L} becomes flatter when ρ_{YL} increases and when $\rho_{YL} = 0.9$ it is practically flat. This indicates that selection can be said to be confounded in a weaker sense: The sign on $\Delta\bar{z}_L$ allows us to infer the sign of β_{w,z_L}, but the magnitude of $\Delta\bar{z}_L$ contains no information regarding the magnitude of β_{w,z_L}.

The results show how observed evolutionary change may not convey the suspected information about optimality. As an example, consider a population where there is selection on at least two characteristics: the relative size of business units' administrative department and business units' total wage costs. On the former there is stabilising selection towards and optimum size and on the latter there is directional selection towards lower costs. If these traits correlate, so that more administration is correlated with higher labour costs, then we would not be able to infer the actual selection process from observed evolution: the size of the average business units' administration would be decreasing even below the optimal size. And the pace at which the average labour costs of business units decrease says very little about the importance of labour costs for business units' growth. A study focussing on the representative agent or assuming that business units all have the optimal configuration from rational choice and/or competitive pressure would miss this important point.

References

Agrawal, A. F. and Stinchcombe, J. R. (2009) "How much do genetic covariances alter the rate of adaptation?" *Proceedings of the Royal Society B*, 276:1183–1191.

Andersen, E. S. (2004) "Population thinking, Price's equation and the analysis of economic evolution". *Evolutionary and Institutional Economics Review,* 1(1):127–148.

Andersen, E. S. and Holm, J. R. (2014) "The signs of change in economic evolution: An analysis of directional, stabilizing and diversifying selection based on Price's equation". *Journal of Evolutionary Economics*, 24:291–316.

Baldwin, J. R. and Gu, W. (2006) "Competition, firm turnover and productivity growth". Statistics Canada, Economic Analysis Research Papers. 11F0027MIE No. 042.

Bottazzi, G., Dosi, G., Jacoby, N., Secchi, A. and Tamagni, F. (2010) "Corporate performance and market selection: Some comparative evidence". *Industrial and Corporate Change*, 19(6):1953–1996.

Brodie, E. D., Moore, A. J. and Janzen, F. J. (1995) "Visualizing and quantifying natural selection". *Trends in Ecology and Evolution*, 10:313–318.

Coad, A. (2007) "Testing the principle of 'growth of the fitter': The relationship between profits and firm growth". *Structural Change and Economic Dynamics*, 18(3):370–386.

Coad, A. and Teruel, M. (2013) "Inter-firm rivalry and firm growth: is there any evidence of direct competition between firms?" *Industrial and Corporate Change*, 22(2):397–425.

Conner, J. K. and Hartl, D. L. (2004) *A Primer of Ecological Genetics*. Sunderland, Mass.: Sinauer.

Endler, J. A. (1986) *Natural Selection in the Wild*. Princeton, N.J.: Princeton University Press.

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Foster, L., Haltiwanger, J. and Krizan, C. J. (1998) "Aggregate productivity growth: Lessons from microeconomic evidence". NBER Working Paper Series, 6803.

Frank, S. A. (2013) "Natural selection. VI. Partitioning the information in fitness and characters by path analysis", *Journal of Evolutionary Biology*, 26:457–471.

Hodgson, G. M. and Knudsen, T. (2006) "Dismantling Lamarckism: why descriptions of socio-economic evolution as Lamarckian are misleading". *Journal of Evolutionary Economics*, 16(4):343–366.

Holm, J. R. (2014) "The significance of structural transformation in productivity growth: How to account for different levels in economic selection". *Journal of Evolutionary Economics*, forthcoming.

Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M. et al. (2001) "The strength of phenotypic selection in natural populations". A*merican Naturalist*, 157:245–261.

Kingsolver, J. G. and Pfennig, D. W. (2007) "Patterns and power of phenotypic selection in nature". *BioScience*, 57:561–572.

Lande, R. and Arnold, S. J. (1983) "The measurement of selection on correlated characters". *Evolution*, 37: 1212–1226.

Metcalfe, J. S. (1994) "Competition, Fisher's principle and increasing returns in the selection process". *Journal of Evolutionary Economics,* 4(4):327–346.

Metcalfe, J. S. (1997) "Labour Markets and Competition as an Evolutionary Process". Aretis, P., Palma, G. and Sawyer, M (eds) *Markets, Unemployment and Economic Policy. Essays in Honour of Geoff Harcourt*, vol. 2, pp. 328–343. London and New York: Routledge.

Metcalfe, J. S. and Ramlogan, R. (2006). "Creative destruction and the measurement of productivity change". *Revue de l'OFCE*, 97:373–397.

Metcalfe, J. S. (2010) "Investment in the Theory of Evolutionary Competition". *FINNOV Discussion Paper* D3.8

Nelson, R. R. and Winter, S. G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge, Mass. and London: Harvard University Press.

Price, G. R. (1970) "Selection and covariance". *Nature*, 227:520–521.

Price, G. R. (1995) "The nature of selection". *Journal of Theoretical Biology*, 175:389–396.

Rice, S. H. (2004) *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sunderland, Mass.: Sinauer.