

# Determinação dos fatores econométricos de impacto na produção cafeeira em Minas Gerais

Determination of econometric factors impacting coffee production in Minas Gerais

Determinación de factores econométricos que impactan la producción de café en Minas Gerais

Recebido: 16/04/2022 | Revisado: 24/04/2022 | Aceito: 29/04/2022 | Publicado: 01/05/2022

**Luciano Ribeiro Galvão**

ORCID: <https://orcid.org/0000-0001-7930-1049>

Universidade de São Paulo, Brasil

E-mail: [lucianogalvao@usp.br](mailto:lucianogalvao@usp.br)

**Júlio Sílvio de Sousa Bueno Filho**

ORCID: <https://orcid.org/0000-0002-3243-0959>

Universidade Federal de Lavras, Brasil

E-mail: [jssbueno@ufla.br](mailto:jssbueno@ufla.br)

**Andrezza Kellen Alves Pamplona**

ORCID: <https://orcid.org/0000-0002-8809-0172>

Instituto Federal do Triângulo Mineiro, Brasil

E-mail: [andrezzapamplona@iftm.edu.br](mailto:andrezzapamplona@iftm.edu.br)

**Caio Peixoto Chain**

ORCID: <https://orcid.org/0000-0002-9520-1547>

Universidade Federal Rural do Rio de Janeiro, Brasil

E-mail: [caiopeixotochain@gmail.com](mailto:caiopeixotochain@gmail.com)

## Resumo

O estudo das relações existentes entre os componentes econométricos da produção de café, estimadas por correlações, por exemplo, é de grande relevância. Estas métricas fornecem informações úteis para o processo de decisão na cadeia produtiva dessa *commodity*. Todavia, a quantificação e a interpretação da magnitude das correlações não implicam efeitos diretos e indiretos aplicáveis a realidade do agronegócio. Nesse contexto, a análise de trilha apresenta-se como uma alternativa viável. O objetivo deste trabalho foi, através da análise de trilha, determinar os efeitos diretos e indiretos de componentes econométricos sobre a produção de café. Os dados utilizados são de municípios mineiros produtores de café, no período de 2008 a 2013, nos quais foi observada a produção de café como a variável básica (dependente) e como variáveis independentes (explicativas) a área colhida do grão, a idade média dos trabalhadores no campo, a remuneração média dos trabalhadores na atividade, o preço pago pelo produto e o número de propriedades produtoras por município. Observou-se forte variação com efeitos de ano e de município, sendo que o restante da explicação se refere mais às variáveis primárias (área plantada e número de propriedades do município - coeficientes de trilha 0,38 e 0,05, respectivamente). As demais variáveis interferem indiretamente, por meio da modificação dessas duas. A análise de trilha mostrou-se útil em elucidar parte da variabilidade da cadeia produtiva do café e poderá ser usada como auxiliar na tomada de decisões de negócio do setor.

**Palavras-chave:** Café; Análise de trilha; Correlação; Componentes econométricos.

## Abstract

The study of the relationships between the econometric components of coffee production, estimated by correlations, for example, is of great relevance. These metrics provide useful information for the decision process in the production chain of this commodity. However, the quantification and interpretation of the correlation's magnitude do not imply direct and indirect effects applicable to the agribusiness reality. In this context, trail analysis presents as a viable alternative. The objective of this work was, through trail analysis, to determine the direct and indirect effects of econometric components on coffee production. The data used are from coffee producing municipalities in Minas Gerais, in the period from 2008 to 2013, in which coffee production was observed as the basic (dependent) variable and as independent (explanatory) variables the harvested area of the grain, the average age of the workers in the field, the average remuneration of workers in the activity, the price paid for the product and the number of producing properties per municipality. There was a strong variation with year and municipality effects, with the rest of the explanation referring more to the primary variables (planted area and number of properties in the municipality), trail coefficients 0.38 and 0.05, respectively. The other variables interfere indirectly, through the modification of both. Path analysis proved to be useful in elucidating part of the coffee production chain variability and can be used as an aid in making business decisions in the sector.

**Keywords:** Coffee; Path analysis; Correlation; Econometric components.

## Resumen

El estudio de las relaciones entre los componentes econométricos de la producción de café, estimados por correlaciones, es de gran relevancia. Ellas brindan información útil para el proceso de decisión en la cadena productiva de este producto básico. La cuantificación e interpretación de la magnitud de las correlaciones no implican efectos directos e indirectos aplicables a la realidad de los agronegocios. Así, el análisis de senderos se presenta como una alternativa viable. Este trabajo tuvo como objetivo, a través del análisis de senderos, determinar los efectos directos e indirectos de los componentes econométricos en la producción de café. Los datos son de municipios productores de café de Minas Gerais, período de 2008 a 2013, teniendo como variable básica (dependiente) la producción de café y como variables independientes (explicativas) el área cosechada del grano, la edad promedio de los trabajadores del campo, la remuneración promedio de los trabajadores de la actividad, el precio pagado por el producto y el número de propiedades productoras por municipio. Hubo una fuerte variación con los efectos de año y municipio, con el resto de la explicación refiriéndose más a las variables primarias (área sembrada y número de propiedades en el municipio), coeficientes de rastro 0.38 y 0.05, respectivamente. Las otras variables interfieren indirectamente, a través de la modificación de estas dos. El análisis de trayectoria demostró ser útil para dilucidar parte de la variabilidad de la cadena productiva del café y puede utilizarse como ayuda para la toma de decisiones comerciales en el sector.

**Palabras clave:** Café; Análisis de senderos; Correlación; Componentes econométricos.

## 1. Introdução

O café é uma das bebidas mais apreciadas e consumidas em todo o mundo (Abic, 2021). O café pertence ao gênero *Coffea* e possui duas espécies mais comercializadas, *Coffea arabica* e *Coffea canephora*, conhecidos como café arábica e café conilon, respectivamente. Assim como acontece com a produção mundial, o café arábica corresponde a mais de 70% da produção brasileira devido à sua qualidade superior de bebida, enquanto o café conilon tem características mais robustas.

Não se sabe exatamente a data de ingresso da cultura de café no Brasil, mas é certo que isso ocorreu no século XVIII e se tornou uma das matérias-primas agrícolas mais importantes do país. Atualmente, o café é o quinto item agrícola mais exportado do país (Conab, 2015). Sua produção e comercialização ainda hoje são de grande importância para a balança comercial brasileira, sendo o Brasil o maior produtor e exportador mundial, com produção de mais de 63 milhões de sacas beneficiadas em 2020 (Conab, 2020; 2021)

De acordo com o levantamento de safra da Companhia Nacional de Abastecimento – Conab (2021), Minas Gerais ocupa o posto de maior Estado produtor de café, com mais de 55% da produção nacional, sendo o café arábica o cultivo predominante. No ano de 2021, foram produzidas cerca de 34,6 milhões de sacas de café no estado.

Quando se estuda mercadorias do setor agrícola, a produção é um dos focos mais importantes, pois impulsiona o mercado e mantém os produtores ativos. Entretanto, a produtividade é um caráter complexo e resultante da expressão e associação de diferentes componentes (Carvalho et al., 2002), o que torna necessário entender e estudar o grau de associação entre esses componentes e a produtividade, por meio de estudos de correlações.

Apesar de auxiliar no entendimento de um caráter complexo, como a produtividade, a análise da correlação não permite determinar os efeitos diretos e indiretos dos componentes que a compõem. Assim, não é possível constatar se sua estimativa foi estabelecida por verdadeiras relações de causa e efeito e a interpretação direta de suas magnitudes pode resultar em equívocos na estratégia de decisão mercadológica, pois uma alta correlação entre dois caracteres pode ser resultado do efeito de um terceiro componente, ou de um grupo de componentes, sobre eles. (Wright, 1921)

Com o objetivo de compreender melhor as causas envolvidas nas associações de componentes, Wright (1921) propôs um método de desdobramento das correlações, estimando efeitos diretos e indiretos das variáveis explicativas (independentes) sobre uma variável básica (dependente), denominado análise de trilha. Embora a correlação seja uma característica intrínseca a dois componentes em dada condição experimental, sua decomposição depende do conjunto de componentes estudados que, normalmente, são avaliados pelo conhecimento prévio de suas importâncias e de possíveis inter-relações expressas em diagramas de trilha (Cruz et al., 2004).

Em seu trabalho, Souza (2014) apresenta recursos estatísticos para a inferência causal usando a análise de trilha. Desse modo, essa técnica tem como característica exprimir a associação e, caso exista, a sua magnitude, tanto entre as variáveis independentes e a dependente, como também entre as próprias variáveis independentes.

Um aspecto importante sobre a análise de trilha é o artifício visual, que é de grande auxílio para o leitor, uma vez que o diagrama de trilha, ou diagrama causal, segue um padrão e esse, quando bem executado, traz em si toda informação necessária sobre o estudo. Desse modo, quando o coeficiente de trilha possui alto ajuste entre variável dependente e independente em questão, sua representação no diagrama de trilha é feita com setas, unidirecionais, de traço contínuo (De Carvalho, 1994). Para coeficientes de trilha de baixo valor de interação da variável independente com a variável dependente, a representação é feita com setas tracejadas (Li, 1975). Já entre as próprias variáveis independentes, a correlação é mútua, representada por arcos de dois sentidos, conectando-as (Kline, 2011).

A análise de trilha, apesar de vastamente aplicada no melhoramento genético de várias culturas como o milho e os processos de pós-colheita (De Oliveira et al., 2021), o algodão (De Souza, 2021) e o próprio o café. Nessa técnica de seleção são associadas variáveis de produtividade e linhagem (Severino et al., 2008; Teixeira et al., 2012) ou adicionando métodos de cultivo e desempenho de progênies (Martinez et al., 2008) com produção. O estudo de trilha ainda encontra pouca inserção em análises de componentes econométricos. O presente trabalho visa propor o uso dessa importante ferramenta de análise estatística para a determinação dos componentes produtivos importantes na atividade cafeeira.

## 2. Metodologia

Esse trabalho fez uso de dados do setor cafeeiro do Estado de Minas Gerais, obtidos de maneira secundária nos sites oficiais da Relação Anual de Informações Sociais – RAIS (Brasil, 2015a) e do Cadastro Geral de Empregados e Desempregados – CAGED (Brasil, 2015b). Foi observado o período de 2008 a 2013 em que havia, segundo os dados obtidos, 605 municípios produtores do Estado de Minas Gerais. Os dados foram organizados para a análise com o *software* R (R Core Team, 2015). A escolha dos autores se deu por este recorte ter amostras de sazonalidade suficientes para que a estimação seja feita e, durante este período (5 anos), as bases estavam completas e disponíveis com pouca incidência de valores faltantes, por exemplo.

Para esse trabalho foram escolhidas seis variáveis, sendo a variável dependente Produção e as cinco independentes sendo: Área colhida do grão, Idade média dos trabalhadores no campo, Remuneração média dos trabalhadores na atividade, Preço pago pelo produto e Número de propriedades produtoras por município. Utilizou-se um modelo aditivo para a execução de análises univariadas. Cada variável do estudo foi comparada aos fatores Ano e Cidade.

Verificada a normalidade dos resíduos usando o teste de Shapiro-Wilk, conduziu-se a análise multivariada envolvendo as variáveis do estudo em função de Cidade e Ano. Para tal, utilizou-se o teste de Hotelling-Lawley, testando a hipótese de ausência de efeitos de Cidade e Ano.

Para a obtenção das correlações entre as variáveis independentes, sem a interferência da variável básica (dependente), utilizou-se a técnica de correlação múltipla, que é tida como uma extensão direta do método de Pearson (Montgomery et al., 2012) para correlações simples. Nesta etapa, devido ao grande número de observações por variável, não ignoramos as correlações de baixo valor no cálculo dos coeficientes de trilha.

Além das correlações lineares, fez-se uso de correlações parciais, onde os efeitos de associação entre variáveis levam em conta a participação de uma ou mais sobre as duas inicialmente correlacionadas, nesta condição pode-se estabelecer a significância das trilhas calculando os efeitos das variáveis dependentes entre si e com relação à variável dependente. Assim, as correlações foram testadas duas a duas com o controle de cada uma das variáveis que não participaram diretamente nas correlações, e o resultado foi submetido ao teste de hipótese, segundo propõe Carvalho (et al., 2004). A partir das matrizes de

correlação foram calculados os coeficientes de trilha (Wright, 1934) e construídos os diagramas de trilha (Li, 1956), onde as variáveis independentes se definem primárias e secundárias. Procedeu-se à análise causal de trilha usando a metodologia apresentada por Li (1956) e seguindo os passos descritos em Souza (2014).

### 3. Resultados e Discussão

De acordo com o teste de Shapiro-Wilk, os dados são normalmente distribuídos. Logo, apresentamos o resumo da análise univariada para as variáveis de interesse em função do Ano e da Cidade (Tabela 1), onde observa-se que o fator Ano foi significativo para as variáveis Produção, Idade do trabalhador, Área colhida, Número de propriedades produtoras, Remuneração do trabalhador e Valor pago pelo produto. A sazonalidade naturalmente observada no café (bienal) foi considerada. O fator Cidade foi significativo para as seguintes variáveis: Produção, Idade do trabalhador, Área colhida, Número de propriedades produtoras e Remuneração do trabalhador, ou seja, existem variações importantes de município em município para estas variáveis. Para a variável Valor pago pelo produto, o fator Cidade foi não significativo, assim sendo, não se rejeitou a hipótese nula de que o valor que se paga pelo café varia em função da cidade onde é produzido podendo ser atribuído a questões de frete e valor de praça, que de fato, existem como constante na rotina do produtor. Esse resultado corrobora a noção prévia dos autores, pois o preço do café é cotado em bolsa e não sofre variação (significativa) nas praças regionais onde é, de fato, comercializado. A variável Valor pago ao produtor varia somente em função do ano, já que os motivos desta volatilidade de preços estão relacionados a diversos fatores, por exemplo, não sistêmicos - clima - ou ligados ao próprio mercado ou a alterações no cenário econômico mundial.

Por haver apenas uma observação por ano e cidade, não se investigou a interação entre o ano estudado e a cidade onde os dados foram extraídos.

**Tabela 1:** Análise univariada das variáveis em estudo em função de Ano e Cidade.

		Produção	Área colhida	Número de propriedades	Remuneração do trabalhador	Idade do trabalhador	Valor pago pelo produto
Fator	GL	Valor F					
Ano	5	81,1***	509,7***	360,4***	1,2*	21,3***	0,9*
Cidade	604	27,9***	4,5***	6,3***	2,7*	11,7***	2,5 <sup>NS</sup>
GL resíduo		2983	2983	1733	1735	1733	2973

Valores de significância pelo teste de FISHER: \*\*\* 0.001; \*\*0.01; \*0.1. Fonte: Autores (2022).

Uma vez realizada a análise univariada, seguiu-se para a análise multivariada confirmando-se os resultados obtidos com análises realizadas anteriormente. Na Tabela 2 apresenta-se a análise multivariada, na qual se observa forte significância de ambas as variáveis (Ano e Cidade) com alta significância ( $p < 0,01$ ) com relação todas as variáveis estudadas.

**Tabela 2:** Análise multivariada e teste de Hotelling-Lawley.

Fator	GL	Hotelling-Lawley	Probabilidade de Significância
Cidade	401	161,609	< 0,001
Ano	5	0,193	< 0,001
Resíduos	1721		

Fonte: Autores (2022).

A partir deste modelo multivariado, obteve-se a matriz residual das variáveis estudadas (Tabela 3). Os resíduos do modelo em função de Ano e Cidade são usados para isolar os efeitos desses fatores sobre as demais variáveis, sem que se tenha perda de informação. Assim, a matriz residual é usada para testar as correlações entre as variáveis, independentemente das variações ocorridas em Ano e Cidade. Para isso, definiram-se essas correlações entre as variáveis explicativas e a variável básica (Tabela 3b). Essas correlações serão usadas nos sistemas matriciais com as correlações entre as próprias variáveis independentes (Tabela 3a).

**Tabela 3:** Tabela de correlações residuais: (a) entre as variáveis independentes; (b) entre as variáveis independentes e a variável dependente Produção.

Variáveis em estudo	Idade	Número de Propriedades	Remuneração	Valor	Área
Idade	1				
Número de Propriedades	- 0,07**	1			
(a) Remuneração	0,05*	0,00	1		
Valor	- 0,01	0,01	- 0,01	1	
Área	- 0,01	0,16**	0,01	- 0,01	1
(b) Produção	0,01	0,11**	0,01	0,01	0,39**

Valores de significância pelo teste unilateral de t com 603 graus de liberdade: \*0,05; \*\*0,01. Fonte: Autores (2022).

Ainda na Tabela 3, é possível observar que as maiores correlações positivas com a variável Produção estão entre Área (0,39) e Número de propriedades (0,11), enquanto a correlação entre variáveis independentes de maior valor é entre Número de propriedades e Área colhida (0,16). Mesmo as correlações de menores intensidades são interessantes ao estudo, uma vez que o número de observações é grande (mais de 2.000) para cada uma das seis variáveis. Verificou-se que a correlação multivariada é mais importante entre algumas das variáveis em estudo.

O fator Idade média dos trabalhadores obteve valores negativos nas correlações com Número de propriedades, Valor pago pelo produto e Área colhida, e valores positivos para Produção e Remuneração dos trabalhadores, ou seja, trabalhadores mais jovens recebem, em média, menos que os mais velhos e, do mesmo modo, em regiões onde se concentram grande número de propriedades observa-se uma tendência de trabalhadores mais jovens na atividade cafeeicultora.

O Número de propriedades gerou um coeficiente positivo quando correlacionado com todas as variáveis em estudo, com alta significância para Produção e Área colhida, indicando que a produção é maior em municípios com muitas propriedades e que têm grandes áreas de lavoura permanente. Observa-se que a correlação mais alta (0,39) foi entre Área

colhida e Produção, mostrando que quanto maior a área destinada a lavouras de café, maior será sua produção independentemente das variações naturalmente ocorridas entre Ano e Cidade.

A seguir apresenta-se a matriz de correlações parciais para modelos de trilha. O grau de associação entre as variáveis independentes com a dependente é influenciado pela contribuição de outras variáveis explicativas sobre a variável básica. Desse modo, uma correlação parcial significativa mostra a provável intensidade do efeito de trilha entre estas variáveis.

**Tabela 4:** Correlações parciais entre as variáveis independentes:

Variáveis em estudo	Idade	Número de Propriedades	Remuneração	Valor	Área	Produção
Idade	1					
Número de Propriedades	- 0,071**	1				
Remuneração	0,049*	0,003	1			
Valor	- 0,013	0,011	- 0,005	1		
Área	- 0,001	0,122**	0,001	- 0,011	1	
Produção	0,008	0,05**	0,011	0,001	0,376**	1

Valores de significância pelo teste unilateral de *t* com 2127 graus de liberdade: \*0,05; \*\*0,01. Fonte: Autores (2022).

De posse das correlações parciais, obtêm-se, então, os valores relativos aos coeficientes de trilha, dispostos na Tabela 4, sendo que os efeitos diretos estão na diagonal da matriz e os efeitos indiretos nas demais linhas e colunas. Ainda que se possam observar valores baixos (o que exclui a necessidade de interpretação prática) deve-se, porém, observar apenas a aplicação do método, que foi o foco desse trabalho. Desse modo, a matriz de coeficientes de trilha está disposta no diagrama de trilha apresentado na Figura 1.

Para compor corretamente o diagrama de trilha, deve-se também gerar mais duas tabelas de coeficientes, uma para cada variável, que foi em um segundo momento, determinada como primária e definido seu valor de coeficiente de trilha com relação às demais (secundárias). O estudo aprofundado das variáveis secundárias é de pouco interesse nesse trabalho, pois, como é esperado, elas possuem baixos valores de correlação com a variável básica. Assim, seguem a Tabela 5 que é referente a variável primária área (tida aqui como básica) e a Tabela 6, na qual observamos os valores de coeficiente de trilha para a variável primária Número de propriedades e o efeito das demais variáveis (secundárias) sobre ela.

**Tabela 5:** Matriz de efeitos diretos (diagonal) e indiretos dos coeficientes de trilha para as variáveis:

Variáveis	Idade	Número de Propriedades	Remuneração	Valor	Área
Idade	0,0070	- 0,0040	0,0003	- 0,0004	- 0,0038
Número de Propriedades	- 0,0004	0,0500	0,0000	0,0004	0,0620
Remuneração	- 0,0003	0,0000	0,0060	- 0,0004	0,0040
Valor	- 0,0007	0,0005	- 0,0006	- 0,0004	- 0,0040
Área	0,0007	0,0070	0,0006	- 0,0004	0,3800

Resíduo quadrático = 0,8455

Fonte: Autores (2022).

**Tabela 6:** Matriz de efeitos diretos (diagonal) e indiretos dos coeficientes de trilha para as variáveis secundárias em função de Área plantada:

Variáveis	Idade	Remuneração	Valor
Idade	-0,01	0,0005	0,0001
Remuneração	-0,0005	0,01	0,0001
Valor	0,0001	-0,0001	-0,01

Resíduo quadrático = 0,99

Fonte: Autores (2022).

**Tabela 7:** Matriz de efeitos diretos (diagonal) e indiretos dos coeficientes de trilha para as variáveis secundárias em função do Número de propriedades:

Variáveis	Idade	Remuneração	Valor
Idade	-0,07	0,00001	0,00001
Remuneração	-0,003	0,00003	0,00001
Valor	0,0007	-0,0003	-0,00001

Resíduo quadrático = 0,99

Fonte: Autores (2022).

Para melhor elucidar, os valores que são exibidos no grafo unidirecional acíclico, a Tabela 8 traz os valores obtidos com a estimação de trilha, confrontando-os com a correlação direta. Essa comparação é tão somente para efeito de verificação, uma vez que a trilha e seus efeitos indiretos somados devem ser exatamente o valor de correlação. A diferença está apenas que no grafo unidirecional acíclico elas são dispostas de forma que explicam sua direção de ocorrência e intensidade de correlação.

**Tabela 8:** Coeficientes de correlação de Pearson e coeficientes de trilha das variáveis primárias.

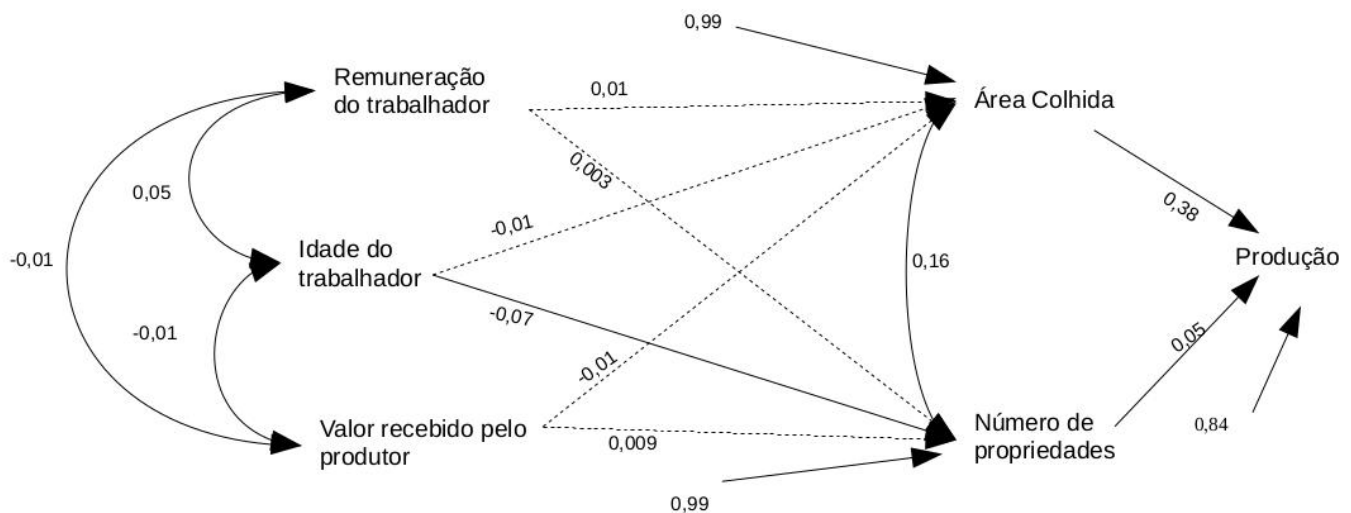
Variáveis primárias	Vias de associação	Estimador	Estimativa
Área colhida	Efeito direto sobre produção	r01	0,38
	Efeito indireto via Número de propriedades	r02r12	0,01
Total		r01	0,39
Número de propriedades	Efeito direto sobre produção	r02	0,05
	Efeito indireto via Área colhida	r01r12	0,06
Total		r02	0,11
Efeito residual ( $r\epsilon 0$ )			0,84

Fonte: Autores.

A Figura 1 apresenta o diagrama de trilha, expondo as variáveis primárias definidas, a priori, por meio de revisão de literatura e conhecimento prévio do assunto. Para esse trabalho, definiu-se Área colhida e Número total de propriedades produtoras como primárias e as demais como secundárias, de modo que os coeficientes de trilha foram maiores para as primárias, e baixos para as secundárias.

As linhas de traço contínuo representam a relação unidirecional de trilha significativa com base na análise da correlação parcial. Desse modo, as variáveis Número de propriedades e Área colhida possuem seta contínua no sentido de produção. A variável Idade do trabalhador possui seta contínua para Número de propriedades, demonstrando uma relação unidirecional e consistente de trilha. As demais variáveis, que são relacionadas com linhas tracejadas, possuem efeitos de correlação parcial com significância superior a 5%. Dessa forma, não foram consideradas para efeito direto na composição da trilha.

**Figura 1:** Diagrama de trilha: Área colhida (1) e Número de propriedades (2) são variáveis primárias; Remuneração do trabalhador (3), Idade do trabalhador (4) e Valor pago pelo produto (5) são variáveis secundárias. O resíduo quadrático da estimação de trilha é representado por “ $\epsilon_0$ ”.



Fonte: Autores (2022).

As variáveis são fortemente influenciadas pelos fatores Ano e Cidade, o que afetou os coeficientes de trilha quando se isolam esses fatores, expresso no alto valor para resíduo de estimação (0,84). De fato, a cada ano o valor do grão é cotado em bolsa e apresenta grande oscilação, além do regime de chuvas e políticas de apoio governamental, fatores que também influenciam a produção e alteram os preços anuais do café. Com relação às cidades, a própria cultura local define o volume e o modo de produção. Entende-se que em cidades tradicionalmente produtoras, o cultivo do café segue uma dinâmica diferenciada quando comparada a locais onde a atividade é recente. Ainda, propriedades situadas em cidades próximas a grandes indústrias de beneficiamento, vias de escoamento da produção e centros de pesquisa e apoio ao produtor irão certamente gerar grande variação sobre as variáveis estudadas.

#### 4. Conclusão

O uso da análise de trilha para se trabalhar com componentes econométricos do café revelou pouca dependência entre as variáveis explicativas secundárias, este sendo maior para as variáveis explicativas primárias. A dependência do Número de propriedade de uma região para o alto potencial produtivo é facilmente entendida, pois cooperativas e centros de beneficiamentos são fortes propulsores de crescimento da atividade numa região.

Nossos achados sugerem que estudos posteriores em que se considere número maior de variáveis primárias, elencadas por importância à priori, podem apresentar melhores resultados na interpretação. O número de variáveis e sua ordem definidas pelo pesquisador ajudará a elucidar os efeitos dessas na cadeia produtiva do café, por meio da análise de trilha.



## Referências

- Abic. (2021). Associação Brasileira da Indústria de Café. *Indicadores da Indústria de Café*. <https://estatisticas.abic.com.br/estatisticas/indicadores-da-industria/indicadores-da-industria-de-cafe-2021>.
- Brasil. (2015a). Ministério do Trabalho e Emprego. *Cadastro Geral de Empregados e Desempregados: CAGED estatístico 2012*. <http://bi.mte.gov.br/bgcaged/rais.php>.
- Brasil. (2015b). Ministério do Trabalho e Emprego. *Relação Anual de Informações Sociais: RAIS vínculos 2012*. <http://bi.mte.gov.br/bgcaged/rais.php>.
- De Carvalho, S. P. (1994). *Métodos alternativos de estimação de coeficientes de trilha e índices de seleção, sob multicolinearidade*. (Tese de Doutorado). Universidade Federal de Viçosa. Viçosa.
- Carvalho, C. G. P., Arias, C. A. A., De Toledo, J. F. F., De Oliveira, M. F., & Vello, N. A. (2002). Correlações e análise de trilha em linhagens de soja semeadas em diferentes épocas. *Pesquisa Agropecuária Brasileira*, 37(3), 311-320.
- Carvalho, F. I. F., Lorencetti, C., & Benin, G. (2004). *Estimativas e implicações da correlação no melhoramento vegetal*. Pelotas: UFPel.
- Conab. (2021). Companhia Nacional de Abastecimento. *Acompanhamento da Safra Brasileira de Café*. 8(1) – Safra 2021 – Primeiro levantamento. Brasília, 1-71.
- Conab. (2022). Companhia Nacional de Abastecimento. *Acompanhamento da Safra Brasileira de Café*. 9(1) – Safra 2022 – Primeiro Levantamento. Brasília, 1-60.
- Cruz, C. D. et al. (2004). Análise de trilha. In: Cruz, C. D., Regazzi, A. J., & Carneiro, P. C. S. (Ed.). *Modelos biométricos aplicados ao melhoramento genético*. Viçosa, 180-193.
- De Oliveira, R. M., de Oliveira, R. A., Neves, J. C. L., de Oliveira, E. M., Botelho, M. E., & de Oliveira, J. T. (2021). Relação entre índice de colheita do milho e eficiências de fósforo por meio de análise de trilha. *Research, Society and Development*, 10(2), e13510212358-e13510212358.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford Press.
- Li, C. C. (1956). The concept of path coefficient and its impact on population genetics. *Biometrics*, 12(2), 190-210.
- Li, C. C. (1975). *Path Analysis: a primer*. Pacific Grove: The Boxwood Press.
- Lira, S. A. *Análise de correlação: análise teórica e de construídos coeficientes com aplicações*. (Dissertação de Mestrado). Universidade Federal do Paraná. Curitiba.
- Martinez, H. E. P., Augusto, H. S., Cruz, C. D., Pedrosa, A. W., & Sampaio, N. F. (2008). <b>Crescimento vegetativo de cultivares de café (<em>Coffea arabica</em> L.) e sua correlação com a produção em espaçamentos adensados</b> - 10.4025/actasciagron.v29i4.404. *Acta Scientiarum, Agronomy*, 29(4), 481-489. <https://doi.org/10.4025/actasciagron.v29i4.404>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis*. J. Wiley & Sons.
- R Core Team. (2022). R: a language and environment for statistical computing. *Vienna: R Foundation for Statistical Computing*, 2015. <http://www.R-project.org/>.
- Severino, L. S., Sakiyama, N. S., Pereira, A. A., Miranda, G. V., Zambolim, L., & Barros, U. V. (2008). Associações da produtividade com outras características agrônômicas de café (<em>Coffea arabica</em> L. “Catimor”). *Acta Scientiarum, Agronomy*, 24, 1467-71. 10.4025/actasciagron.v24i0.2405
- De Souza, L. A. (2021). *Correlação genética e análise de trilha em genótipos de algodoeiro de fibra colorida*. 21 p. Trabalho de Conclusão de Curso (Graduação em Agronomia) - Universidade Federal de Uberlândia. Uberlândia.
- Souza, T. V., Ribeiro, C. M., Scalon, J. D., & Guedes, F. L. (2014). Relações entre componentes de rendimento e características morfológicas de milho. *Magistra*, 26(4), 495-506.
- Teixeira, A. L., Gonçalves, F. M. A., Rezende, J. C. D., Carvalho, S. P. D., Pereira, A. A., Moraes, B. F. X. D., & Teixeira, L. G. V. (2012). Seleção precoce para produção de grãos em café arábica pela avaliação de caracteres morfológicos. *Pesquisa agropecuária brasileira*, 47, 1110-1117.
- Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20(7), 557-585.