

Tópicos em História Econômica

Notas de Aula - 2025

1 Introdução

Nas últimas duas décadas, uma nova literatura empírica surgiu para investigar o impacto de eventos históricos no desenvolvimento econômico. Três linhas de pesquisa iniciadas nos anos 1990 foram fundamentais: Engerman e Sokoloff (1997, 2002) analisaram o papel das dotações de fatores e do colonialismo no desenvolvimento das Américas; Acemoglu, Johnson e Robinson (2001, 2002) exploraram a relação entre instituições históricas e crescimento econômico de longo prazo; e La Porta et al. (1997, 1998) examinaram o impacto das instituições jurídicas coloniais na proteção de investidores e no desenvolvimento financeiro.

Esses estudos pioneiros destacaram o papel do colonialismo e das instituições no desenvolvimento de longo prazo, impulsionando uma ampla literatura empírica. Inicialmente focada em correlações entre eventos históricos e desenvolvimento, essa pesquisa evoluiu com novos dados históricos e metodologias econométricas mais robustas para estabelecer relações causais. Além disso, essa literatura passou a investigar os mecanismos pelos quais eventos históricos continuam a afetar o presente, destacando a importância de instituições, normas culturais e conhecimento tecnológico. Outro avanço foi a análise da interação entre geografia e história, mostrando que a geografia influencia o desenvolvimento principalmente por meio de seus efeitos históricos.

Este ensaio apresenta uma análise crítica da literatura, iniciando pelos estudos fundamentais e suas contribuições, passando pelas pesquisas sobre “**tema**” (e.g., *O Papel das Instituições e Origens Coloniais no Crescimento Econômico*) e examinando os mecanismos pelos quais a história impacta o desenvolvimento. Por fim, apresenta as direções futuras para essa área de pesquisa.

2 História do Desenvolvimento Econômico

Origens Históricas do Desenvolvimento Econômico. A literatura que relaciona história e desenvolvimento econômico tem suas raízes em três linhas de pesquisa interligadas: Acemoglu et al. (2001), La Porta et al. (1997, 1998) e Engerman e Sokoloff (1997, 2002). Esses estudos analisam o impacto duradouro da colonização europeia nos antigos territórios coloniais, destacando como o

domínio colonial moldou instituições domésticas que persistiram após a independência.

Embora compartilhem a visão de que as instituições são determinantes cruciais para o desenvolvimento econômico de longo prazo, esses estudos diferem na ênfase sobre quais aspectos coloniais foram mais relevantes. Para La Porta et al., o fator decisivo foi a identidade do colonizador, que influenciou a adoção de sistemas jurídicos de common law ou civil law, com impactos duradouros no desenvolvimento econômico. Eles observam que as ex-colônias britânicas adotaram o common law, enquanto as colônias de países como Espanha, França e Portugal adotaram o civil law. Usando essa variação histórica, os autores investigam o efeito da proteção aos direitos dos investidores no desenvolvimento financeiro. Como a origem legal das ex-colônias é exógena, ela pode ser usada como um instrumento para estimar esse impacto. As análises indicam que países com sistemas de common law oferecem melhor proteção aos investidores, o que resulta em mercados de dívida e ações mais desenvolvidos.

Engerman e Sokoloff, assim como Acemoglu et al., argumentam que as características da região colonizada foram determinantes no longo prazo. Acemoglu et al. ressaltam que o ambiente inicial de doenças teve um impacto significativo na formação de direitos de propriedade seguros, influenciando o crescimento econômico. Em colônias com baixa mortalidade, os europeus se estabeleceram e implementaram instituições que protegiam esses direitos, fomentando o crescimento econômico. Por outro lado, em regiões com alta mortalidade, foram adotadas instituições extrativistas voltadas para a exploração. Para estimar o efeito causal dessas instituições na renda per capita atual, os autores utilizam a taxa de mortalidade dos colonos como um instrumento. Os resultados indicam que uma mortalidade inicial elevada está associada à presença de instituições de menor qualidade hoje, e que estas instituições exercem um efeito positivo sobre a renda.

Esse fenômeno pode ser relacionado ao “reversal of fortunes” identificado em um estudo posterior (Acemoglu, Johnson, e Robinson, 2002), onde os locais mais ricos do mundo em 1500 se tornaram alguns dos mais pobres até o final do século XX. As regiões com recursos naturais facilmente exploráveis foram aquelas que sofreram com instituições coloniais deficientes, associadas a um fraco desenvolvimento econômico a longo prazo. De acordo com os autores, o reverso da fortuna ocorreu entre o final do século XVIII e o início do século XIX e está intimamente ligado ao processo de industrialização. Segundo eles, a interação entre instituições e a oportunidade de industrializar desempenhou um papel central no desenvolvimento a longo prazo das antigas colônias.

No entanto, alguns autores contestam o argumento da origem colonial das instituições de AJR.

Albouy (2004, 2006) critica a mensuração das taxas de mortalidade de colonos utilizadas, observando que na amostra original, apenas 28 dos 64 países possuem dados reais, enquanto os outros 36 são baseados em conjecturas sobre condições de saúde. Ao excluir essas observações, ele conclui que o impacto da mortalidade de colonos no risco de expropriação (proxy para instituições) é muito menor. Logo, a taxa de mortalidade dos colonos seria um instrumento fraco para analisar o efeito causal das instituições sobre o crescimento. Glaeser et al. (2004) acrescentam que a utilização da mortalidade de colonos europeus como instrumento para instituições viola a hipótese de exogeneidade, argumentando que os colonizadores trouxeram habilidades e capital humano que persistiram ao longo do tempo. Eles afirmam que essa persistência de habilidades é a responsável pela variação na renda per capita atualmente, e não as instituições.

Engerman e Sokoloff, por sua vez, destacam o papel das dotações geográficas, especialmente a viabilidade de produzir culturas comerciais lucrativas em grandes plantações, o que gerou desigualdade econômica e política e dificultou o desenvolvimento de instituições inclusivas. Regiões favorecidas para o cultivo de commodities valiosas, como o açúcar, viram a formação de grandes plantações baseadas em trabalho escravizado, resultando em sociedades extremamente desiguais. Nas ex-colônias espanholas, a abundância de recursos minerais intensificou essa concentração de poder econômico e político, levando à criação de instituições que protegiam os privilégios das elites e restringiam a participação da população na economia.

Essas três abordagens pioneiras deram origem a uma vasta literatura, explorando como eventos históricos moldaram instituições e influenciaram o desenvolvimento ao longo do tempo.

Por que a História Importa? Eventos históricos podem influenciar o desenvolvimento econômico de longo prazo se alterarem determinantes fundamentais do crescimento. O principal determinante destacado na literatura são as instituições. Essa ênfase é evidente nos estudos fundamentais de Acemoglu et al. (2001, 2002), Engerman e Sokoloff (1997, 2002) e La Porta et al. (1997, 1998), que apontam as instituições domésticas como o principal mecanismo pelo qual o colonialismo impacta o desenvolvimento atual. Afinal, o que são instituições?

Douglass North define instituições como restrições criadas pelos seres humanos que moldam as interações políticas, econômicas e sociais. Elas podem ser informais, como costumes e códigos de conduta, ou formais, como leis e direitos de propriedade. Em essência, instituições representam as “regras do jogo” de uma sociedade, promovendo ordem e reduzindo incertezas. No entanto, um dos desafios da definição de North é a ausência de uma explicação clara sobre como as regras

institucionais são aplicadas. Avner Greif complementa essa abordagem ao destacar o papel das crenças, normas e organizações para entender por que as pessoas obedecem às regras. Para ele, as instituições são sistemas que combinam regras formais, normas sociais e crenças compartilhadas, garantindo a previsibilidade do comportamento. Dessa forma, uma definição mais completa de instituições deve considerar não apenas as regras em si, mas também os mecanismos que garantem sua aplicação e perpetuação.

Se as instituições representam as “regras do jogo” que orientam o comportamento das pessoas, uma razão pela qual essas regras variam ao redor do mundo é a influência das instituições políticas de cada sociedade. Economistas institucionais argumentam que as instituições políticas desempenham um papel crucial no crescimento econômico, pois são responsáveis por moldar os incentivos e a distribuição dos recursos em uma sociedade. Nesse contexto, surge a questão: quais instituições são mais importantes para o crescimento de longo prazo de uma sociedade: as econômicas ou as políticas?

As instituições políticas determinam quem detém o poder na sociedade, mas elas frequentemente dependem de recursos econômicos e, portanto, das instituições econômicas passadas. Aqueles que controlam as instituições políticas podem usá-las para expandir seu controle sobre os recursos econômicos, ampliando assim seu poder político (Acemoglu e Robinson, 2006). Essa dinâmica é uma das razões pelas quais as instituições podem ser uma fonte de pobreza persistente. Em sociedades onde uma elite restrita domina o sistema político, ela pode estruturar instituições econômicas que favorecem seus próprios interesses, ao mesmo tempo em que limitam oportunidades e aprofundam a pobreza do restante da população.

Acemoglu e Robinson (2012) argumentam que o impacto das instituições políticas no crescimento econômico depende de seu caráter inclusivo ou extrativista. Instituições inclusivas promovem amplos incentivos econômicos e oportunidades, sendo marcadas por uma distribuição equilibrada do poder político. Elas garantem a participação de diferentes grupos sociais, étnicos e culturais nos processos políticos, reduzindo a possibilidade de dominação por uma única elite. Em contraste, instituições extrativistas concentram o poder em uma elite restrita, impondo poucas restrições ao seu domínio. Quando o poder é excessivamente centralizado ou o Estado é frágil, criam-se condições propícias para a exploração econômica e social, dificultando o desenvolvimento sustentável.

Um problema dessa taxonomia é que instituições inclusivas para alguns podem ser extrativistas para outros. Para resolver esse paradoxo, Acemoglu e Robinson (2019) introduziram o conceito

de “corredor estreito para a liberdade”. Eles destacam que o desenvolvimento político é um processo dinâmico, no qual instituições inclusivas exigem um Estado funcional, capaz de conter elites corruptas e grupos de interesse poderosos. No entanto, um Estado excessivamente forte pode desestabilizar essas instituições. Para evitar esse risco, a sociedade civil deve crescer junto com o poder estatal, pois é ela que impõe limites ao Estado. Se o equilíbrio se rompe e o Estado se torna dominante, o país pode sair do “corredor estreito”, resultando em despotismo e instituições extrativistas. Por outro lado, um Estado fraco demais leva ao anarquismo, ao tribalismo e à incapacidade de prover bens públicos de forma eficaz.

Portanto, a grande questão a ser respondida é: como as sociedades fazem a transição para instituições inclusivas? Como elas entram no “corredor estreito” descrito por Acemoglu e Robinson?

3 Tópicos em História Econômica

3.1 O Papel das Instituições e Origens Coloniais no Crescimento Econômico

Por Que as Instituições Importam? A importância das instituições como mecanismo causal continuou a ser explorada em pesquisas subsequentes aos trabalhos seminais mencionados anteriormente. Um exemplo disso é o estudo de Acemoglu et al. (2005) sobre o impacto do comércio atlântico histórico na Europa. Os autores argumentam que, nos países com acesso ao lucrativo comércio triangular atlântico, o poder econômico e político deslocou-se em favor dos interesses comerciais. À medida que a classe mercantil se fortalecia, ela conseguiu modificar as instituições domésticas para proteger seus interesses contra os da realeza, promovendo mudanças institucionais que tiveram efeitos positivos sobre a prosperidade de longo prazo.

O estudo de Gennaioli e Rainer (2007) apresenta evidências da persistência das instituições domésticas no contexto africano, mostrando que sociedades com maior desenvolvimento estatal pré-colonial tendem a prover mais bens públicos atualmente. Esse resultado converge com as conclusões de Nunn (2008), que associa o tráfico de escravos ao enfraquecimento dos estados emergentes e ao aumento da fragmentação étnica, resultando em normas de desconfiança que dificultaram o comércio e a consolidação estatal (Nunn e Wantchekon, 2011). A combinação dessas evidências sugere que o tráfico de escravos debilitou as instituições políticas locais, gerando efeitos negativos de longo prazo no desenvolvimento econômico.

A religião também exerce um papel fundamental na formação de instituições, definindo orga-

nizações, regras e normas de comportamento. No mundo islâmico, por exemplo, leis comerciais e a relação entre governantes e autoridades religiosas tiveram impactos profundos nas instituições. A lei islâmica, por exemplo, dificultava a criação de grandes sociedades por ações (Kuran, 2012). Além disso, a dependência dos governantes das autoridades religiosas para legitimar seu poder levou a restrições que limitaram o desenvolvimento econômico, como a proibição da imprensa (Rubin, 2017). Assim, esses estudos reforçam a ideia de que a história influencia o desenvolvimento por meio da evolução e persistência das instituições.

Embora a literatura sobre instituições apresente evidências convincentes sobre o papel das instituições no desenvolvimento econômico, ela também enfrenta críticas, especialmente pela generalização dos resultados a partir de dados limitados. Um exemplo é a tese de que o “Padrão de Casamento Europeu” teria impulsionado o crescimento econômico. Antes da Revolução Industrial, em algumas regiões da Europa Ocidental, as famílias eram nucleares, as taxas de celibato eram altas e as mulheres se casavam tarde—padrões demográficos usados para explicar a industrialização precoce (Voigtländer e Voth 2006, 2013; De Moor e van Zanden 2010; Foreman-Peck 2011). Embora essa hipótese se encaixe na teoria do crescimento unificado (Galor e Weil, 1999; Galor, 2005), Dennison e Ogilvie (2014) analisaram dados de 39 países europeus e concluíram que os casos mais claros desse padrão estão correlacionados com estagnação econômica, e não com crescimento.

Outro problema recorrente é a falta de evidências que conectem diretamente uma causa institucional no passado a um resultado econômico atual. Muitas pesquisas utilizam regressões baseadas em características institucionais históricas ou fatores que influenciaram sua formação, mas a escassez de dados dificulta a análise de sua evolução ao longo do tempo. Sem essas informações, não é possível determinar se instituições extrativistas persistiram após o fim do regime colonial, apesar da adoção de constituições inspiradas em democracias ricas.

Por exemplo, Dell (2010) analisa os efeitos de longo prazo do sistema de trabalho forçado da mita, imposto pelos espanhóis no Peru e na Bolívia entre 1573 e 1812. A mita obrigava mais de 200 comunidades a fornecer um sétimo de sua população masculina adulta para trabalhar nas minas de prata de Potosí e de mercúrio em Huancavelica. Utilizando uma estratégia de regressão descontínua espacial, Dell mostra que ex-distritos da mita apresentam atualmente um nível médio de consumo domiciliar cerca de 25% inferior ao de distritos não submetidos ao sistema. A autora atribui essa diferença a níveis educacionais mais baixos e infraestrutura viária menos desenvolvida. Apesar de convincente, a análise levanta novas questões, especialmente no período entre 1812 e

a era moderna: após a abolição da mita em 1812, por que os trabalhadores não migraram para regiões mais prósperas? Se apenas os mais pobres e menos qualificados permaneceram, isso poderia indicar um mecanismo alternativo para a baixa renda atual. Portanto, mais dados históricos seriam necessários para confirmar essa explicação.

Outra crítica comum a essa literatura é a de que as instituições nas ex-colônias passaram por transformações após a independência, sugerindo explicações alternativas para o desenvolvimento econômico. Sachs (2012), por exemplo, argumenta que as doenças que influenciaram a escolha das instituições pelos colonizadores europeus também foram letais para as populações nativas, reduzindo sua produtividade e prejudicando o desenvolvimento, independentemente de qualquer efeito indireto por meio das instituições políticas. Além disso, é desafiador separar os efeitos das instituições coloniais dos impactos das habilidades dos colonos europeus (Easterly e Levine, 2016). Assim, o que parece ser um impacto direto das instituições pode, em parte, refletir o efeito do capital humano dos próprios imigrantes europeus sobre o desenvolvimento econômico (Glaeser et al., 2004).

Explicando as Mudanças Institucionais. Uma questão central nas críticas à literatura sobre instituições e desenvolvimento econômico é a explicação para a mudança institucional. Se instituições ruins no passado resultam em pobreza ou autocracia hoje, o que impediu que essas instituições mudassem? E se boas instituições geram riqueza e democracia, o que levou ao seu surgimento? Em outras palavras, por que algumas instituições mudam enquanto outras persistem mesmo quando prejudicam a sociedade? Greif (2006) argumenta que algumas instituições são auto-reforçadoras: à medida que mais indivíduos adotam as normas associadas a elas, sua continuidade se fortalece. Por outro lado, certas instituições podem se enfraquecer e desaparecer quando deixam de ser sustentáveis.

Para Acemoglu e Robinson (2008), as instituições econômicas de equilíbrio emergem da interação entre o poder político *de jure*, alocado pelas instituições políticas, e o poder *de facto*, obtido por grupos através de riqueza, armas ou da capacidade em resolver problemas de ação coletiva. Essa interação pode explicar como mudanças nas instituições políticas não resultam necessariamente em mudanças nas instituições econômicas, desde que o poder *de facto* se ajuste de forma compensatória (ex., na forma de subornos, captura de partidos políticos ou uso de forças paramilitares). Isso é particularmente verdade quando as reformas políticas não alteram radicalmente a estrutura política ou a identidade das elites. Assim, como as elites políticas têm uma vantagem

comparativa na resolução do problema da ação coletiva, elas podem manter seu poder *de facto*, mesmo com a redistribuição do poder *de jure* para a maioria. Portanto, esse fenômeno pode levar a uma persistência institucional semelhante em democracias e regimes não democráticos.

Para ilustrar os elementos que sustentam a persistência e a mudança institucional, considere uma sociedade composta por n grupos de indivíduos, onde o grupo i detém uma participação α_i . As preferências econômicas são definidas em um espaço de políticas, assumindo que todos os indivíduos dentro de um grupo compartilham um ponto ideal b_i e uma função de utilidade em relação à política p , expressa como:

$$u_i(p) = -(p - b_i)^2 + a_i,$$

onde a_i é um termo que depende do arranjo institucional vigente e pode influenciar os retornos individuais. Essas preferências podem evoluir ao longo do tempo em função de alterações na estrutura econômica, mobilidade social e choques políticos que transferem poder entre os grupos. Os retornos totais são calculados como uma soma descontada, com um fator de desconto $\beta \in [0, 1)$. Ao supor que os pontos ideais $\{b_i\}$ são crescentes em i , é possível inferir que os grupos estão ordenados, refletindo uma estrutura ideológica ou econômica. O foco está nas instituições que estabelecem regras de decisão, representadas por s_j , onde cada grupo recebe um “peso” w_i^j na decisão coletiva; instituições com pesos iguais correspondem a democracias, enquanto aquelas que conferem peso positivo apenas ao grupo i caracterizam uma ditadura.

Conforme discutido por Acemoglu et al. (2015), podemos buscar os Equilíbrios de Votação de Markov (MVE), que refletem as escolhas dos eleitores medianos efetivos em relação às instituições, antecipando as decisões dos futuros eleitores medianos. Suponha que iniciamos com a instituição s^j no período t . Tanto as decisões políticas quanto as escolhas institucionais (como reformas) são moldadas pela distribuição de poder resultante dessa instituição. Suponha que essas decisões sejam feitas pelo “eleitor mediano efetivo”, pertencente ao grupo $k = k(s^j)$, de modo que os grupos à esquerda e à direita de k não formem uma maioria ponderada. Como a política não é tratada como uma variável de estado, as escolhas políticas são determinadas pelas preferências do eleitor mediano efetivo, resultando em uma política definida pelo ponto ideal desse grupo, $p = b_k$.

A escolha institucional, no entanto, é mais complexa. As instituições afetam os payoffs individuais de duas formas: primeiro, determinando a distribuição de poder político e, consequentemente, quem definirá a política; segundo, influenciando diretamente os payoffs ao estabelecer regras para

interações econômicas ou possibilitar inovações e investimentos. Essa dinâmica é capturada ao considerar que o termo a_i é uma função da instituição atual, ou seja, $a_i(s^j)$.

Resultado 1 (Estagnação institucional): Suponha que, sob a instituição s^* , o eleitor mediano efetivo pertença ao grupo i^* , com ponto ideal $b_{i^*}^*$ e $a_{i^*}^*(s^*) \geq a_{i^*}^*(s)$ para todas as outras instituições s . Nesse caso, haverá estagnação institucional — o grupo i^* sempre escolherá manter as instituições s^* e, portanto, continuará sendo o eleitor mediano efetivo.

A forma mais simples de persistência institucional ocorre quando uma instituição, uma vez estabelecida, se auto-replica, resultando na ausência de mudanças significativas. Essa permanência é facilitada pela congruência entre poder político e interesses econômicos, combinada com a falta de choques que poderiam alterar essa relação. Exemplos históricos, como a elite de agricultores no Sul dos EUA antes da Guerra Civil, ilustram essa estagnação institucional, já que essa elite controlava uma estrutura institucional que perpetuava seu poder, apoiada pelo sistema escravagista. Entretanto, as instituições não permanecem completamente estáticas

Resultado 2 (Incongruência entre poder político e econômico): Suponha que, sob a instituição s^* , o eleitor mediano efetivo pertença ao grupo i^* , com ponto ideal $b_{i^*}^*$. Se existir uma alternativa s' tal que $a_{i^*}^*(s')$ é significativamente maior que $a_{i^*}^*(s^*)$ e β for pequeno, de modo que os indivíduos são suficientemente míopes, então haverá uma transição para outra instituição.

Essa mudança ocorre porque o grupo no poder prefere um novo arranjo institucional; um exemplo histórico é o dos aristocratas em regimes monárquicos, que perceberam que seus interesses econômicos poderiam ser melhor atendidos ao incluir novos grupos no poder. A emergência das instituições parlamentares na Inglaterra, discutida por North e Weingast (1989), exemplifica essa dinâmica, uma vez que essas mudanças estavam ligadas a compromissos com os mercados financeiros. Teoricamente, para que a mudança institucional ocorra, a nova instituição deve ser suficientemente superior à atual para compensar a perda de poder político. Além disso, a condição de que β seja pequeno é essencial para que o grupo no poder não se preocupe com consequências adversas futuras.

Resultado 3 (Ameaça de revolução): Suponha que, sob a instituição s^* , o eleitor mediano efetivo pertença ao grupo i^* , com ponto ideal $b_{i^*}^*$ e $a_{i^*}^*(s^*) \geq a_{i^*}^*(s)$ para todas as outras s . No entanto, há uma “ameaça de revolução”, o que significa que, se as instituições atuais permanecerem em s^* , há uma probabilidade q de que o poder político se desloque para outro grupo i' , que promoverá uma transição para instituições s' tais que $a_{i^*}^*(s')$ seja suficientemente menor que $a_{i^*}^*(s^*)$. Além disso, existe outra instituição s^d tal que $a_{i^*}^*(s^d)$ é suficientemente maior que $a_{i^*}^*(s')$ e a instituição

s^d é, por sua vez, persistente. Nesse caso, haverá mudança institucional.

Intuitivamente, a ameaça de um resultado institucional pior motiva os atuais detentores de poder a promover mudanças para evitar a revolução. Um exemplo é a teoria da emergência da democracia, onde instituições controladas pela elite podem ser desestabilizadas se não-elites ameaçarem agir coletivamente. Acemoglu e Robinson (2000, 2006b) argumentam que essa ameaça foi crucial para a ascensão da democracia na Europa do século XIX, enquanto as reações a movimentos sociais em outras regiões muitas vezes resultaram em repressão.

A estabilidade institucional também pode ser resultado de cálculos estratégicos. Nesse sentido, a estabilidade estratégica refere-se à situação em que instituições atuais são mantidas por grupos que temem que mudanças institucionais futuras possam diminuir seu poder político ou dificultar sua resistência a mudanças adversas. Um exemplo é o efeito dos “perdedores políticos”, conforme descrito por Acemoglu e Robinson (2006a), onde grupos bloqueiam mudanças por receio de perda de influência.

Resultado 4 (Estabilidade estratégica): Suponha que, sob a instituição s^* , o eleitor mediano efetivo pertença ao grupo i^* , com ponto ideal b_{i^*} , e que $a_{i^*}(s^*)$ seja significativamente menor que $a_{i^*}(s')$ para algum s' , indicando ausência de estagnação institucional. Haverá estabilidade estratégica se, para qualquer s' desse tipo, ocorrer uma transição subsequente para instituições $\{s'_1 = s', s'_2, \dots\}$ com ponto limite s'_∞ tal que $a_{i^*}(s'_\infty)$ seja consideravelmente menor que $a_{i^*}(s^*)$ e o fator de desconto β seja suficientemente alto.

Em outras palavras, existe ausência de estagnação institucional e algumas alternativas potencialmente melhores para o eleitor mediano efetivo. Porém, após a mudança para s' , novas alterações levarão a s'_∞ , que será indesejável para o grupo no poder. Logo, o primeiro passo de mudança institucional desencadeia uma sequência de alterações indesejadas por grupos poderosos. A condição de que β seja suficientemente grande é crucial, pois a estabilidade estratégica decorre de cálculos de longo prazo. Se os agentes não se preocupassem com o futuro (β pequeno), essas considerações não teriam impacto e a estabilidade estratégica não seria alcançada.

Resultado 5 (Instituições mais estáveis): Arranjos institucionais que impedem pequenas mudanças tendem a ser mais duradouros. A estabilidade das instituições é assegurada pela ausência de alternativas estáveis que recebam apoio suficiente de grupos poderosos. Assim, a estabilidade é mais provável quando se torna mais difícil formar coalizões que apoiem instituições alternativas. Quando pequenas reformas são viáveis, elas tendem a ser menos ameaçadoras para grupos poderosos, facilitando a formação de coalizões em favor delas, o que, por sua vez, pode tornar

as instituições existentes menos estáveis. Em contrapartida, um sistema institucional que só pode ser alterado em direções mais radicais levará grupos influentes a bloquear mudanças, tornando as instituições atuais mais estáveis.

Diferentemente da estagnação institucional, a maioria dos casos de persistência institucional não envolve a ausência de mudança, mas sim o modo como as instituições iniciais influenciam essa mudança — o que Acemoglu e Robinson (2012) chamam de ‘mudança dependente da trajetória’. Considere uma sequência estocástica de instituições ao longo do tempo, partindo de uma instituição s ou s' . Se, ao longo do tempo, as duas trajetórias convergirem, não há dependência da trajetória. Caso contrário, existe dependência da trajetória. No entanto, a mudança dependente da trajetória só ocorre quando há uma probabilidade positiva de mudança institucional ao longo do tempo. Isso pode acontecer de duas maneiras:

1. Mudança dependente da trajetória intrínseca: decorre de dinâmicas internas ligadas às condições iniciais.
2. Mudança dependente da trajetória extrínseca: resulta de choques externos, com as instituições respondendo de forma diferente a esses choques dependendo de sua posição inicial.

Em outras palavras, a mudança dependente da trajetória intrínseca leva a uma evolução diferente das instituições a partir de condições iniciais diferentes. Isso ocorre mesmo na ausência de choques, ou seja, estamos analisando uma sequência determinística de equilíbrios institucionais. Já a mudança dependente da trajetória extrínseca é definida para o caso em que não há mudança dependente da trajetória intrínseca e, portanto, sem choques, os caminhos de equilíbrio que partem de condições iniciais diferentes convergirão para os mesmos estados estacionários. No entanto, a chegada de choques durante o processo de convergência institucional pode interromper essa convergência e gerar uma divergência, pois as diferenças institucionais no momento em que o choque ocorre moldarão fundamentalmente a resposta a ele.

3.2 O Papel da Geografia no Crescimento Econômico

A geografia pode, de fato, influenciar o destino das nações? A geografia é frequentemente apresentada como um determinante alternativo do desenvolvimento econômico, muitas vezes contrastando com fatores históricos. Os defensores da hipótese geográfica argumentam que elementos naturais fixos, como clima, ecologia e recursos naturais, desempenham um papel fundamental no

desenvolvimento de longo prazo. Segundo essa perspectiva, as diferenças na distribuição desses fatores geraram incentivos distintos para a produção e o investimento, ampliando as disparidades econômicas entre as regiões.

O estudo seminal de Davis e Weinstein (2002) sobre o desenvolvimento urbano japonês após a Segunda Guerra Mundial oferece evidências convincentes da importância dos fundamentos geográficos. Os autores analisaram a distribuição dos centros urbanos, como Hiroshima e Nagasaki, após a devastação provocada pelas bombas atômicas. Eles mostram que, mesmo após a destruição total, essas cidades recuperaram suas posições relativas na distribuição urbana do Japão em um período de vinte anos. Esse resultado sugere que um grande choque na distribuição da população não foi suficiente para superar as vantagens geográficas intrínsecas desses locais.

Existem diferentes mecanismos pelos quais a geografia influencia a produtividade e as escolhas humanas. O primeiro deles é fator agroclimático: um país com escassez de chuvas, sem condições adequadas de irrigação e solo pobre enfrentará dificuldades maiores para alimentar uma grande população em comparação a um país rico em terras cultiváveis e férteis. Jared Diamond (1997) argumenta que no século XV, a evolução tecnológica da Eurásia em relação a outras regiões do mundo se deve à sua seleção favorável de plantas nativas e animais que podiam ser domesticados, resultando em altos rendimentos. Essa vantagem gerou excedentes alimentares, permitindo a expansão do comércio e a especialização do trabalho, o que desencadeou o desenvolvimento de habilidades, tecnologias e sociedades politicamente organizadas.

Segundo o autor, a domesticação de plantas e animais se espalhou rapidamente pelo Leste e Oeste da Eurásia, mas avançou lentamente para o Sul da África. Isso se deve ao fato de que, ao se mover para o Leste ou Oeste, a duração do dia e o clima não variam drasticamente, ao contrário da movimentação para o Norte ou Sul, onde essas condições mudam significativamente. Assim, em continentes com orientação Norte-Sul, como as Américas e a África, o progresso na domesticação e no avanço tecnológico foi mais lento do que na Eurásia, cuja orientação Leste-Oeste favorecia essa difusão. Além disso, a domesticação precoce de animais na Eurásia propiciou a convivência próxima entre humanos e animais, facilitando o surgimento de doenças, como sarampo, tuberculose, gripe e varíola. Com o tempo, as populações desenvolveram resistência genética a essas doenças, em contraste com regiões sem animais domesticados. Esse fenômeno explica por que as doenças europeias devastaram as populações nativas, e não o inverso.

Embora a análise de Diamond ofereça evidências descritivas sobre o impacto histórico da geografia, ela não quantifica a importância relativa dos efeitos históricos em comparação com os

efeitos diretos e contemporâneos da geografia no desenvolvimento. Nunn e Puga (2012) procuram estimar a magnitude desses dois canais relacionados à rugosidade do terreno, medida pela inclinação média do solo de um país. O estudo identifica dois efeitos principais da rugosidade no desenvolvimento econômico: o primeiro é um efeito direto, já que terrenos acidentados dificultam a construção de infraestrutura e a agricultura, elevando os custos do comércio e afetando negativamente a renda; o segundo é um efeito histórico, que associa o uso de terrenos acidentados à proteção de sociedades da ação de traficantes de escravos durante o comércio de escravos na África, favorecendo o crescimento econômico a longo prazo. Os autores estimam que o efeito histórico é duas vezes maior que o efeito direto contemporâneo, sugerindo que a geografia, por meio da história, exerce um impacto mais significativo do que no presente.

Olsson e Hibbs (2005) também testaram a hipótese de Diamond usando países modernos como unidade de análise. Os autores descobriram que nações com ambientes biológicos e geográficos mais favoráveis fizeram a transição para a agricultura mais cedo e apresentaram um PIB per capita superior em 1997, corroborando as hipóteses de Diamond. Mais recentemente, Pavlik e Young (2019) revisaram os argumentos de Diamond sobre a difusão tecnológica entre regiões e encontraram evidências de que tecnologias se espalharam mais facilmente entre vizinhos ao longo do eixo leste-oeste do que entre aqueles localizados ao longo do eixo norte-sul.

O segundo mecanismo pelo qual a geografia influencia o desempenho econômico é através de doenças. Doenças reduzem a disponibilidade e a qualidade do principal ativo das pessoas mais pobres, que é o tempo de trabalho. Por outro lado, ao reduzir a expectativa de vida das pessoas, uma alta incidência de doenças implica em um período de retorno menor para investimentos em capital humano, fazendo com que as pessoas respondam investindo menos em educação. Além disso, em casos de extrema pobreza, altas taxas de mortalidade causadas por doenças podem induzir taxas de fertilidade mais altas, retardando a transição demográfica. Embora se possa argumentar que a economia direciona o padrão de doenças, muitas doenças são determinadas por fatores geográficos, como temperatura, precipitação e qualidade do solo. Em particular, a malária, que é endêmica nas zonas tropicais e não pode sobreviver em outros lugares, tem sido apontada como um dos principais fatores que explicam por que os países mais pobres do mundo estão localizados em áreas tropicais.

Sachs e Malaney (2002) destacam que os países da África Subsaariana, localizados no “cinturão da malária”, enfrentam sérios desafios econômicos, incluindo taxas de crescimento reduzidas, alta mortalidade infantil e desincentivo a investimentos em capital humano e físico. Além disso, doenças como a infecção transmitida pela mosca tsé-tsé impactam tanto a população hu-

mana quanto o gado, dificultando a domesticação. Alsan (2015) mostra que, nas áreas afetadas pela mosca, o desenvolvimento da agricultura intensiva foi limitado, o uso do arado foi menos comum e a domesticação de animais foi restrita, resultando em um atraso no progresso econômico e na formação de estados centralizados.

Um terceiro mecanismo são os altos custos de transporte. Economias sem saída para o mar, pequenas ilhas, regiões montanhosas ou distantes dos principais mercados enfrentam custos de transporte mais elevados do que economias centrais. Esses altos custos funcionam como tarifas no comércio internacional, reduzindo a extensão do mercado; assim, uma região com elevados custos de transporte se beneficia menos da divisão do trabalho e da difusão tecnológica em comparação àquelas fortemente integradas ao comércio global. Em linha com esse argumento, Redding e Venables (2004) destacam que a geografia é um fator crucial para o tamanho dos mercados atuais, observando que mercados enfrentam dificuldades em se formar em áreas onde o transporte de bens e pessoas é custoso. Utilizando um modelo estrutural de geografia econômica, os autores demonstram que as estimativas baseadas em fluxos comerciais bilaterais fornecem medidas de acesso a mercados e fornecedores em cada país, influenciando os preços dos fatores que as empresas de manufatura podem pagar. Essas medidas explicam uma parte significativa da variação da renda per capita entre os países. Por exemplo, o acesso à costa e políticas de comércio aberto resultam em aumentos previstos de mais de 20% na renda per capita, enquanto reduzir pela metade a distância de um país em relação a todos os seus parceiros comerciais leva a um aumento de cerca de 25%. Além disso, mesmo após controlar diversos outros determinantes da renda per capita – incluindo dotações de recursos naturais, características geográficas e fatores institucionais, sociais e políticos – os efeitos da geografia econômica se mostraram estatisticamente significativos e quantitativamente relevantes.

Por outro lado, evidências indicam que, no curto prazo, a primeira onda de globalização (1850-1913) teve impactos positivos para um pequeno grupo de países, mas efeitos adversos em termos de crescimento econômico para outros. Por exemplo, Pascali (2017) analisa a mudança exógena, variada ao longo do tempo, na abertura comercial causada pela introdução dos navios a vapor. O autor observou, em média, um impacto negativo da abertura comercial em diversas medidas de desenvolvimento econômico. Embora este estudo não aborde especificamente a proteção da indústria nascente, seus resultados são consistentes com a ideia de que, no final do século XIX, países que experimentaram menor abertura ao comércio tiveram maior crescimento.

Esse resultado se alinha aos modelos de indústria nascente que incorporam economias de es-

cala externas, que preveem que países mais abertos desenvolveriam um crescimento mais lento da renda devido à sua contínua especialização em uma agricultura de baixo crescimento, como sugerido por Matsuyama (1992). Em contrapartida, a teoria neoclássica do comércio sugere que o aumento do comércio levaria a uma renda mais alta por meio dos efeitos da especialização. Embora os modelos de indústria nascente não sejam a única abordagem para interpretar esses padrões empíricos—outros fatores incluem o grau de inclusão das instituições políticas e o contexto imperial do século XIX—essa interpretação é significativa, pois sugere que, nesse período crítico, os ganhos da especialização baseada na vantagem comparativa estática podem ter sido superados pelo desenvolvimento da vantagem comparativa dinâmica na manufatura moderna.

Essas evidências levantam uma questão crucial: como superar as restrições geográficas e climáticas? A literatura sugere que investimentos em infraestrutura podem mitigar os efeitos da “maldição da geografia”. Hornbeck e Rotemberg (2021), por exemplo, mostram que o acesso ferroviário teve um impacto significativo na produção manufatureira nos Estados Unidos, estimando que, sem as ferrovias, a produtividade agregada dos EUA teria sido 25% mais baixa em 1890. No contexto da Revolução Industrial na Grã-Bretanha, Bogart et al. (2017) revelam que melhorias no acesso aos mercados impactaram diretamente o crescimento populacional. Sem essas melhorias no transporte, o crescimento urbano entre 1680 e 1831 teria sido 25 pontos percentuais menor, e 97 pontos percentuais a menos entre 1831 e 1851.

Portanto, de acordo com a hipótese da geografia, fatores exógenos relacionados à localização específica de um país desempenham um papel central na determinação de seu potencial de crescimento. Como os fatores geográficos são invariantes ao longo do tempo, essa teoria incorpora um grande grau de determinismo. No entanto, se investimentos em infraestrutura podem superar uma geografia desfavorável, por que alguns estados investem e outros não? Uma questão fundamental é que para superar a “má” geografia são necessárias instituições fortes e eficientes, capazes de alocar recursos adequadamente para a construção e manutenção da infraestrutura.

O Debate Geografia vs Instituições. Na teoria do crescimento econômico, duas forças fundamentais são identificadas como determinantes do desempenho econômico a longo prazo: instituições e geografia. Embora o desenvolvimento econômico seja um fenômeno complexo e nenhum fator isolado possa explicar completamente as disparidades econômicas globais, compreender a importância relativa desses dois elementos é vital para as políticas públicas. Enquanto a geografia se refere a condições que as sociedades não podem modificar, as instituições, criadas pelo ser hu-

mano, têm o potencial de serem alteradas por ações coletivas. Assim, não é surpreendente que uma literatura empírica tenha surgido para investigar se a geografia ou as instituições desempenham um papel mais crucial no desenvolvimento econômico.

O debate surge em parte porque a geografia influencia tanto as ações humanas no passado quanto no presente. Ou seja, além de afetar diretamente a renda, a geografia também molda a história, que, por sua vez, impacta a renda atual. Existem diversos exemplos de como pequenas diferenças geográficas, amplificadas por eventos históricos, resultam em grandes efeitos no desenvolvimento econômico de longo prazo. Um exemplo são as diferenças no solo e clima, que tornaram a agricultura de plantation e sua dependência da escravidão mais ou menos lucrativas em diferentes regiões das Américas (Engerman e Sokoloff, 1997, 2002).

A tensão entre as visões de que a “geografia importa” e a “história importa” também é evidente no estudo de Acemoglu et al. (2001), que utilizaram a mortalidade dos colonos como um instrumento para demonstrar o efeito causal das instituições na renda. Jeffrey Sachs (2003), um defensor da perspectiva geográfica, argumenta que as estimativas do método de variável instrumental utilizadas por Acemoglu et al. são enviesadas, pois a geografia exerce um efeito direto na renda atual. Ele afirma que regiões com alta mortalidade de colonos durante o período colonial continuam a enfrentar ambientes com doenças letais, resultando em níveis de renda mais baixos atualmente. Assim, como a mortalidade inicial dos colonos se relaciona com a renda por meio de um canal distinto das instituições, a restrição de exclusão não é satisfeita.

Acemoglu et al. reconhecem essa preocupação e argumentam que as principais doenças que afetaram os europeus, como malária e febre amarela, tinham um impacto limitado nas populações indígenas, que desenvolveram imunidade. Devido a essa imunidade, é improvável que as taxas de mortalidade dos colonos europeus estejam fortemente correlacionadas com as cargas de doenças enfrentadas pelas populações nativas, tanto no passado quanto atualmente. Além disso, os autores demonstram que seus resultados permanecem robustos mesmo após incluir diversos controles que medem características geográficas e o ambiente local de doenças.

Acemoglu e seus coautores apresentam um argumento adicional em favor da hipótese institucional, afirmando que a colonização europeia resultou em várias “reversões de fortuna”, onde regiões ricas se tornaram pobres e vice-versa. Se fatores como geografia, clima ou ecologia fossem responsáveis pela riqueza ou pobreza, essas regiões deveriam ter permanecido em seus estados originais. Os autores argumentam que a principal explicação para essas mudanças reside na qualidade das instituições, as quais tiveram um papel importante entre o final do século XVIII e o início

do século XIX durante o processo de industrialização. No entanto, é importante reforçar que a escolha dos europeus por diferentes modelos institucionais nas colônias está ligada às condições geográficas e recursos disponíveis. Assim, embora as "reversões de fortuna" tenham sido impulsionadas por mudanças institucionais significativas, a geografia teve um papel fundamental na configuração da qualidade dessas instituições.

Assim, a geografia pode, no máximo, explicar a localização da atividade econômica, mas oferece pouco entendimento sobre o timing do crescimento econômico. Isso sugere que os efeitos mais significativos da geografia atuam através das instituições econômicas e políticas. Em certos casos, condições geográficas favoráveis podem até se tornar obstáculos ao desenvolvimento econômico. A literatura sobre a "maldição dos recursos" (Sachs e Warner, 2001; Mehlum, Moene e Torvik, 2006; Robinson, Torvik e Verdier, 2006; Ross, 2015; Desierto, 2018) investiga por que países ricos em recursos naturais frequentemente continuam pobres, uma questão particularmente relevante na África Subsaariana. A dificuldade enfrentada por alguns países em converter suas riquezas naturais em prosperidade está intimamente relacionada às instituições políticas. Quando essas instituições são extrativas, os recursos naturais tendem a fortalecer o controle das elites, que os utilizam para beneficiar seus aliados e enriquecê-las, perpetuando assim instituições ainda mais extrativas. Assim, enquanto os recursos naturais podem ser uma bênção, também têm o potencial de se tornar uma maldição.

Portanto, a geografia apresenta a vantagem de ser exógena, ou seja, não é diretamente influenciada por outras variáveis, como cultura e instituições. Contudo, sua imutabilidade é uma limitação significativa, desafiando teorias que a conectam ao crescimento econômico a longo prazo. Muitas disparidades de renda observadas hoje passaram por transformações ao longo do tempo, dificultando a explicação de reviravoltas econômicas apenas com base na geografia. Em geral, as evidências indicam que os maiores efeitos da geografia no desenvolvimento econômico contemporâneo se manifestam por meio de sua influência em eventos passados, em vez de por seus efeitos diretos nos resultados econômicos atuais.

3.3 Herança Colonial e Persistência Econômica nas Américas

A Era Colonial foi marcada pelo enriquecimento das potências coloniais e pela miséria dos povos nativos e africanos escravizados. Regiões inicialmente menos desenvolvidas, com tecnologias mais simples, experimentaram um crescimento econômico acelerado após a colonização. Em

contraste, áreas densamente povoadas vivenciaram um desenvolvimento mais lento durante a era moderna. Como o legado do colonialismo impactou o desenvolvimento a longo prazo? Por que algumas colônias se tornaram nações prósperas, enquanto outras permanecem imersas na pobreza? Este ensaio busca abordar essas questões, com ênfase no papel das instituições implantadas pelos europeus durante o período colonial nas Américas.

Persistência Econômica nas Américas. Os estudos que analisam os impactos do domínio colonial nas Américas tendem a se concentrar em testar a hipótese de que as dotações iniciais influenciaram o grau de desigualdade econômica e política, ambas prejudiciais ao desenvolvimento econômico de longo prazo (Engerman e Sokoloff, 1997). Em um estudo posterior, Engerman e Sokoloff (2005) fornecem evidências adicionais para sua hipótese ao documentar uma relação positiva entre desigualdade econômica e desigualdade política, esta última medida pela expansão nos direitos ao voto. Por sua vez, Sokoloff e Zolt (2007) fornecem evidências de que a desigualdade na América Latina está relacionada a uma tributação mais baixa sobre riqueza e renda, bem como a menores investimentos em bens públicos, como educação.

Embora a hipótese de Engerman e Sokoloff se baseie amplamente em narrativas históricas e estatísticas descritivas, diversos estudos realizaram testes mais formais. Bruhn e Gallego (2012) analisam 345 regiões de 17 países da América do Norte e do Sul e identificam uma forte correlação negativa entre o desenvolvimento de longo prazo e a especialização colonial em atividades consideradas “ruins” — aquelas que, segundo Engerman e Sokoloff (1997), apresentavam economias de escala e dependiam fortemente do trabalho explorado, como a produção de açúcar, café, arroz, algodão e mineração. Por outro lado, os autores mostram que atividades como agricultura de subsistência, criação de gado e manufatura não estavam negativamente associadas ao desenvolvimento de longo prazo, a menos que envolvessem grandes populações nativas exploradas no processo produtivo.

Naritomi et al. (2012) fornecem evidências semelhantes ao analisar a influência do ciclo da cana-de-açúcar (1530–1760) e do ciclo do ouro (durante os anos 1700) nos municípios brasileiros. Esses períodos representaram as principais expansões econômicas da história colonial do Brasil, relacionados à ocupação inicial de certas regiões e ao desenvolvimento de atividades extrativas promovidas por Portugal. Esses ciclos podem ser delimitados cronologicamente e geograficamente, permitindo que alguns municípios rastreiem suas origens a episódios específicos de exploração extrativa. Com isso, surge a questão: os municípios afetados por esses eventos históricos apresentam

características sistematicamente diferentes atualmente?

A análise foca em três características municipais: a distribuição do poder econômico, ligada à distribuição de terras; o acesso à justiça, medido pela disponibilidade de tribunais locais; e a qualidade das práticas de governo, quantificada por um índice de eficiência administrativa. Os efeitos estimados para cada episódio extrativo são consistentes com as características socioeconômicas e políticas frequentemente discutidas na literatura histórica. Municípios com origens ligadas à economia oligárquica do ciclo da cana-de-açúcar apresentam maior concentração fundiária. Por outro lado, os municípios associados ao ciclo do ouro mostram atualmente práticas de governança mais fracas e menor acesso à justiça. Em ambos os casos, os efeitos negativos são significativamente mais pronunciados quanto mais próximos os municípios estão de Portugal, destacando a influência prejudicial da interferência da metrópole, especialmente quando associada a atividades extrativas.

O mecanismo central da hipótese de Engerman e Sokoloff é que a desigualdade econômica e política resultante das estruturas coloniais teria prejudicado o desenvolvimento. No entanto, essa visão é contestada por diversos estudos. Dell (2010) analisa os efeitos de longo prazo do sistema de trabalho forçado da mita, imposto pelos espanhóis no Peru e na Bolívia entre 1573 e 1812. A mita obrigava mais de 200 comunidades a fornecer um sétimo de sua população masculina adulta para trabalhar nas minas de prata de Potosí e de mercúrio em Huancavelica. Utilizando uma estratégia de regressão descontínua espacial, que envolve a comparação dos resultados em ambos os lados de uma seção da fronteira geográfica da área afetada, Dell mostra que os ex-distritos da mita apresentam atualmente um nível médio de consumo domiciliar aproximadamente 25% inferior ao dos distritos que não foram submetidos a esse sistema. A autora atribui essa diferença a níveis educacionais mais baixos e infraestrutura viária menos desenvolvida.

Dell argumenta que o subdesenvolvimento das regiões da mita decorreu da ausência de grandes haciendas (grandes propriedades rurais baseadas no trabalho forçado) devido a uma política colonial que restringia a formação. Em outras áreas, as haciendas pressionavam a Coroa por investimentos em bens públicos, como estradas e escolas, e frequentemente ofereciam esses serviços diretamente. Assim, em contraste com a hipótese de Engerman e Sokoloff, Dell encontra melhores resultados de longo prazo em locais com grandes haciendas e maior desigualdade. No entanto, sua análise levanta novas questões, especialmente no período entre 1812 e a era moderna: após a abolição da mita em 1812, por que os trabalhadores não migraram para regiões mais prósperas? Se apenas os mais pobres e menos qualificados permaneceram, isso poderia indicar um mecanismo

alternativo para a baixa renda atual. Mais dados históricos seriam necessários para confirmar essa explicação.

Acemoglu et al. (2008) também questionam a hipótese de Engerman e Sokoloff. Ao analisar municípios de Cundinamarca, na Colômbia, os autores mostram que a desigualdade fundiária no final do século XIX estava positivamente associada à matrícula no ensino secundário no final do século XX. Além disso, questionam a ideia de que desigualdade econômica e política caminham juntas. A partir da identidade de prefeitos em 4.763 nomeações entre 1875 e 1895, os autores demonstram que a desigualdade econômica não estava correlacionada com a desigualdade política. Pelo contrário, argumentam que maior desigualdade fundiária poderia ter sido benéfica para o desenvolvimento de longo prazo, pois grandes proprietários de terra atuavam como contrapesos ao poder da elite política.

Nunn (2007) também examina o papel da desigualdade ao comparar estados e condados dos EUA, bem como países das Américas. Embora encontre uma relação negativa entre o uso passado da escravidão e a renda atual, consistente com Engerman e Sokoloff, ele não identifica a desigualdade como canal intermediário desse efeito. Embora a escravidão esteja positivamente correlacionada com a desigualdade histórica e atual, controlar pela desigualdade fundiária não reduz o impacto negativo da escravidão sobre a renda. Além disso, não há evidências de uma relação robusta entre desigualdade e renda, nem no passado nem no presente.

É importante destacar que, embora as evidências sugiram que as instituições adotadas nas regiões colonizadas tenham tido um efeito duradouro no desenvolvimento econômico das ex-colônias, essa interpretação enfrenta críticas relevantes. Sachs (2012), por exemplo, argumenta que as doenças que influenciaram a escolha das instituições pelos colonizadores europeus também foram letais para as populações nativas, reduzindo sua produtividade e prejudicando o desenvolvimento, independentemente de qualquer efeito indireto por meio das instituições políticas. Além disso, é desafiador separar os efeitos das instituições coloniais dos impactos das habilidades dos colonos europeus (Easterly e Levine, 2016). Assim, o que parece ser um impacto direto das instituições pode, em parte, refletir o efeito do capital humano dos próprios imigrantes europeus sobre o desenvolvimento econômico (Glaeser et al., 2004).

Explicando as Mudanças Institucionais. Embora muitas abordagens sobre o desenvolvimento nas Américas assumam que as instituições persistem por longos períodos, instituições específicas podem variar consideravelmente ao longo do tempo. Países menos desenvolvidos, especialmente

na América Latina e África, frequentemente alteraram suas instituições políticas, alternando entre democracia e ditadura e implementando reformas constitucionais. Esse padrão também se aplica às instituições econômicas. Embora práticas como a escravidão e a mita tenham desaparecido, os sistemas econômicos frequentemente exibem continuidade.

Isso levanta a questão de como a continuidade dos equilíbrios institucionais pode coexistir com mudanças em instituições específicas. Acemoglu e Robinson (2008) propõem uma explicação simples: as instituições econômicas de equilíbrio emergem da interação entre o poder político *de jure*, alocado pelas instituições políticas, e o poder *de facto*, obtido por grupos através de riqueza, armas ou da capacidade em resolver problemas de ação coletiva.

Essa interação pode explicar como mudanças nas instituições políticas não resultam necessariamente em mudanças nas instituições econômicas, desde que o poder *de facto* se ajuste de forma compensatória (ex., na forma de subornos, captura de partidos políticos ou uso de forças paramilitares). Isso é particularmente verdade quando as reformas políticas não alteram radicalmente a estrutura política ou a identidade das elites. Assim, como as elites políticas têm uma vantagem comparativa na resolução do problema da ação coletiva, elas podem manter seu poder *de facto*, mesmo com a redistribuição do poder *de jure* para a maioria. Portanto, esse fenômeno pode levar a uma persistência institucional semelhante em democracias e regimes não democráticos.

3.4 Origens Étnicas, Nacionais e Históricas no Continente Africano

A Conferência de Berlim (1884-1885) marcou o início do colonialismo em larga escala no continente africano, com potências europeias dividindo arbitrariamente o território e administrando suas colônias. Embora o período colonial tenha sido relativamente breve em comparação a outras regiões, estudos sobre os impactos desse domínio revelam efeitos duradouros que persistem até hoje. Uma questão relevante é se o desenvolvimento econômico da África teria sido diferente na ausência do colonialismo e se o continente seria mais rico atualmente. Algumas evidências desafiam as críticas ao colonialismo, destacando melhorias em indicadores econômicos durante esse período. Por exemplo, utilizando uma cesta de bens padronizada para calcular os salários reais, Frankema e Van Waijenburg (2012) mostram que os padrões de vida dos africanos ocidentais eram surpreendentemente altos em comparação com os trabalhadores não qualificados de países asiáticos durante grande parte do período colonial.

Isso leva à reflexão: o domínio colonial foi menos predatório do que geralmente se argu-

menta? Poderia ser uma evidência de que o colonialismo beneficiou o desenvolvimento? Alguns pesquisadores afirmam que a resposta é provavelmente não (Heldring e Robinson, 2012). Este ensaio investiga brevemente as razões para o baixo desenvolvimento econômico da África, organizando-se em duas partes: a primeira analisa os impactos de eventos históricos no desenvolvimento econômico de longo prazo, enquanto a segunda investiga os efeitos do comércio de escravos no continente e seus possíveis mecanismos.

As Origens do Desenvolvimento Africano. A África nem sempre foi o continente mais pobre, mas medir suas oscilações econômicas ao longo do tempo não tem sido uma tarefa fácil. A falta de registros escritos, especialmente no período pré-colonial, complica qualquer análise de longo prazo e tem levado historiadores econômicos a buscar abordagens criativas para medir a trajetória dos padrões de vida ao longo do tempo. Nas últimas décadas, vários pesquisadores têm investigado como fatores históricos profundos ainda moldam o desenvolvimento econômico da África. A capacidade de mapear fatores ambientais, políticos e culturais na história e relacioná-los a resultados contemporâneos permite que economistas revelem correlações e até causalidade, algo que não era viável antes devido à falta de recursos computacionais e técnicas econométricas rigorosas.

Alguns pesquisadores argumentam que a diversidade das condições ambientais desempenhou um papel crucial no desenvolvimento da região. Por exemplo, a mosca tsé-tsé, encontrada em áreas tropicais da África, mas inexistente em outros continentes, transmite um parasita prejudicial aos humanos e letal para o gado. Alsan (2014) investigou empiricamente os efeitos de longo prazo dessa mosca sobre os resultados econômicos. Ela mostrou que a mosca predominava em regiões adequadas para a agricultura, diminuindo o excedente agrícola, pois os agricultores eram menos propensos a usar animais domesticados e arados, resultando em densidades populacionais mais baixas e menor centralização política.

Condições climáticas e ambientais também influenciaram a adoção de práticas agrícolas. Michalopoulos et al. (2016) descobriram que pessoas de etnias cuja subsistência dependia mais da agricultura no período pré-colonial tendem a ser mais educadas e ricas hoje, o que pode ser atribuído a diferenças em atitudes, crenças e tratamentos por outros grupos. Outro mecanismo para a persistência dessas práticas agrícolas antigas é a complexidade dos regimes políticos pré-coloniais que delas se originaram. Gennaioli e Rainer (2007), utilizando dados etnográficos para medir o desenvolvimento estatal nas sociedades africanas pré-coloniais, mostram que a força das instituições políticas desse período foi crucial para a capacidade dos governos coloniais e pós-

coloniais em fornecer bens públicos.

Em um estudo inovador, Michalopoulos e Papaioannou (2013) mostram como a distribuição espacial das etnias pré-coloniais afeta o desempenho econômico contemporâneo. Regiões que tinham níveis mais altos de centralização política no período pré-colonial exibem hoje maior atividade econômica, evidenciada por uma maior densidade de luz noturna capturado por imagens de satélite. Os autores afirmam que essa associação é independente de características geográficas e de outras variáveis culturais e econômicas associadas às etnias.

Outro estudo importante sobre os impactos de longo prazo do domínio colonial na África é o de Huillery (2009). A autora combina dados de documentos históricos provenientes dos arquivos de Paris e Dacar com pesquisas domiciliares da década de 1990. Sua análise dos distritos da África Ocidental Francesa revela uma correlação positiva entre os investimentos coloniais iniciais em educação, saúde e infraestrutura e os níveis atuais de escolaridade, indicadores de saúde e acesso à eletricidade, água e combustíveis.

Um aspecto interessante da pesquisa é a evidência de persistência específica para cada tipo de bem público. Huillery mostra que investimentos mais elevados em educação durante o período colonial estão associados a níveis mais altos de escolaridade no período pós-colonial, enquanto esses investimentos não impactam os indicadores de saúde ou infraestrutura. Da mesma forma, os investimentos em infraestrutura realizados durante a colonização estão relacionados a um maior acesso à infraestrutura atualmente, mas não afetam os outros indicadores. Por outro lado, os investimentos em saúde resultam em melhores indicadores de saúde, mas não influenciam as demais áreas analisadas.

Embora os mecanismos por trás dessa acentuada persistência ainda não sejam totalmente compreendidos, Huillery fornece evidências de que os investimentos iniciais geraram mais investimentos do mesmo tipo ao longo do tempo. Uma explicação alternativa sugere que fatores persistentes que afetam especificamente determinados bens públicos possam estar impulsionando esses resultados. No entanto, a autora adota várias estratégias em sua análise para descartar essa hipótese, incluindo o pareamento de distritos com base na proximidade geográfica.

Uma consequência adversa frequentemente associada ao domínio colonial europeu na África é a criação de fronteiras nacionais que desconsideraram os reinos, estados e grupos étnicos preexistentes. Embora a ideia de que um dos legados do colonialismo seja a natureza artificial dessas fronteiras tenha sido discutida há muito tempo, essa afirmação foi formalmente testada apenas recentemente. Michalopoulos e Papaioannou (2011) combinaram dados sobre a localização pré-

colonial de 834 grupos étnicos, baseada em Murdock (1959), com as fronteiras atuais dos Estados nacionais, analisando as diferenças entre grupos étnicos separados por fronteiras e aqueles que não foram afetados.

Para cada etnia, os autores criaram duas variáveis indicadoras. A primeira assume o valor “1” se a etnia foi dividida por uma fronteira, com mais de 10% de sua área situada de ambos os lados da fronteira. A segunda atribui o valor “1” se a etnia também foi dividida, mas menos de 10% de sua área está de um lado da fronteira. Essas medidas visam abordar possíveis erros de mensuração decorrentes da imprecisão na localização dos grupos étnicos conforme mapeados por Murdock (1959). Mesmo que as fronteiras não dividam efetivamente os grupos, esses erros de mensuração podem criar divisões, sendo mais provável que isso ocorra na segunda medida (menos de 10%).

Os autores analisaram dois conjuntos de resultados em nível étnico: o desenvolvimento econômico, medido pela densidade de luz noturna, e a incidência de conflitos civis entre 1970 e 2005. Os resultados mostraram que grupos étnicos divididos por fronteiras estão associados a um menor desenvolvimento econômico, refletido em uma menor densidade de luz, e a uma maior incidência de guerras civis. Ambas as medidas de partição apresentaram significância estatística; embora a magnitude da medida com menos de 10% geralmente seja menor, isso é consistente com uma maior probabilidade de erro de mensuração. Assim, os resultados confirmam a crença de que o domínio colonial, devido à maneira artificial como o continente foi dividido entre as potências europeias, teve efeitos prejudiciais.

Uma das possíveis consequências adversas do domínio colonial foi o agravamento das hostilidades entre grupos étnicos. Em Ruanda, as políticas coloniais aprofundaram deliberadamente as diferenças raciais entre Hutus e Tutsis. O censo de 1933–1934 institucionalizou essa distinção ao introduzir cartões de identidade que registravam a etnia de cada indivíduo. Além disso, foi estabelecido um sistema educacional segregado, com vias separadas para Hutus e Tutsis. Antes da colonização europeia, a divisão entre Hutus e Tutsis era mais social do que racial, permitindo mobilidade entre os grupos. Diante da dificuldade de classificar rigidamente a população, as autoridades belgas adotaram a “Regra das Dez Vacas”: aqueles que possuíam mais de dez vacas eram classificados como “Tutsi”, enquanto os demais eram considerados “Hutu”.

Diversos autores, com destaque para Mamdani (2001), defendem que as tensões étnicas entre Hutus e Tutsis — que culminaram no genocídio de 1994 — têm suas origens nas políticas coloniais belgas. No entanto, essa interpretação não é unânime. Vansina (2004), apoiado principalmente em relatos orais e fontes históricas, argumenta que as divisões entre os grupos já existiam antes da

chegada dos europeus. Já a análise empírica de Anagizawa-Drott (2014) mostrou que o acesso ao rádio, e à propaganda de ódio anti-Tutsi transmitida por meio dele, fez com que a sociedade passasse de um equilíbrio normativo caracterizado pela paz para outro, marcado por assassinatos em massa. Os programas de rádio não apenas incentivaram diretamente os indivíduos a matar Tutsis, mas também tiveram o efeito indireto de tornar os assassinatos mais frequentes, alterando as normas em relação aos homicídios e ampliando ainda mais sua ocorrência.

O Comércio de Escravos: Causas, Consequências e Controvérsias. Os historiadores estudam há muito a devastação do tráfico de escravos na África. Entre 1400 e 1900, aproximadamente 12 milhões de africanos foram enviados pelo tráfico transatlântico e outros 6 milhões por rotas como o transaariano, o Mar Vermelho e o Oceano Índico. Estimativas indicam que a população africana poderia ser metade do que seria sem essas práticas (Manning 1990).

A demanda e oferta explicam a predominância do tráfico transatlântico no século XVI, com a mão de obra africana sendo altamente produtiva no Novo Mundo. A resistência dos africanos a doenças tropicais e sua proximidade com as Américas os tornaram mais atraentes que trabalhadores europeus ou asiáticos (Bertocchi e Dimico 2014; Angeles 2013). Angeles (2013) destaca que os baixos custos de captura, em grande parte realizados por africanos, e a fragmentação étnica contribuíram para a lucratividade do tráfico. Condições climáticas também impactaram a oferta: Fenske e Kala (2015), por exemplo, descobriram que exportações aumentavam em anos mais frios, pois temperaturas mais baixas diminuía a mortalidade e os custos de transporte. A seca do século XIX, apontada por Levi Boxell (2017), incentivou o aumento das exportações e elevou a probabilidade de conflitos nas regiões exportadoras.

Enquanto historiadores analisam as causas da escravidão, economistas se concentram em suas consequências. Nunn (2008) foi pioneiro ao estabelecer uma relação causal entre o tamanho da população escravizada e a pobreza atual de muitos países africanos, argumentando que as regiões mais prósperas eram as mais envolvidas no tráfico. Seu diferencial foi o uso de uma variável instrumental: a distância de cada país africano aos mercados escravistas das Américas, assumindo que essa distância não influencia os resultados econômicos atuais, exceto pelo impacto do tráfico de escravizados. As estimativas IV de Nunn corroboram suas estimativas OLS, evidenciando que o tráfico teve efeitos negativos duradouros nas economias africanas.

A análise de Nunn gerou interesse em identificar os mecanismos pelos quais a escravidão produziu efeitos persistentes. Por exemplo, Nunn e Puga (2012) argumentam que a geografia

pode ser um mecanismo pelo qual o comércio de escravos afetou o desenvolvimento africano. Eles mostram que a rugosidade do terreno, medida pela inclinação média do solo de um país, teve um efeito positivo sobre a renda, uma vez que dificultava os impactos do comércio de escravos; terrenos acidentados eram mais difíceis de serem atravessados pelos invasores e mais adequados para esconderijos de pessoas perseguidas.

Nunn e Wantchekon (2011) exploram a desconfiança étnica como mecanismo central para os efeitos de longo prazo do comércio de escravos. Eles mostram que descendentes de grupos étnicos fortemente expostos ao tráfico apresentam níveis mais baixos de confiança, tanto em relação a parentes e vizinhos quanto ao governo local, afetando negativamente o desenvolvimento econômico atual. O estudo investiga dois canais potenciais desse impacto: um que altera as normas culturais, tornando grupos menos confiantes, e outro que deteriora instituições legais e políticas, aumentando a desconfiança interpessoal. Para distinguir entre esses canais, os autores realizam testes que avaliam a confiança na governança local levando em consideração a qualidade institucional baseada na percepção da eficiência do governo, corrupção e provisão de bens públicos.

Um dos testes usa o número médio de escravizados retirados da localidade do indivíduo, em contraste com a média do grupo étnico, permitindo diferenciar entre os impactos cultural (normas internalizadas) e institucional (instituições externas) do tráfico. Isso ocorre porque, ao se mudar, o indivíduo leva consigo suas normas internas, mas deixa para trás o ambiente institucional. Os resultados revelam que ambos os canais afetaram a confiança, com o canal cultural apresentando um impacto mais significativo.

Existem diversas razões pelas quais o fracionamento étnico pode estar negativamente associado ao crescimento econômico. Em estados fragmentados, alguns grupos governam sobre outros, o que pode ser desastroso, especialmente se houver um histórico de animosidade entre eles. Nesse contexto, o fracionamento pode gerar conflitos civis, reduzir a provisão de bens públicos e diminuir a confiança. Nesse sentido, Obikili (2016) mostra que vilarejos e cidades de grupos étnicos com maiores exportações de escravizados eram mais fragmentados politicamente na era pré-colonial, uma característica que ainda se reflete em resultados políticos atuais.

Outro impacto cultural do comércio de escravizados é a poligamia, resultado da notável disparidade de gênero criada pela captura de um número significativamente maior de homens. Dalton e Leung (2014) e Fenske (2015) encontraram uma correlação positiva entre as exportações de escravizados e a prevalência da poligamia, sugerindo que isso contribuiu para o crescimento econômico lento, afetando poupanças, investimentos em capital humano das meninas e a oferta de trabalho de

homens solteiros.

Finalmente, Lowes et al. (2017) investigam a persistência de normas culturais em um estudo na República Democrática do Congo, onde os descendentes do Reino Kuba, estabelecido no século XVII, mostraram-se menos propensos a seguir regras e mais inclinados a cometer furtos. Os autores argumentam que, sob o domínio do Kuba, as normas de conformidade eram menos necessárias, pois as punições eram aplicadas pelas instituições do reino. Essa evidência reforça a ideia de que normas culturais podem persistir muito tempo após mudanças no ambiente institucional e político.

3.5 A Revolução Industrial Britânica: Comércio, Invenção Induzida e Progresso Tecnológico

A Revolução Industrial Britânica desencadeou o crescimento econômico moderno. Mas por que esse processo começou na Inglaterra entre os séculos XVIII e XIX? Este ensaio revisa a literatura para responder a essa questão, destacando pré-condições fundamentais e os argumentos de Allen e Mokyr. Allen atribui as inovações industriais aos altos salários e baixos custos de energia, enquanto Mokyr enfatiza o papel do desenvolvimento cultural e intelectual europeu, aliado ao capital humano elevado.

Pré-Condições para a Industrialização. Por volta de 1700, a Grã-Bretanha já possuía uma economia de mercado avançada, impulsionada por melhorias em infraestrutura. A ascensão do consumidor, conhecida como Revolução do Consumidor, marcou a transição da produção doméstica para um modelo orientado pelo mercado, impulsionando o crescimento econômico da época. De Vries (1993, 2008) argumenta que a expansão dos mercados estimulou o consumo, resultando em uma “revolução industriosa”, na qual as famílias começaram a trabalhar mais para adquirir bens, reduzindo a produção doméstica e aumentando a oferta de trabalho (Koyama, 2012). Essa transformação foi um precursor essencial para a Revolução Industrial, ao elevar o PIB per capita e criar bases para o sistema fabril (Clark, 1994).

No século anterior à Revolução Industrial, a agricultura inglesa já era comercializada, com grandes fazendas e trabalhadores assalariados. O processo de cercamentos (enclosures), acelerado no século XVIII, aumentou a produtividade e possibilitou a mecanização agrícola (Heldring, Robinson e Vollmer, 2021). O excedente alimentar gerado favoreceu a urbanização, embora não

tenha sido suficiente para garantir um crescimento econômico sustentado após 1800. O ambiente político britânico era outro fator que favorecia o crescimento econômico, pois o país possuía governos representativos e limitados. Contudo, as instituições por si só não explicam a Revolução Industrial. Historiadores argumentam que o que realmente tornou essa transformação uma “revolução” foi a onda de inovações que se intensificou na Grã-Bretanha a partir da segunda metade do século XVIII. Nesse contexto, a limitação do poder executivo na Grã-Bretanha assegurou a proteção da propriedade privada, criando um ambiente propício à inovação.

A economia atlântica também alterou a estrutura econômica britânica. O comércio impulsionou os portos da costa oeste e estimulou a manufatura em regiões como Yorkshire e Lancashire. A introdução de novos produtos, como chá, tabaco e açúcar, expandiu a Revolução do Consumidor. Inicialmente bens de luxo, esses produtos logo se tornaram acessíveis a uma parcela maior da população. Historiadores como Williams (1944) e Inikori (2002) destacam o papel do comércio de escravos na economia britânica, associando a colonização europeia à Revolução Industrial. Contudo, setores como carvão, ferro e têxteis tiveram um impacto maior do que o açúcar (Eltis e Engerman, 2000). Esses centros industriais exigiam uma grande quantidade de trabalhadores altamente qualificados, cuja concentração em cidades pode ter sido um dos fatores que diferenciaram a Grã-Bretanha. Do mesmo modo, embora Inikori (2002) sugira que o comércio exterior aumentou a produtividade, não há evidências definitivas dessa relação.

Já Beckert (2014) destaca o algodão como essencial para a economia industrial, atribuindo seu sucesso ao colonialismo e à intervenção estatal. No entanto, Mokyr (2009) sustenta que a Revolução Industrial não dependia exclusivamente da indústria têxtil. Alguns economistas sugerem que o tamanho do mercado britânico foi crucial para incentivar a inovação e impulsionar o crescimento. Desmet e Parente (2012) associam o aumento dos mercados ao crescimento das empresas, fatores interligados que teriam sido determinantes para a Revolução Industrial. Entre os séculos XVII e XVIII, o crescimento dos mercados urbanos favoreceu a especialização regional e o desenvolvimento manufatureiro em áreas como o Noroeste e os Midlands, onde a terra era menos propícia à agricultura (Kelly, Mokyr e Ó Gráda, 2020). No entanto, esses fatores isolados não explicam completamente a Revolução Industrial, já que outras regiões também possuíam mercados integrados sem gerar crescimento sustentado (Bateman, 2011, 2012).

A capacidade estatal britânica também foi um diferencial importante, pois permitiu a sustentação da dívida pública sem a necessidade de impostos excessivos, enquanto a República Holandesa enfrentava desafios fiscais e limitações estruturais ao crescimento econômico. No entanto,

isso explica apenas a estagnação holandesa, não a inovação britânica. Outro diferencial essencial da Grã-Bretanha foi a abundância de trabalhadores qualificados, fundamental para a Revolução Industrial (Kelly, Mokyr e Ó Gráda, 2014, 2020). O sistema de aprendizado, mais flexível e voltado à promoção de alto capital humano, contrastava com as corporações de ofício europeias (Humphries, 2010; Wallis, 2008). Esse sistema ajudou a reduzir a disparidade salarial entre trabalhadores qualificados e não qualificados, estimulando o desenvolvimento técnico (van Zanden, 2009).

Em resumo, a Grã-Bretanha possuía muitas das precondições para a industrialização: instituições políticas que protegiam a propriedade, um ambiente receptivo à inovação, acesso ao comércio atlântico e às economias escravistas, sem cair na “maldição dos recursos”. O mercado de trabalho evitava as restrições das guildas e oferecia treinamento em larga escala. No entanto, a questão central permanece: como exatamente a Grã-Bretanha se industrializou?

Altos Salários e Inovação Induzida. Allen (2009) argumenta que a Revolução Industrial na Grã-Bretanha foi uma resposta aos preços relativos de trabalho, capital e energia, sendo explicada pela teoria da “inovação induzida” ou “mudança tecnológica tendenciosa”. Segundo Allen, as tecnologias que economizavam trabalho foram lucrativas na Grã-Bretanha devido aos altos salários, enquanto o capital e a energia eram relativamente baratos. Sua teoria se desenvolve em duas fases: na primeira, os preços relativos dos fatores determinam a escolha de tecnologias, favorecendo aquelas intensivas em capital quando o trabalho é caro; na segunda, ocorre rápido progresso tecnológico nas tecnologias intensivas em capital, tornando-as viáveis até em locais com custos de trabalho baixos.

Allen atribui os altos salários britânicos ao sucesso comercial nos séculos XVII e XVIII, impulsionado pelo comércio exterior, e à energia barata, proveniente da substituição de madeira por carvão devido à escassez de madeira no século XVI. No entanto, sua teoria foi alvo de críticas. Primeiramente, mudanças tecnológicas que buscam economizar o fator mais caro dependem da substituíbilidade entre trabalho e capital, e ainda não há evidências claras de que essas condições se aplicaram na Grã-Bretanha durante a Revolução Industrial. Em segundo lugar, estudos como o de Styles (2021) mostram que a adoção de novas tecnologias não foi determinada pelo nível geral de salários, mas sim pelo custo específico do trabalho em centros têxteis como Lancashire. Além disso, Humphries e Schneider (2019) refutam a ideia de Allen de que as mulheres tecelãs recebiam salários elevados.

Allen também associa altos salários ao custo elevado do trabalho, mas ignora que salários eleva-

dos podem refletir altos níveis de capital humano, e não apenas o custo do trabalho. Trabalhadores britânicos eram mais bem alimentados, mais altos, mais qualificados e com maior capacidade de trabalho, o que pode explicar seus salários mais altos (Kelly e Ó Gráda, 2014). Por fim, a explicação de Allen não considera a grande oferta de trabalhadores qualificados nem o papel crucial dos avanços científicos, que foram determinantes na Revolução Industrial, especialmente no desenvolvimento do motor a vapor e na metalurgia.

Uma Economia Iluminista. Mokyr (2009, 2016) propõe uma explicação para esse enigma, destacando o papel das elites culturais do Iluminismo na inovação tecnológica europeia, especialmente na Grã-Bretanha. Ele enfatiza a importância da República das Letras, um fórum intelectual europeu onde ideias científicas e filosóficas eram debatidas e refinadas. Esse ambiente incentivava a inovação ao eliminar ideias ruins e difundir as boas. Embora cientistas notáveis tenham sido fundamentais, Mokyr (2016) argumenta que a disseminação da ideologia do progresso envolveu um amplo grupo de intelectuais menos conhecidos, culminando na criação de um mercado de ideias. Essa cultura de crescimento foi impulsionada por fatores como o desenvolvimento de serviços postais e a fragmentação política europeia, que permitiu a fuga de pensadores inovadores de governos opressores, promovendo uma cultura iluminista pan-europeia.

Mas por que essa rede intelectual contribuiu para a industrialização na Grã-Bretanha e não em outros países? Mokyr (2009) destaca que não bastavam apenas ideias iluministas; era essencial uma classe de artesãos e técnicos qualificados para aplicá-las. No limiar da Revolução Industrial, a Inglaterra possuía essa mão de obra em abundância, ao contrário de outras regiões. Crafts (2011) sugere que as visões de Allen e Mokyr sobre a Revolução Industrial podem ser conciliadas. O progresso tecnológico pode ter sido uma resposta aos altos custos do trabalho, como argumenta Allen, enquanto a transformação cultural descrita por Mokyr foi crucial para expandir a oferta de inventores e empreendedores. Embora a influência cultural pareça mais determinante, ambas as abordagens contribuem para entender por que a industrialização começou na Grã-Bretanha.

Em resumo, apenas a Grã-Bretanha reunia todas as condições necessárias para a industrialização: um governo relativamente limitado e representativo, um grande mercado interno, acesso às economias atlânticas e uma ampla base de trabalhadores mecânicos altamente qualificados. Muitas dessas condições interagiam entre si, tornando-se relevantes apenas na presença de outras.

Duas teorias principais explicam por que a Grã-Bretanha foi pioneira na industrialização. Allen enfatiza os altos salários britânicos e a energia relativamente barata, fatores que estimularam a ino-

vação em tecnologias poupadoras de trabalho. No entanto, sua explicação não esclarece por que a Grã-Bretanha possuía tantos trabalhadores mecânicos altamente qualificados na véspera da industrialização – essenciais para a criação, implementação e manutenção das invenções da época. Além disso, o argumento de Allen não considera o papel da ciência em diversas inovações fundamentais da Revolução Industrial.

Esses aspectos são abordados por Mokyr, que propõe a existência de um “Iluminismo Industrial” na Grã-Bretanha entre o final do século XVIII e o início do XIX. Segundo ele, o país soube aproveitar os avanços científicos europeus e, ao contrário de outras regiões, conseguiu aliar conhecimento científico e habilidades mecânicas, criando um ambiente propício à inovação contínua. Essa combinação garantiu que a industrialização britânica não fosse um fenômeno passageiro, mas sim o ponto de partida para um ciclo de inovação que persiste até os dias de hoje.

3.6 Demografia, Fecundidade e Crescimento Econômico

A população tem sido tanto causa quanto efeito do desenvolvimento econômico. Melhores padrões de vida geralmente incentivam o aumento populacional, mas uma população maior, com uma área fixa de terra para cultivo, tende a reduzir os padrões de vida. Até o início do crescimento econômico moderno, os padrões de vida permaneceram estagnados ou variaram ao redor de um nível baixo devido a guerras, fomes, pestes e mudanças climáticas. Com o aumento persistente da produtividade, muitas vezes associado a revoluções industriais, uma segunda fase começa, na qual a mortalidade cai e a fecundidade aumenta, fazendo a população crescer rapidamente. Em uma terceira fase, conhecida como transição demográfica, mudanças sociais e tecnológicas reduzem a fecundidade, estabilizando a população em um nível mais alto. A Inglaterra é geralmente vista como o exemplo desse padrão, enquanto a França é uma exceção, com a fecundidade caindo antes de evidências de crescimento sustentado da produtividade.

A história econômica tenta vincular essas mudanças demográficas históricas ao desenvolvimento econômico e ao aumento dos salários reais, explicando a fecundidade e o acúmulo de capital humano como resultado de decisões racionais das famílias, que influenciam a evolução de seus ambientes. Porém, a população de um país pode ser influenciada não apenas pelo aumento natural, mas também pela migração. A partir do século XIX, com maior controle estatal e interesse por estatísticas, dados migratórios se tornaram mais disponíveis.

Período Malthusiano. A primeira fase da economia pode ser analisada a partir do modelo fundamental de Malthus (1798), que utiliza duas equações com salários e população endógenos. Malthus acreditava que o aumento populacional restringia os salários a níveis de subsistência no longo prazo, o que foi uma contribuição importante para o entendimento da interação entre economia e demografia. Embora a teoria de Malthus tenha se mostrado limitada com o tempo, ela ainda fornece uma estrutura útil para estudar o crescimento econômico. Sua teoria sugere que o crescimento populacional, quando não controlado, leva a uma queda nos salários para níveis de subsistência.

Malthus identificou dois tipos de “freios” para controlar o crescimento populacional: freios positivas, como fome e peste, que reduzem a população, e freios preventivas, como o adiamento do casamento e o controle da natalidade. Ele argumentou que políticas de assistência social, como o aumento dos pagamentos da Lei dos Pobres com base no número de filhos na Inglaterra, incentivariam o crescimento populacional e minariam a independência econômica. De fato, Boyer (1989) mostrou que salários mais altos estavam associados a mais nascimentos e que os subsídios por filhos estimulavam mais nascimentos. No entanto, Kelly e O Grada (2014) mostraram que a introdução de ajuda sistemática aos pobres na Inglaterra contribuiu para quebrar a ligação entre quebras de colheitas e mortalidade em massa.

Considere o seguinte modelo de tempo discreto do processo Malthusiano. A relação entre os salários e a população, onde os salários caem com o aumento da população devido à diminuição da produtividade marginal do trabalho, pode ser representada por:

$$w_t = a - bP_t + u_t \quad (1)$$

Onde w_t é o salário, P_t é a população, a (função da urbanização e da produtividade total dos fatores na agricultura) e b (resposta dos salários à pressão populacional) são parâmetros, e u_t representa choques (como peste ou fome) com média zero.

A população é influenciada pelos salários do período anterior, com um aumento nos salários levando a um aumento na população devido a um maior número de casamentos e nascimentos (filhos são um bem normal):

$$P_t - P_{t-1} = c + dw_{t-1} + v_t, \quad c < 0, d > 0 \quad (2)$$

onde c (negativo, pois quando o salário é menor que o nível de subsistência, o número de mortes excede o número de nascimentos e a população diminui) e d são parâmetros e v_t representa choques

com média zero.

Das equações (1) e (2), podemos encontrar o equilíbrio de longo prazo, onde a população e os salários convergem para valores estáveis. Substituindo (1) em (2) e assumindo que os termos de choque assumem o valor médio deles:

$$P_t + (bd - 1)P_{t-1} = c + ad \quad (3)$$

O equilíbrio de longo prazo (steady-state) é

$$P^* = (c + ad)/bd \quad (4)$$

A população convergirá para esse valor se choques causarem desvios iniciais em relação a ele. Por exemplo, uma idade de casamento mais alta (c menor) reduz o crescimento populacional e a população de equilíbrio de longo prazo. A equação de salários é:

$$w_t + (bd - 1)w_{t-1} = -bc \quad (5)$$

O equilíbrio de longo prazo (steady-state) é

$$w^* = -c/d \quad (6)$$

Um aumento em c (queda na mortalidade) reduz o salário de equilíbrio de longo prazo. Por outro lado, se a urbanização for suficientemente importante para aumentar as taxas de mortalidade, c diminuiria e, portanto, os salários de equilíbrio aumentariam.

Embora, pela equação (1), padrões de vida mais altos (w) sejam alcançados no curto prazo devido a um aumento na produtividade (a), a pressão populacional garante que eles não sejam sustentados no longo prazo (a menos que a aumente continuamente devido ao progresso técnico):

$$\frac{\partial w^*}{\partial a} = 0$$

No entanto, a população aumenta com melhorias na produtividade tanto no longo prazo quanto no curto prazo:

$$\frac{\partial P^*}{\partial a} = \frac{1}{b}$$

Em resumo, duas características desse modelo dinâmico simples são que os ajustes a choques são de longa duração e que aumentos na produtividade elevam a população e os salários no curto prazo, mas apenas a população no longo prazo. Existem evidências de que, durante a maior parte da história mundial, existiu um equilíbrio Malthusiano estável entre salários e população (Ashraf e Galor, 2011).

Crescimento Econômico e Transição Demográfica. O equilíbrio Malthusiano pode ser definido como a primeira fase da interação entre demografia e economia. A segunda fase refere-se à saída desse equilíbrio, quando a população cresce significativamente e os salários reais começam a aumentar. Esse padrão sugere uma aceleração na mudança técnica, representada pela elevação contínua do coeficiente a (e possivelmente uma redução de b), permitindo que ganhos de produtividade compensem os retornos decrescentes do crescimento populacional. Na terceira fase, a transição demográfica, o parâmetro d da equação (2) se aproxima de zero, destacando a importância do controle da fecundidade, tornando irrelevante o modelo Malthusiano e sustentando o crescimento econômico.

Historiadores econômicos frequentemente utilizam a escolha racional das famílias para explicar a transição do equilíbrio Malthusiano, considerando trade-offs entre número de filhos e investimento em capital humano. Evidências sugerem que casais pré-modernos influenciavam sua taxa de natalidade e priorizavam a qualidade dos filhos (Cinnirella et al 2017; Klemp e Weisdorf 2019), embora essa visão seja contestada por alguns autores (Clark e Cummins 2019). Nesse sentido, o casamento tardio feminino pode ter limitado a fecundidade e favorecido o crescimento econômico. Porém, seu papel exato ainda é controverso.

Paradoxalmente, em uma versão da teoria do crescimento unificado, o aumento populacional, que tem um efeito paralisante na teoria malthusiana, proporciona uma saída (Galor e Weil 2000). Um gatilho alternativo para o crescimento econômico (aumento de a) é a seleção natural da população em favor de uma maior qualidade dos filhos (Galor e Moav 2002; Ashraf e Galor 2018; Foreman-Peck e Zhou 2020). Outro estímulo para a ruptura com o equilíbrio malthusiano é o declínio da mortalidade no século XIX que, ao contrário das previsões pessimistas, levou a menores taxas de natalidade, maior expectativa de vida e maiores retornos à educação, acelerando o progresso técnico (Boucekkine et al. 2003; Lagerlöf 2003; Cervellati e Sunde 2005).

A transição demográfica na fase 3 pode ser interpretada como uma inovação (disseminação do conhecimento sobre contracepção) ou uma adaptação a novas condições econômicas (restrições

ao trabalho infantil, aumento do custo dos filhos devido à maior importância do capital humano, mudanças culturais e queda na mortalidade). Uma explicação central para a transição demográfica é o aumento da demanda por capital humano (Galor 2012). O progresso tecnológico elevou os salários reais, aumentando o custo de oportunidade dos filhos e reduzindo a taxa de crescimento populacional.

3.7 Cultura, Religião e Crescimento Econômico

Recentemente, economistas e historiadores econômicos “redescobriram” a religião como um fator significativo do crescimento econômico de longo prazo. O ensaio a seguir apresenta um breve levantamento de importantes contribuições para essa literatura com o objetivo de explorar a relação entre religião, instituições religiosas e crescimento econômico. Ênfase especial é dada ao papel da religião na acumulação histórica de capital humano e à interseção entre religião e economia política histórica, dois canais através dos quais a religião teve um efeito particularmente pronunciado no crescimento econômico. Importante, essa literatura dá pouca ênfase à doutrina dessas religiões, focando, em vez disso, nos mecanismos institucionais e políticos.

O ensaio está estruturado da seguinte forma: a próxima seção discute a relação entre religião e capital humano na história, abordando as três principais religiões abraâmicas (Cristianismo, Islamismo e Judaísmo) e o papel dos missionários na disseminação do capital humano. A seção seguinte revisa a literatura sobre religião na economia política histórica, focando no papel da religião na legitimação do Estado nas sociedades islâmicas e protestantes e o impacto disso nas trajetórias econômicas desses Estados.

Religião e Capital Humano. A religião impacta a acumulação de capital humano, principalmente por meio de incentivos à educação. Até o final do século XIX, escolas religiosas predominavam na Europa e nos EUA, com missionários desempenhando papel crucial nas colônias europeias. Por dois milênios, os judeus tiveram vantagem de capital humano em relação a cristãos e muçulmanos, o que lhes permitiu se especializar em áreas como finanças e medicina. Moser et al. (2014) mostram que emigrantes judeus alemães revolucionaram invenções nos EUA. Essa vantagem persiste, como evidenciado pela alta proporção de judeus entre laureados com o Nobel. Botticini e Eckstein (2005, 2007) argumentam que a destruição do Segundo Templo em Jerusalém em 70 d.C. mudaram as trajetórias institucionais e culturais do judaísmo, favorecendo a acumulação de capital humano. Já Abramitzky e Halaburda (2020), por sua vez, alertam que a maior

educação dos judeus na Polônia entre as duas guerras era em parte um efeito composicional, já que os judeus eram mais propensos a viver em áreas urbanas, onde as taxas de educação eram mais altas.

O Protestantismo é a religião mais estudada na literatura sobre religião e capital humano, em grande parte devido ao trabalho pioneiro de Becker e Woessmann (2009), os quais buscaram uma explicação alternativa para a famosa tese da ética protestante de Weber. Estudando a mesma região onde Weber cresceu (a Prússia do final do século XIX), seus dados censitários indicam que os protestantes tinham melhores empregos e salários mais altos do que os católicos. Becker e Woessmann argumentam que a vantagem protestante estava enraizada na educação, e não em uma "ética de trabalho" superior. Eles atribuem essa vantagem à ênfase que Martin Lutero colocou na leitura da Bíblia, o que levou os protestantes a dar mais atenção à educação e a aumentar o acesso à educação por meio de escolas, oferecendo mais acesso à educação para meninas do que nas sociedades católicas, o que resultou em uma lacuna de gênero menor na educação nos países protestantes. Além disso, Binzel et al. (2022) argumentam que a reforma levou à "vernacularização" da língua, com Lutero e outros reformadores escrevendo em alemão, o que aumentou a produção de obras em línguas vernáculas e foi associado ao crescimento do capital humano de alto nível e do crescimento populacional.

A vantagem educacional protestante também se estende além da Prússia do século XIX. Em estudos complementares da Suíça do final do século XIX, Boppart et al. (2013, 2014) encontraram uma vantagem educacional protestante significativa sobre os católicos. Seus resultados qualificam os encontrados por Becker e Woessmann em certa medida. Boppart et al. (2013) estudaram as diferenças no fornecimento de educação, encontrando que o gasto com educação era, de fato, mais baixo em regiões predominantemente católicas, mas esse resultado só se aplicava em áreas conservadoras que se opunham à secularização e às regulamentações de trabalho. Boppart et al. (2014) descobriram que a vantagem protestante se estendia além da educação, abrangendo matemática, redação e história.

Um dos mecanismos mais importantes através dos quais as igrejas protestante e católica afetaram a acumulação de capital humano foi o trabalho dos missionários. Um dos principais efeitos das atividades missionárias na Ásia, África e América Latina foi a disseminação do capital humano, particularmente a alfabetização. Estudos sobre atividades missionárias em regiões específicas, como na África, mostram que a competição entre missões levou a mais escolas e centros de saúde, com missões protestantes oferecendo mais escolas do que as católicas (Wantchekon et al.

2015; Cagé e Rueda 2016). Na América do Sul, missões jesuítas também tiveram impactos positivos de longo prazo na educação e na renda (Valencia Caicedo, 2019). Esse efeito persistente se deve a mecanismos como transmissão cultural, transformação econômica e adoção de tecnologias agrícolas. Na Índia, missões protestantes aumentaram a educação, especialmente entre mulheres (Calvi et al., 2022).

No Islã, a relação com o capital humano apresenta efeitos mistos. Chaney (2016) documenta uma transição no Oriente Médio, onde, até cerca de 1100 d.C., obras científicas predominavam no discurso islâmico, sendo depois substituídas por obras religiosas. O Islã também influenciou a acumulação de capital humano por meio das políticas educacionais de governantes muçulmanos, que priorizavam a educação religiosa em detrimento da secular. Chaudhary e Rubin (2016) analisaram o impacto da identidade religiosa dos governantes na educação, constatando que, em territórios muçulmanos, os hindus apresentavam taxas de alfabetização mais baixas, enquanto os muçulmanos mantinham taxas semelhantes, independentemente do regime.

Religião e Economia Política. A religião tem influenciado os resultados econômicos por meio da economia política de diversas maneiras. Entre os mecanismos principais estão a legitimação religiosa, a presença de leis religiosas e a governança territorial. Desde os primórdios das religiões institucionalizadas, a religião tem sido uma ferramenta para legitimar o poder político. As autoridades religiosas, com sua sabedoria espiritual, podem influenciar as crenças e ações de seus seguidores, o que as torna poderosas para fins de controle social. Em troca, podem obter influência política, o que tem implicações econômicas importantes.

A legitimação religiosa foi crucial para os resultados políticos e econômicos, especialmente no mundo muçulmano. Rubin (2017) propõe um modelo explicando que as pessoas seguem um governante por coerção ou legitimidade. A religião tem sido uma importante fonte de legitimidade, com autoridades religiosas atuando como “agentes de legitimação”. No caso do Islã, a religião surgiu com o império, fornecendo uma base sólida para a legitimação do governo. Isso contrastou com o cristianismo, que se originou em um império com instituições políticas pré-existentes.

No Ocidente, a Revolução Comercial de 1066 levou os governantes a buscar fontes alternativas de legitimação, especialmente após a Reforma Protestante enfraquecer a Igreja. No mundo islâmico, no entanto, a legitimidade religiosa permaneceu forte até o século XX, com as autoridades religiosas tendo mais poder político. Estudos como o de Kuru (2019) e Blaydes et al. (2018) corroboram a ideia de uma aliança entre o clero e o estado islâmico, que manteve as elites econômicas

fora das coalizões de poder até o século XX. No entanto, Rubin (2011) destaca que a abordagem dos empréstimos e a proibição de juros nos dois mundos religiosos (cristão e islâmico) foram moldadas por essas diferenças de legitimação religiosa.

Em países muçulmanos, o controle religioso sobre a lei comercial foi duradouro. Kuran (2005, 2011) explica que a lei islâmica, combinada com práticas como as parcerias e os waqfs (fundos religiosos), limitava a demanda por mudanças legais, o que dificultava a criação de organizações comerciais maiores e mais complexas. No entanto, essas formas rígidas de organização, como o waqf, também protegiam a riqueza, o que fortaleceu a posição das autoridades religiosas, como visto na Indonésia (Bazzi et al., 2020). Além disso, as cortes islâmicas eram parciais, favorecendo homens, muçulmanos e elites, o que gerava desigualdade e impedia transações mais fluidas com os grupos desfavorecidos. Isso teve implicações econômicas importantes, pois elevou as taxas de juros para os privilegiados, afetando as oportunidades de negócios.

Fora do mundo islâmico, a religião também teve um papel central. A Reforma Protestante levou os governantes protestantes a buscar novas fontes de legitimidade, principalmente entre as elites econômicas e parlamentos, o que impulsionou políticas favoráveis ao comércio. Isso se traduziu em mudanças significativas, como a transição para a monarquia constitucional na Inglaterra e o fortalecimento da economia na Holanda (Rubin, 2017). Na Alemanha medieval, as mudanças nas formas de legitimação política ajudaram a secularizar o poder, o que também teve impacto econômico, como o aumento da educação secular e do poder do estado (Cantoni et al., 2018). Em países protestantes, a maior tolerância religiosa e a alocação mais eficiente de recursos agrícolas também contribuíram para o crescimento econômico, especialmente durante a Revolução Agrícola (Heldring et al., 2021).

Porém, a religião também contribuiu para a perseguição de minorias. Johnson e Koyama (2019) argumentam que, antes da Reforma, os judeus eram “tolerados condicionalmente” pelos governantes, mas com a secularização após a Reforma, a tolerância se generalizou. No entanto, em algumas regiões protestantes, a concorrência religiosa aumentou a perseguição, como no caso dos judeus na Alemanha (Becker e Pascali, 2019). Por fim, a religião também influenciou o surgimento do autogoverno e do estado moderno. Bentzen e Gokman (2023) destacam que estados com forte dependência da legitimidade religiosa eram mais autocráticos, e a Igreja medieval, com seu poder de legitimação, desempenhou um papel crucial na fragmentação política da Europa.

3.8 Crise de 1929 e Grande Depressão

Nos Estados Unidos, o termo “Grande Depressão” refere-se a um extenso período de atividade econômica deprimida que se estendeu de 1929 até o início da Segunda Guerra Mundial. O colapso começou com uma queda após o pico de 1929, alcançando seu ponto mais baixo em 1933. Embora tenha ocorrido uma rápida expansão, antes que o emprego e a produção retornassem às tendências anteriores a 1929, uma nova contração se manifestou em 1937.

Historicamente, a Grande Depressão é frequentemente estudada como um exemplo de políticas fiscais e monetárias inadequadas. Este ensaio explora a crise de 1929 à luz da teoria macroeconômica novo-Keynesiana, visando identificar eventos e mecanismos relevantes. A depressão de 1929-1933 e a recuperação posterior podem ser explicadas pelos modelos novo-Keynesianos que associam a atividade econômica e a inflação a choques nas taxas de juros e às imperfeições do mercado financeiro. A queda inicial em agosto de 1929 foi causada por um aumento nas taxas de juros entre 1928 e 1929, quando o Fed tentou conter a bolha do mercado de ações. A profundidade da depressão subsequente foi amplificada por novos aumentos nas taxas de juros em 1931 e no início de 1933, juntamente com uma crise financeira que resultou no fechamento de quase todos os bancos, interrompendo o crédito.

A recuperação da economia em 1933 deveu-se a uma política monetária mais frouxa e à reabertura dos bancos após a desvalorização do dólar promovida por Roosevelt. A inflação e os salários reais durante a recuperação foram consistentes com outros ciclos econômicos, exceto por surtos anômalos de inflação em 1933-1934 e em 1936-1937, que foram atribuídos a intervenções regulatórias do New Deal, como o NIRA e o Wagner Act.

Este ensaio começa com uma base teórica novo-Keynesiana para fundamentar a discussão sobre o período da Grande Depressão de 1929-1933 e sua recuperação subsequente de 1933-1936.

Modelo Novo Keynesiano. Considere um modelo simples de economia fechada de estrutura novo-keynesiana, que possui a seguinte equação “IS novo-keynesiana”:

$$y_t = y_{t+1}^e - \alpha r_t + \varepsilon_{1t} \quad (7)$$

e a equação da “curva de Phillips novo-keynesiana”:

$$\pi_t = \delta \pi_{t+1}^e + \gamma y_t, \text{ onde } 0 < \delta < 1$$

Neste contexto, x^e representa a expectativa dos agentes sobre uma variável futura x , y é o hiato do produto, ou seja, a diferença entre o logaritmo do PIB real e o logaritmo do PIB real no estado estacionário de longo prazo não estocástico da economia. r é o diferencial entre a taxa de juros real de curto prazo (sem risco de inadimplência) e a “taxa natural de juros” (a taxa na qual $y = y_{t+1}^e = 0$ quando $\varepsilon_1 = 0$). Por fim, π é a taxa de inflação do nível de preços.

Essas equações podem ser derivadas de um modelo em que não há investimento em capital (ou seja, o estoque de capital é fixo, ou a produção é gerada apenas pelo trabalho), e todos os agentes têm liberdade para tomar emprestado ou emprestar à taxa de juros sem risco de inadimplência, com o objetivo de maximizar uma função padrão de utilidade esperada de consumo e lazer. O regime monetário é configurado de modo que a taxa de inflação seja zero no estado estacionário de longo prazo não estocástico. Os valores esperados das variáveis pelos agentes, em um horizonte suficientemente distante, são aproximados pelos valores de longo prazo. Na maioria dos modelos, as expectativas dos agentes são racionais no sentido usual e uma condição mais fraca (correspondente à “lei das expectativas iteradas”) é suficiente para que as relações centrais do modelo se mantenham válidas.

O termo de perturbação ε na equação IS reflete flutuações nos parâmetros da função de utilidade ou compras de produção pelo governo, financiadas por receita tributária sob a equivalência ricardiana. As suposições que tornam o modelo keynesiano criam “rigidez nominal” nos preços: os mercados de produtos são monopólios, os ajustes nos preços estão sujeitos a uma restrição de Rotemberg (1982) ou Calvo (1983), e o valor padrão de um preço não ajustado é o preço do período anterior.

Resolvendo a equação (1) a partir do estado estacionário de longo prazo esperado, obtemos

$$y_t = -\alpha r_t + \varepsilon_{1t} + \sum_{\tau=0}^{\infty} [-\alpha r_{t+1+\tau}^e + \varepsilon_{1t+1+\tau}^e] = \sum_{\tau=0}^{\infty} [-\alpha r_{t+\tau}^e + \varepsilon_{1t+\tau}^e] \quad (8)$$

A atividade real depende da trajetória esperada das taxas de juros reais de curto prazo desde o presente até o infinito. Isso difere significativamente da curva IS keynesiana tradicional, que expressava uma relação negativa entre a atividade real e as taxas de juros reais correntes ou defasadas. De maneira similar, a equação (2) resulta em:

$$\pi_t = \gamma y_t + \gamma \sum_{\tau=0}^{\infty} \delta^{\tau+1} y_{t+\tau+1}^e = \gamma \sum_{\tau=0}^{\infty} \delta^{\tau} y_{t+\tau}^e \quad (9)$$

A inflação depende da trajetória esperada da atividade real até um horizonte infinito. Na curva de Phillips keynesiana tradicional, a inflação estava associada à atividade real corrente. No entanto, sob uma condição plausível, as equações (3) e (4) geram resultados semelhantes às relações keynesianas tradicionais. A condição é que os valores esperados de produção futura pelo público sejam geralmente próximos às previsões autoregressivas simples. Isso não implica uma afirmação sobre o grau de racionalidade das expectativas do público, mas sobre a correlação incondicional entre previsões e o produto corrente, como observado nos dados. De fato, muitos estudos demonstram que previsões de produto futuro (como o PIB real) geradas por métodos sofisticados são próximas das previsões AR, devido à forte correlação serial da atividade real (Chauvet e Potter 2013). Assim, assume-se que os valores esperados do público para a produção futura possam ser descritos como:

$$y_{t+\tau}^e = \rho^\tau y_t + z_{t+\tau}^e, \text{ onde } 0 < \rho < 1 \quad (10)$$

onde $z_{t+\tau}^e$ captura as diferenças entre uma previsão AR(1) e a previsão real do público. Portanto, temos:

$$y_t = -\frac{\alpha}{1-\rho} r_t + \varepsilon_{2t}, \text{ onde } \varepsilon_{2t} = \frac{1}{1-\rho} [\varepsilon_{1t} + z_{t+1}^e] \quad (11)$$

$$\pi_t = \gamma \frac{1}{1-\delta\rho} y_t + \gamma \sum_{\tau=1}^{\infty} \delta^\tau z_{t+\tau}^e \quad (12)$$

Os dados gerados por essa economia exibiriam uma relação negativa entre a atividade real e a taxa de juros real corrente, bem como uma relação positiva entre inflação e atividade real corrente. Anomalias nessas relações surgiriam em momentos em que eventos incomuns levassem o público a formular previsões atípicas.

Modelos novo-keynesianos complexos incorporam vários outros elementos como, por exemplo, investimento em um estoque de capital variável e fricções salariais semelhantes às de preços. Pode existir uma taxa de desemprego negativamente relacionada ao produto, influenciada por choques de mark-up salarial. Nos modelos que incluem "imperfeições do mercado financeiro", restrições de informação limitam o acesso ao crédito, tornando o consumo de agentes com crédito restrito sensível à renda corrente, amplificando os efeitos das variações nas taxas de juros reais sobre a atividade econômica.

Adicionando alguma combinação desses elementos a essa dinâmica, uma relação entre a ativi-

dade real (y) e as taxas de juros (r) pode ser expressa como:

$$y_t = - \sum_{\tau=0}^i \beta_{yr\tau} r_{t-\tau} + \sum_{\tau=1}^j \beta_{yy\tau} y_{t-\tau} + \varepsilon_{yt} \quad (13)$$

O termo de perturbação, ε_y , é influenciado pelos gastos do governo e por previsões atípicas da atividade real, além de outros fatores, como a política fiscal e eventos externos. A atividade real também pode ser afetada por redistribuições de riqueza que impactam agentes com alta propensão a consumir (Auclert 2019). As curvas de Phillips observáveis podem ser expressas da seguinte forma:

$$\pi_t^P = \beta_{\pi py} y_t + \varepsilon_{\pi pt} \quad (14)$$

$$\pi_t^W = \beta_{\pi wy} y_t + \varepsilon_{\pi wt} \quad (15)$$

Onde π^P e π^W referem-se à inflação de preços e salários, respectivamente. O comportamento cíclico dos salários reais é determinado pela rigidez nominal dos salários em relação aos preços. Assim, os salários reais podem ser pró-cíclicos ($\beta_{\pi wy} > \beta_{\pi py}$), contra-cíclicos ($\beta_{\pi wy} < \beta_{\pi py}$) ou acíclicos ($\beta_{\pi wy} \approx \beta_{\pi py}$). As perturbações na Curva de Phillips, $\varepsilon_{\pi P}$ e $\varepsilon_{\pi W}$, podem ser afetadas por choques exógenos nos mark-ups de salários e preços, bem como por previsões sobre a atividade real futura.

Crises Financeiras e o Mecanismo de Expectativas Novo-Keynesiano Uma “crise financeira” se manifesta quando ocorre uma corrida em massa contra intermediários financeiros, resultando na retirada rápida de empréstimos por credores de curto prazo, o que força os intermediários a liquidar ativos, incluindo empréstimos “ilíquidos” que são difíceis de vender sem perdas significativas. Essa situação pode levar à inadimplência ou à suspensão de pagamentos, reduzindo a oferta de crédito para tomadores com restrições financeiras.

O modelo novo-keynesiano apresentado por Gertler et al. (2017) explica que tais corridas, que são impulsionadas por condições econômicas adversas, podem reduzir a atividade real e amplificar os efeitos de choques de demanda. A ocorrência de corridas pode ser descontínua e, quando uma ocorre, aumenta a probabilidade de outras corridas devido ao contágio e à complementaridade estratégica. Para prevenir ou mitigar crises financeiras, podem ser implementados sistemas de seguro para compensar credores, ou o uso de um "emprestador de última instância" que fornece recursos

a intermediários em troca de ativos ilíquidos a preços superiores aos de venda forçada; o banco central é a instituição mais qualificada para desempenhar essa função devido à sua capacidade de criar fundos.

Adicionalmente, a armadilha de liquidez ocorre quando a taxa de juros nominal atinge zero, resultando na perda de eficácia da política monetária. Nesse cenário, a autoridade monetária não consegue estimular a atividade econômica apenas ajustando a taxa de juros. Isso agrava a situação de recessão, potencializando um cenário de deflação.

Nos modelos novo-keynesianos, assume-se que os agentes acreditam que a armadilha de liquidez é temporária, com a expectativa de que o hiato do produto e a inflação retornarão a zero em algum momento. Além disso, esses modelos indicam que o hiato do produto atual se baseia nas expectativas sobre as taxas de juros reais futuras. Mesmo com a taxa de juros nominal em seu limite inferior, os formuladores de políticas podem influenciar a atividade econômica ao manipular as expectativas do público sobre políticas futuras. Eles podem impulsionar a atividade imediata convencendo o público de que promoverão um aumento da taxa de inflação acima de zero por um período, antes de retorná-la a zero, como se espera no estado estacionário a longo prazo. Em resumo, os formuladores de políticas precisam elevar as expectativas do público em relação ao nível de preços que prevalecerá no equilíbrio de longo prazo.

Para ilustrar, considere um ponto no tempo $t+T$ e retorne ao modelo simples descrito nas equações (1) a (4). O valor esperado do logaritmo do nível de preços em $t+T$ é:

$$p_{t+T}^e = p_{t-1} + \frac{1}{1-\rho} \sum_{\tau=0}^T (1 - \delta^{1+\tau} \gamma y_{t+\tau}^e) \quad (16)$$

Essa expressão indica que um aumento em p_{t+T}^e é equivalente a um aumento nas expectativas do hiato do produto em períodos futuros. Com base na equação (3), isso resulta em uma redução na taxa de juros real esperada para esses períodos, o que, igualmente a partir de (3), deve elevar a atividade real atual, independentemente da taxa de juros real vigente. Além disso, o aumento em p_{t+T}^e gera um incremento nas expectativas de produto futuro acumulado, o que impulsiona a inflação atual (equação (4)). Esse efeito tende a reduzir a taxa de juros real atual, estimulando ainda mais a atividade econômica. Assim, os formuladores de políticas podem tirar uma economia de uma armadilha de liquidez por meio da manipulação das expectativas, conhecido como "mecanismo de expectativas novo-keynesiano".

A Grande Depressão 1929-1933 Os modelos novo-keynesianos destacam que o regime monetário e a estrutura dos mercados financeiros de uma economia influenciam sua resposta a eventos exógenos. Às vésperas da Grande Depressão, os Estados Unidos eram mais vulneráveis a crises financeiras do que países como Reino Unido, Austrália e Canadá, onde os bancos operavam em redes nacionais, minimizando riscos de inadimplência. Nesses países, os bancos utilizavam títulos de curto prazo garantidos por empresas solventes e contavam com o Banco da Inglaterra como prestador de última instância, essencial para a liquidez.

Em contraste, a ausência de um banco central nos EUA até 1914 e as regulamentações que restringiam os bancos a operarem em uma única região tornavam seu sistema financeiro mais frágil. A falta de títulos de curto prazo e a dependência de depósitos interbancários contribuía para crises recorrentes, resultando em suspensões de pagamentos e paralisando a atividade econômica. No início da Grande Depressão, mesmo com restrições às filiais bancárias, os EUA já contavam com um banco central, representando uma mudança estrutural significativa em seu sistema financeiro.

O Sistema do Federal Reserve, ou "Fed", iniciou suas operações em 1914 como uma confederação de 12 bancos supervisionados por um conselho central em Washington. Os bancos membros eram obrigados a manter reservas mínimas em contas de reservas, utilizadas para transações entre si. No final da década de 1920, surgiu um mercado de empréstimos overnight, conhecido como "federal funds", junto ao tradicional mercado de call money. Embora os bancos nacionais fossem obrigados a aderir ao Fed, muitos bancos estaduais optaram por não fazê-lo devido a custos regulatórios e à falta de juros sobre saldos.

Os idealizadores do Fed pretendiam que ele atuasse como o Banco da Inglaterra, interagindo com os bancos por meio do mercado de aceites, mas a participação de bancos fora de Nova York foi limitada. A legislação restringiu sua capacidade de emprestar, permitindo que apenas bancos membros recebessem empréstimos, garantidos por aceites bancários, dívida federal ou certos "papeis elegíveis" de curto prazo. As taxas de juros para essas operações eram chamadas de "taxas de desconto", e cada banco poderia estabelecer sua própria taxa, sujeita à aprovação do Conselho. No final da década de 1920, essas operações passaram a ser geridas pelo Comitê Federal de Mercado Aberto (FOMC).

A Doutrina Riefler-Burgess sugere que, se o Fed oferecesse menos reservas do que os bancos desejavam manter, estes precisariam tomar empréstimos na janela de redesconto, mesmo relutantemente. A oferta de "reservas livres" influenciava as taxas de juros do mercado monetário, e as alterações nas taxas de redesconto afetavam a disposição dos bancos em permanecer endividados.

Essa teoria implica que há um limite para a influência do Fed sobre as taxas de juros, levando a uma armadilha de liquidez: uma vez que a oferta de reservas livres é suficiente para eliminar os empréstimos na janela de redesconto, aumentos adicionais na oferta não impactam as condições do mercado financeiro.

No final da década de 1920, os Estados Unidos e muitos outros países adotaram o padrão-ouro, onde as autoridades monetárias trocavam moeda por ouro a preços fixos, afetando diretamente a oferta monetária e estabelecendo a paridade das taxas de câmbio com base nos preços do ouro, resultando em um sistema de taxas estáveis. Durante esse período, esperava-se que o preço do ouro em dólares permanecesse fixo, resultando em baixa inflação a longo prazo. O padrão-ouro impunha restrições às taxas de juros, forçando países deficitários a aumentá-las e países superavitários a reduzi-las. Esse sistema desempenhou um papel central na gestão das questões monetárias até 1928 nos Estados Unidos.

Em 1928, embora não houvesse sinais de inflação e a atividade econômica fosse moderadamente fraca, preocupações com um boom no mercado de ações levaram o Fed a aumentar as taxas de redesconto e vender títulos, visando restringir o crédito. Após um pico em agosto de 1929, a economia entrou em recessão, agravada pelo aumento das taxas de juros, caracterizando um choque exógeno que não estava relacionado à inflação. Modelos novo-keynesianos sugerem que a alta das taxas de juros teve um impacto acentuado, especialmente quando as expectativas de recessão se intensificaram, embora não haja evidência de expectativas pessimistas naquele momento.

O crash da bolsa em outubro de 1929 resultou em significativas perdas de patrimônio para as famílias, exacerbando a crise. Algumas teorias debatem se esse colapso foi devido a uma bolha irracional ou a um declínio racional nas expectativas de lucros. Além disso, a queda acumulada na riqueza das famílias devido ao endividamento com a compra de bens duráveis e hipotecas impactou o consumo durante a Grande Depressão. Entre 1929 e 1933, a atividade econômica dos EUA continuou a declinar em um período prolongado, acentuando a crise. Fatores como a política fiscal do governo Hoover e a queda na demanda por exportações não foram suficientes para deter esse declínio, enquanto um aumento nas taxas de juros e o crescimento das falências bancárias desempenharam papéis significativos na continuidade da recessão.

Entre 1930 e 1932, os Estados Unidos enfrentaram ondas sucessivas de falências bancárias, com a primeira ocorrendo em outubro de 1930 e a segunda em outubro de 1931, coincidente com o aumento das taxas de juros pelo Federal Reserve após a desvalorização da libra esterlina. Embora essas falências tenham sido restritas a algumas regiões, provocaram uma retirada nacional de fun-

dos, fazendo com que o total de depósitos começasse a cair durante a primeira onda e despencasse na segunda, refletindo o medo generalizado de colapsos bancários (Friedman e Schwartz, 1963, pp. 311, 313).

No final de 1932, a situação se agravou com uma nova onda de falências se espalhando do oeste para o meio-oeste, enquanto em fevereiro de 1933 bancos de Nova York enfrentavam retiradas massivas. Em resposta, os governadores estaduais decretaram “feriados bancários” e, em 6 de março de 1933, o presidente Franklin Roosevelt declarou um feriado bancário nacional, fechando todas as instituições financeiras. As falências bancárias atuaram mais como um mecanismo amplificador da crise do que como causas independentes da Grande Depressão, já que muitos bancos estavam próximos da insolvência antes das retiradas de depósitos, resultado de políticas de investimento inadequadas e da deterioração econômica. A retirada de depósitos e as falências estavam fortemente associadas à queda da atividade econômica durante esse período (Bernanke, 1983). Christiano et al. (2003) propõem um modelo novo-keynesiano que explica como a retirada de depósitos e a demanda crescente por moeda reduziram a produção agregada, elevando os custos do crédito para tomadores com restrições.

Durante a crise de liquidez de 1930 a 1932, esperava-se que o Fed atuasse como prestador de última instância, mas suas capacidades eram limitadas pelo Federal Reserve Act. Muitos bancos necessitados de empréstimos não eram membros do sistema e os que eram frequentemente não tinham os ativos adequados como garantias (Carlson e Wheelock 2016). Alguns bancos, como o Federal Reserve Bank de Atlanta, adotaram estratégias mais flexíveis, resultando em menos falências, enquanto outros, como o de St. Louis, não o fizeram (Richardson e Troost, 2009).

Em resposta à inação do Fed, o presidente Hoover e o Congresso criaram a Reconstruction Finance Corporation (RFC) em 1932, mas sua eficácia foi comprometida por leis que exigiam a divulgação dos nomes dos bancos que recebiam empréstimos, levando a corridas bancárias (Friedman e Schwartz, 1963, p. 325). Muitos formuladores do Fed não acreditavam que deveriam agir como prestadores de última instância. Originalmente, o Fed fora criado para evitar crises financeiras, com a esperança de que a “moeda elástica” se ajustasse às flutuações na demanda, mas membros influentes argumentaram que intervir como um “alívio emergencial” poderia causar mais danos do que benefícios. Meltzer (2003, p. 731) conclui que os formuladores do Fed acreditavam que as falências eram inevitáveis devido a más decisões e excessos especulativos, e que a estabilidade poderia ser restaurada apenas após a eliminação dessas causas.

A recuperação econômica de 1934 a 1936 foi notável, com o PIB real anual crescendo mais

rapidamente do que em qualquer ano após a Segunda Guerra Mundial, impulsionada por três principais fatores:

1. **Política Fiscal:** O governo aumentou as compras de bens e serviços, resultando em uma queda na razão entre impostos líquidos de transferências e a renda nacional, embora essa contribuição tenha sido considerada modesta. Um exemplo relevante de política fiscal foi o pagamento do bônus aos veteranos da Primeira Guerra Mundial em 1936. Hausman (2016) mostrou que, devido à alta propensão ao consumo dos veteranos, o pagamento do bônus resultou em um aumento do consumo, estimulou a produção e impulsionou o emprego.
2. **Política Monetária:** Em março de 1933, os Estados Unidos abandonaram o padrão-ouro, resultando na desvalorização do dólar em relação ao ouro e outras moedas. Em janeiro de 1934, os EUA retornaram ao padrão-ouro sob novas regras, com o Tesouro responsável pela troca e estabelecendo um novo preço do ouro que desvalorizou o dólar em 40%. Isso elevou os preços dos produtos agrícolas nos mercados internacionais, redistribuindo renda para os agricultores, que tinham maior propensão a consumir (Hausman et al., 2019). A desvalorização também melhorou a balança de pagamentos dos EUA, que se manteve superavitária na década de 1930, levando o Tesouro a adquirir grandes quantidades de ouro e aumentando a oferta de moeda. Economistas argumentam que a suspensão dos pagamentos em ouro, a desvalorização do dólar e o aumento da oferta de moeda criaram expectativas de inflação futura (Temin e Wigmore 1990). Durante 1933–1934, Roosevelt prometeu "reflacionar" os preços e modelos novo-keynesianos sugerem que esse compromisso poderia estimular a atividade econômica real. Com a expansão das reservas livres, a taxa overnight caiu para seu mínimo e as taxas de juros de longo prazo começaram uma queda contínua.
3. **Restabelecimento da Confiança nos Bancos:** Após o feriado bancário de março de 1933, a rápida reabertura dos bancos e a criação do seguro de depósitos pela Federal Deposit Insurance Corporation (FDIC) em janeiro de 1934 ajudaram a restaurar a confiança no sistema financeiro (Friedman e Schwartz 1963, p. 421-428). Nesse período, foi aprovada uma legislação que permitia à Reconstruction Finance Corporation (RFC) conceder empréstimos garantidos por ativos de baixa qualidade e realizar investimentos em ações de bancos para fortalecer seu capital.

Além disso, as políticas do New Deal, como o National Industrial Recovery Act (NIRA) e o

Wagner Act, fortaleceram a negociação coletiva e impulsionaram a inflação salarial, desempenhando um papel significativo na recuperação econômica. Usando um modelo novo-keynesiano padrão, Eggertsson (2012) mostra que esses choques podem estimular o produto corrente quando a taxa de juros nominal está no limite inferior, ao elevar a inflação esperada futura e reduzir a taxa de juros real corrente. Ele argumenta que o NIRA pode ter contribuído para a recuperação econômica pós-1933, amplificando o efeito do aumento da inflação esperada resultante das políticas monetárias de Roosevelt.

Por fim, a inflação salarial e de preços aumentou à medida que a atividade econômica se recuperou após março de 1933. Esse aumento, especialmente da inflação salarial, foi maior do que o esperado, considerando o alto desemprego e a produção ainda abaixo da tendência pré-Depressão. Uma explicação para esse fenômeno é o mecanismo de expectativas novo-keynesiano, que pode gerar inflação elevada em relação à atividade econômica atual. Outra explicação, compatível com os modelos novo-keynesianos, é que a inflação foi alimentada por choques de mark-up salarial resultantes das políticas do New Deal, como a NIRA de 1933 e o Wagner Act de 1935, que fortaleceram as disposições pró-sindicais. A reeleição de Roosevelt em novembro de 1936 catalisou uma onda de sindicalização e greves, levando muitos empregadores a formalizar acordos com sindicatos. Hanes (2020) argumenta que a magnitude e o timing da inflação salarial anômala entre 1933 e 1937 são mais consistentes com os efeitos das políticas do New Deal e da sindicalização do que com o mecanismo de expectativas novo-keynesiano.

Conclusão O New Deal de Roosevelt trouxe mudanças institucionais significativas, com o governo federal assumindo a responsabilidade de oferecer proteção contra crises e implementando programas emergenciais e sistemas de seguridade social. Durante os primeiros meses de seu governo, Roosevelt e o Federal Reserve reverteram políticas monetárias da década de 1930, promovendo expectativas de inflação. A reabertura dos bancos após o feriado bancário e a criação do seguro de depósitos pela FDIC restauraram a confiança no sistema financeiro.

Roosevelt retirou os EUA do padrão-ouro e o Fed focou na expansão monetária, reduzindo as taxas de redesconto. A desvalorização do dólar e o fluxo de ouro ajudaram na recuperação econômica, com o PIB real per capita se aproximando dos níveis de 1929 em 1937 e a taxa de desemprego caindo para cerca de 14%. Embora o New Deal tenha aumentado os gastos do governo em quase 70% até 1939, os déficits eram comparáveis aos da administração Hoover, com algumas políticas fiscais, como o aumento dos impostos sobre os mais ricos, afetando as pequenas empresas.

Entre as agências criadas, a Federal Emergency Relief Administration (FERA) e a Works Progress Administration (WPA) se destacaram na assistência aos pobres e na construção de obras públicas. O Social Security Act de 1935 introduziu um sistema de seguro social e assistência a viúvas, idosos e cegos, embora seu impacto nas taxas de mortalidade tenha sido limitado. O Agricultural Adjustment Act (AAA) buscou reduzir a produção agrícola e aumentar os preços, beneficiando grandes agricultores em detrimento dos trabalhadores rurais. Novas regulamentações financeiras foram implementadas após a crise, incluindo a criação da SEC para monitorar os mercados de ações e a FDIC para garantir depósitos, enquanto a Home Owners' Loan Corporation (HOLC) ajudou a refinaranciar hipotecas em risco, mantendo muitos mutuários em suas casas.

3.9 Economia da Escravidão no Brasil e EUA

A escravidão foi uma das principais formas de coerção do trabalho ao longo da história. As sociedades coloniais na América recorreram amplamente à escravidão e essa prática permaneceu predominante na maioria dos países mesmo após a independência das potências europeias. Em 1860, os escravizados ainda representavam cerca de 13% da população dos Estados Unidos. No Brasil, em 1872, a população escravizada correspondia a 15,2% do total de habitantes. Qual foi o impacto da escravidão no crescimento econômico e no desenvolvimento dessas sociedades? Este ensaio busca discutir brevemente essa questão à luz da literatura recente. Para isso, está estruturado em duas partes: a primeira apresenta a literatura sobre os Estados Unidos, enquanto a segunda examina os resultados encontrados para o Brasil.

Economia da Escravidão nos EUA. Uma questão central na literatura sobre os efeitos da escravidão nos EUA é sua real lucratividade para comerciantes e proprietários. Uma visão clássica sugeria que a escravidão era, ou poderia ser, economicamente inviável. Argumentava-se que os senhores de escravos nem sempre tinham uma mentalidade capitalista ou buscavam maximizar lucros, mas mantinham a escravidão por razões culturais, tradicionais ou para preservar uma ordem social hierárquica (Aitken, 1971). No entanto, se a escravidão fosse apenas um resquício não lucrativo de sociedades pré-capitalistas, por que sobreviveu por tanto tempo? Embora fatores sociais, culturais e religiosos tenham desempenhado um papel, os aspectos econômicos estavam no cerne dessa instituição.

Quando Conrad e Meyer escreveram seu artigo em 1958, havia uma crença predominante de que a escravidão não era lucrativa a partir da década de 1840. Além disso, mesmo nas décadas

de 1950, algumas pessoas ainda tentavam retratar a escravidão no Sul como um “serviço nobre” prestado aos negros, que eram considerados “não inteligentes” e “incompetentes”. Esses defensores da escravidão argumentavam que ela não atendia aos interesses econômicos dos proprietários de escravizados, mas que, ao contrário, tornava os negros trabalhadores produtivos, mantinha-os “contentes” e preservava a paz. Conrad e Meyer desafiaram essa visão, utilizando dados do período anterior à Guerra Civil para testar a lucratividade da escravidão. Eles demonstraram que a escravidão era, de fato, lucrativa, refutando as justificativas racistas e errôneas para a prática.

Nas décadas seguintes, outros economistas tentaram replicar os achados de Conrad e Meyer, utilizando novos conjuntos de dados e diferentes modelos baseados em suposições variadas. Esses estudos confirmaram a lucratividade da escravidão, refutando ainda mais as visões equivocadas sobre o sistema escravocrata. No entanto, embora a escravidão fosse lucrativa e viável, isso não implicava que ela fosse compatível com o crescimento ou o desenvolvimento econômico. Uma longa tradição intelectual argumenta que o relativo atraso econômico do Sul dos EUA em comparação com o restante do país (em termos de renda per capita, urbanização, indústria manufatureira e infraestrutura) foi diretamente uma consequência da economia escravista (Genovese, 1965; Wright, 1978; Bateman e Weiss, 1981; Ransom e Sutch, 2001).

Em 1974, o livro *Time on the Cross* de Robert Fogel e Stanley Engerman gerou grande controvérsia por suas conclusões. Fogel e Engerman utilizaram a cliometria para argumentar que as condições de vida dos escravizados não eram muito diferentes das dos trabalhadores livres e que os proprietários usavam incentivos positivos para motivar o trabalho. Eles também sustentaram que o comércio de escravizados não separava famílias e que o trabalho nas plantações era mais eficiente devido ao trabalho em gangue e às economias de escala. Esses argumentos foram amplamente refutados por outros economistas, que apontaram falhas metodológicas e erros significativos na pesquisa de Fogel e Engerman, concluindo que *Time on the Cross* estava repleto de “erros flagrantes”. Wright (1975), por exemplo, argumenta que o ritmo da expansão econômica no Sul dos Estados Unidos era principalmente determinado pela taxa de crescimento da demanda mundial por algodão e não pelas características da escravidão como um sistema de produção.

Engerman e Sokoloff (1997) colocaram a repressão do trabalho no centro de sua teoria comparativa sobre o desenvolvimento de longo prazo das Américas. Segundo os autores, a presença de condições favoráveis ao cultivo de produtos que apresentavam economias de escala e podiam ser explorados lucrativamente com trabalho escravo, como a cana-de-açúcar e o algodão, resultou em um desenvolvimento econômico inferior na América Latina em comparação com a América do

Norte. Sociedades com essa dotação de fatores tendiam a se estruturar de forma altamente desigual e hierárquica, o que dificultava o crescimento econômico.

Nunn (2007) examina empiricamente a hipótese de Engerman e Sokoloff (1997) utilizando dados em níveis nacional e subnacional. Ele conclui que a escravidão está associada a efeitos negativos sobre o desenvolvimento, embora essa relação não pareça ocorrer por meio da desigualdade. Mais recentemente, Naidu (2012) analisou os efeitos da escravidão no sufrágio e na escolarização, enquanto Bertocchi e Dimico (2014) investigaram seus impactos na educação. Gouda e Rigterink (2013) e Buonanno e Vargas (2019) associaram a escravidão a taxas mais altas de criminalidade, enquanto Acharya et al. (2016, 2018) examinaram seu impacto abrangente na política do Sul.

Economia da Escravidão no Brasil. A escravidão foi uma característica marcante e abrangente do desenvolvimento econômico do Brasil. Mais cativos africanos chegaram às costas brasileiras do que a qualquer outro lugar nas Américas. Entre os séculos XVI e XIX, 4,9 milhões de africanos desembarcaram no Brasil (Eltis, 2018). Esse número corresponde a 46% de todas as chegadas ao Novo Mundo. Assim como no contexto do Sul dos Estados Unidos, a escravidão teve consequências negativas para o desenvolvimento do Brasil: enquanto os proprietários de escravos e alguns setores específicos se beneficiaram, a sociedade como um todo saiu prejudicada (Papadia et al. 2020).

Um dos primeiros historiadores a comparar Brasil e Estados Unidos, tomando as regiões escravistas como referência, destacou que as “características marcantes da estrutura econômica do Sul” eram semelhantes às do Brasil (Graham, 1981). Essas características incluíam um sistema de plantation baseado no trabalho escravo, falta de industrialização, urbanização limitada e baixo investimento em capital humano. Estudos posteriores reforçaram a tese de que a escravidão beneficiou uma pequena elite, mas comprometeu o desenvolvimento econômico das sociedades onde existiu, como ocorreu no Sul dos EUA (Wright, 2020).

Seyler (2021) investiga o impacto de longo prazo da escravidão e sua abolição no desenvolvimento, capital social e atitudes políticas no Brasil. O estudo mostra que a escravidão e o apoio a instituições coercitivas — refletidos nas votações dos legisladores sobre projetos de emancipação no final do século XIX — tiveram efeitos negativos persistentes no PIB, pobreza e desigualdade. Indivíduos de municípios historicamente mais dependentes da escravidão apresentam menor confiança generalizada, maior ceticismo em relação à democracia e crenças mais fracas sobre a corrupção. Lambais e Palma (2023) mostram que, à medida que o tráfico de escravos aumentava na

Bahia, os salários reais dos trabalhadores não qualificados caíam e a desigualdade aumentava. Os salários reais só se recuperaram com o fim progressivo do tráfico de escravos.

Um dos testes mais claros da hipótese de Engerman e Sokoloff (1997) para o Brasil é o estudo de Laudares e Caicedo (2023). Para lidar com a endogeneidade da localização da escravidão, os autores empregam um desenho de Regressão Descontínua Espacial (RDD), explorando as fronteiras coloniais entre os impérios português e espanhol. Usando o Tratado de Tordesilhas (1494) como referência, eles mostram que, em 1872, a presença de escravos era significativamente maior no lado português, refletindo a vantagem comparativa de Portugal no tráfico transatlântico. Essa diferença levou a uma desigualdade de renda moderna 0,04 pontos maior no coeficiente de Gini (cerca de 10% da média nacional). O estudo identifica mecanismos como maior disparidade racial de renda, além de diferenças em educação, emprego e preconceito contra negros, descartando o papel da identidade do colonizador e outras explicações históricas.

O fim da escravidão no Brasil foi resultado de um processo gradual e pacífico, impulsionado por mudanças institucionais, políticas e econômicas. A abolição ocorreu durante a expansão da economia cafeeira, com leis que restringiram progressivamente a escravidão, culminando na Lei Áurea em 1888. Dois principais argumentos econômicos explicam o declínio da escravidão: a transição para formas mais eficientes de trabalho dentro do capitalismo emergente e a divisão entre fazendeiros tradicionalistas do Rio de Janeiro e Vale do Paraíba e aqueles do Oeste paulista, mais abertos à imigração europeia e cujo poder político e econômico crescia à medida que a fronteira agrícola avançava com a expansão das ferrovias.

No entanto, uma interpretação alternativa enfatiza que a escravidão continuava lucrativa na economia cafeeira e que a abolição foi forçada por pressões abolicionistas externas, que minaram gradualmente o suporte social à instituição (Mello, 1978). A imigração europeia surgiu como solução para a demanda de mão de obra, sendo intensificada apenas quando a escravidão se tornou inviável. Assim, o fim da escravidão foi menos uma consequência da modernização econômica e mais um reflexo das expectativas dos fazendeiros diante da crescente pressão política e social contra o regime escravista.

3.10 Imigração, Capital Humano e Crescimento no Brasil

Entre 1872 e 1920, o Brasil recebeu mais de 3,3 milhões de imigrantes, predominantemente da Europa, em um contexto de expansão das plantações de café e da substituição do trabalho

escravo no estado de São Paulo. Qual foi o impacto da imigração europeia no desenvolvimento do Brasil? Este ensaio busca explorar essa questão, oferecendo uma visão geral dos principais resultados da literatura, com foco nos mecanismos por meio dos quais a imigração contribuiu para o desenvolvimento econômico, social e político do país.

O Brasil na Era da Grande Imigração. No final do século XIX, a industrialização e as mudanças demográficas na Europa levaram a fluxos significativos de emigração, com muitos europeus em busca de oportunidades no exterior (Hatton and Williamson, 1998). Entre 1872 e 1920, o Brasil se tornou o quarto destino mais importante para os migrantes europeus, recebendo mais de 3,3 milhões de imigrantes. O estado de São Paulo foi a principal região receptora, absorvendo cerca de 1,8 milhão de imigrantes, um número notável considerando sua população inicial de 837.354 habitantes em 1872.¹ A migração em massa da Europa coincidiu com a ascensão de São Paulo como um dos maiores produtores de café do mundo e a região mais desenvolvida do país (Love, 1980).

A imigração europeia para São Paulo aumentou na segunda metade do século XIX em resposta à rápida expansão das plantações de café e ao maior programa de imigração subsidiada do mundo. O apoio à imigração estrangeira cresceu com a iminente abolição da escravatura em 1888, a fim de expandir rapidamente a força de trabalho enquanto mantinha os custos de mão de obra baixos (Andrews, 1988). O governo implementou políticas para atrair famílias agrícolas sem recursos econômicos, oferecendo incentivos financeiros como subsídios para transporte e concessões de terras. Em 1886, os cafeicultores criaram a Sociedade Promotora da Imigração, privada, mas financiada pelo estado, para gerenciar o programa de imigração.² Sua principal responsabilidade era garantir um fornecimento constante de mão de obra para as fazendas de café, promovendo a imigração europeia, o que incluía a disseminação de propaganda, o custeio das despesas de transporte para famílias agrícolas europeias, o fornecimento de moradia e a alocação de imigrantes para as plantações de café em todo o estado (Holloway, 1980; Lanza et al., 2023).

Imigrantes com despesas de viagem subsidiadas pelo estado eram obrigados a entrar em São Paulo através da Hospedaria dos Imigrantes, na capital. A hospedaria fornecia alimentação, hospedagem, assistência médica e passagens de trem gratuitas para seus destinos finais. Também incluía um escritório de trabalho que ajudava os imigrantes a encontrar emprego nas fazendas

¹Durante esse período, o estado recebeu 53,2% de todos os imigrantes e absorveu aproximadamente 80% em anos específicos.

²Em 1895, a Secretaria de Agricultura, Comércio e Obras Públicas de São Paulo incorporou a Sociedade Promotora da Imigração e assumiu a responsabilidade por todo o programa de subsídios (Holloway, 1980).

de café. Após assinar um contrato, os imigrantes se comprometiam a trabalhar em uma fazenda por pelo menos um ano, antes de poderem revisar os termos de seu contrato ou retornar à hospedaria para buscar novas oportunidades de emprego (Holloway, 1980; Lanza et al., 2023). Para facilitar o transporte dos imigrantes para seus destinos finais no interior do estado, a hospedaria estava conectada a uma estação ferroviária. De 1889 a 1900, 80% dos imigrantes para São Paulo tiveram suas passagens subsidiadas pelo governo estadual. Entre 1889 e 1927, cerca de 2.250.000 estrangeiros entraram em São Paulo, com aproximadamente 54% recebendo subsídios estaduais.

A composição dos imigrantes que chegavam a São Paulo variou ao longo do tempo, especialmente após 1900. Desde a abolição da escravidão até o final do século XIX, os italianos foram o grupo predominante, constituindo 73% de todas as chegadas entre 1887 e 1900.³ Os espanhóis também representaram um grupo significativo, com 11% das chegadas entre 1887 e 1900. A imigração portuguesa seguiu tendências semelhantes à dos espanhóis, com picos notáveis na década de 1890 e entre 1904 e 1914. Ao contrário dos italianos e espanhóis, a migração portuguesa foi mais impulsionada por laços familiares e redes estabelecidas do que por subsídios governamentais. Uma mudança significativa nos padrões de imigração ocorreu no início do século XX, levando a uma população de imigrantes mais diversificada. Preocupações com as condições dos migrantes, especialmente nas fazendas de café, levaram várias nações europeias a proibir a emigração subsidiada para o Brasil, sendo o Decreto Prinetti italiano de 1902 particularmente significativo ao tornar esses subsídios ilegais. De 1901 a 1930, a participação dos italianos caiu para 26%, enquanto a dos espanhóis aumentou para 22%, a dos portugueses para 23% e outras nacionalidades chegaram a 28% (Holloway, 1980, p73).

Os Efeitos da Imigração Europeia no Brasil. A pesquisa sobre os efeitos da imigração no Brasil remonta, pelo menos, ao relato clássico de Dean (1969) sobre a industrialização do estado de São Paulo, que se tornaria o motor manufatureiro do Brasil até 1920. Dean (1969) atribui um papel direto aos migrantes, mas o imigrante crucial nesse relato é o imigrante burguês—relativamente rico e qualificado, e principalmente ativo na indústria e no comércio—em vez dos trabalhadores imigrantes predominantes na agricultura. Outros estudiosos quase unanimemente veem a imigração europeia como um fator positivo no desenvolvimento da economia brasileira, especialmente no setor agrícola. Focando no sudeste do Brasil, e especialmente no estado de São Paulo, essa literatura destacou a importância do trabalho imigrante em três desenvolvimentos interconectados no

³Entre 1888 e 1897, o Brasil atraiu mais emigrantes italianos do que qualquer outro país (Holloway, 1980, p71).

final do século XIX—substituição do trabalho escravo, expansão da fronteira agrícola e produção de café, principal produto de exportação do Brasil (Buciferro 2021; Colistete 2015; Holloway 1980; Milliet 1941; Monbeig 1984; Vidal Luna, Klein e Summerhill 2014).

Um número crescente de pesquisas tem documentado os impactos do capital humano dos imigrantes no desenvolvimento do Brasil. de Carvalho Filho e Colistete (2010) e Rocha et al. (2017) documentam efeitos significativos de longo prazo da imigração europeia sobre o capital humano e os recursos educacionais em São Paulo. Notavelmente, Rocha et al. (2017) mostra que as políticas subsidiadas atraíram imigrantes mais educados para regiões específicas, melhorando a renda per capita a longo prazo. Esses resultados são consistentes com Stolz et al. (2013), que documentam efeitos positivos e persistentes dos imigrantes sobre o estoque de capital humano no Brasil. Complementando essa literatura, Craig e Faria (2021) mostram que a demanda por educação varia com as experiências dos imigrantes com a educação pública e seus antecedentes religiosos, enquanto a oferta de educação depende do capital social e do tamanho da comunidade dos imigrantes. Lopes et al. (2024) fornecem mais evidências sobre os efeitos de diferentes ascendências no desempenho acadêmico ao analisar dados em nível individual. Os autores mostram que alunos com ascendência não ibérica apresentam taxas de promoção mais altas e melhores pontuações em testes padronizados nacionais.

Uma linha diferente de literatura enfatiza não os níveis mais altos de educação dos imigrantes, mas as instituições que eles introduziram (Acemoglu et al., 2014). Witzel de Souza (2018) encontra que a presença de imigrantes alemães em São Paulo não afetou diretamente a acumulação de capital humano; em vez disso, a fundação de escolas por esses imigrantes foi crucial para resultados positivos de longo prazo. Da mesma forma, Colistete (2017) mostra que a chegada de imigrantes intensificou as demandas existentes das famílias brasileiras por educação primária em São Paulo, enquanto de Carvalho Filho e Monasterio (2012) encontram que regiões próximas às colônias alemãs no estado do Rio Grande do Sul têm níveis de desigualdade mais baixos e níveis educacionais mais altos atualmente. Os autores atribuem esses resultados a um sistema de distribuição de terras mais igualitário dentro dos assentamentos patrocinados pelo estado, e não ao capital humano mais elevado dos imigrantes.

Focando no impacto da imigração europeia no setor agrícola de São Paulo, Lanza et al. (2023) encontram que uma maior fração de imigrantes europeus estava associada a um aumento na produção de café por fazenda e a uma maior adoção de ferramentas e maquinário agrícola em 1920, no nível municipal. Em um trabalho relacionado, focado no Brasil, Escamilla-Guerrero et al. (2024)

mostram que uma maior presença de imigrantes levou ao aumento do valor das propriedades agrícolas, o que eles atribuem a um cultivo mais intenso das terras. Apesar das crescentes evidências que ligam o capital humano dos imigrantes do passado aos resultados presentes, particularmente através da acumulação de capital humano, ainda sabemos pouco sobre o impacto da imigração na transmissão de preferências políticas para a população local em países com uma longa história de eleições limitadas e não livres.

Em uma pesquisa em andamento, Viaro, Nakaguma e Pereira (2024) investigam o impacto da imigração para São Paulo na competição eleitoral. Ao analisar a variação entre municípios na imigração, resultante da interação entre os fluxos migratórios e a expansão da rede ferroviária paulista, os autores encontram que a maior presença de imigrantes em 1920 contribuiu para o aumento da votação em candidatos alinhados a Getúlio Vargas nas primeiras eleições presidenciais competitivas do século XX (1955 e 1960). Esses resultados sugerem que a imigração desempenhou um papel no fortalecimento da competição eleitoral no Brasil, intensificando as tensões sociais e econômicas que culminaram no golpe militar de 1964.

3.11 Infraestrutura, Instituições e as Origens da Industrialização no Brasil

Apesar de sua abundância de recursos, o Brasil do século XIX ficou para trás na industrialização em comparação aos países do Atlântico Norte. A estagnação econômica resultou de baixos investimentos e baixa produtividade, sendo as precárias condições de transporte o principal obstáculo ao crescimento antes de 1900 (Summerhill, 2003). Com poucos rios navegáveis fora da Bacia Amazônica, produtos agrícolas e manufaturados enfrentavam altos custos de escoamento. A expansão ferroviária na segunda metade do século reduziu gradualmente esses custos, integrando mercados, promovendo a especialização regional e elevando a produtividade e a renda nacional. Além de sustentar o setor exportador, as ferrovias prepararam o país para um crescimento acelerado após 1900, deixando um impacto duradouro na economia (Summerhill, 2003).

Este ensaio examina o papel da modernização dos transportes, iniciada com a construção das primeiras ferrovias na década de 1850, no desenvolvimento econômico do Brasil. A primeira seção discute o impacto da infraestrutura de transporte no crescimento econômico, enquanto a segunda examina o papel das instituições nesse contexto.

O Papel da Infraestrutura no Desenvolvimento Econômico do Brasil. Até meados do século XIX, a principal transformação econômica do Brasil foi sua abertura ao comércio internacional

(Summerhill, 2003). Nesse período, a economia era altamente dependente das exportações agrícolas — sobretudo café, açúcar e borracha — que constituíam a principal fonte de riqueza do país (Dean, 1971). No entanto, esse modelo exportador tornava o Brasil vulnerável à volatilidade dos mercados internacionais de commodities, dificultando um crescimento sustentável. Além disso, os elevados custos de transporte, agravados pela topografia desafiadora, limitavam a circulação de pessoas e mercadorias. Fora da Bacia Amazônica, a navegação fluvial era restrita, pois a maioria dos rios ao sul do Amazonas possuía trechos não navegáveis. A infraestrutura terrestre também era precária: estradas adequadas para carroças eram escassas, resultando em um sistema de transporte caro e ineficiente, que fragmentava os mercados internos e restringia seu dinamismo (Summerhill, 2003). Além de dificultar o comércio, os altos custos logísticos prejudicavam o crescimento das indústrias emergentes, tanto no acesso a matérias-primas quanto na distribuição de bens manufaturados.

Apesar desses desafios, a modernização da infraestrutura terrestre só ganhou impulso na segunda metade do século XIX. O principal obstáculo à expansão ferroviária no período imperial era a dificuldade de acesso a financiamento (Summerhill, 2003). No entanto, com a crescente importância das exportações de café, tanto o governo quanto a elite agrícola passaram a investir na construção de ferrovias para conectar as regiões produtoras aos portos, reduzindo custos de transporte e aumentando a competitividade do setor. A primeira ferrovia foi inaugurada em 1852 e, nas décadas seguintes, capitalistas nacionais, investidores estrangeiros e o próprio governo financiaram projetos ferroviários. A malha ferroviária, que contava com apenas algumas centenas de quilômetros na década de 1860, expandiu-se para mais de 24 mil quilômetros em 1913 (Summerhill, 2005). Entre 1854 e 1913, a extensão operacional das ferrovias cresceu, em média, quase 11% ao ano, sendo o Sudeste — centro da economia cafeeira e da atividade industrial — a região com maior expansão ferroviária. No entanto, a Primeira Guerra Mundial desacelerou esse crescimento ao impactar o comércio global e restringir o financiamento externo.

A rápida expansão da malha ferroviária entre 1850 e 1913 teve um papel central no desenvolvimento econômico do Brasil, ao integrar mercados regionais e reduzir significativamente os custos de transporte. Summerhill (2003) destaca que, entre 1860 e 1900, os custos de embarque caíram drasticamente, o que teve um impacto profundo em uma economia onde o transporte de longa distância era um dos principais entraves ao comércio. Summerhill (2005) estima que as economias sociais proporcionadas pelas ferrovias correspondiam a pelo menos 18% da renda nacional por volta de 1913. Além de reduzir os custos de transação, as ferrovias diminuíram as disparidades

de preços entre cidades conectadas à rede ferroviária, promovendo maior integração do mercado interno, especialmente em regiões que antes enfrentavam grandes dificuldades logísticas devido à ausência de rios navegáveis e à má qualidade das estradas.

Comparado a outros países, o impacto das ferrovias no Brasil foi expressivo. Fogel e Fishlow estimaram que, nos Estados Unidos, a ausência das ferrovias teria reduzido a renda nacional em apenas 3–4% em 1860 ou 1890, sugerindo um impacto relativamente modesto sobre o crescimento econômico. Já na América Latina, as estimativas de Summerhill (2005) e Herranz-Loncán (2014) indicam que, entre 1910 e 1913, as economias sociais geradas pelas ferrovias correspondiam a cerca de 25% do PIB total na Argentina, no México e no Brasil—um impacto significativamente maior do que nos EUA. Em contraste, no Uruguai, no Peru e na Colômbia, as economias sociais foram inferiores a 5%, reflexo do tamanho limitado de suas redes ferroviárias e da baixa qualidade dos serviços oferecidos (Herranz-Loncán, 2014).

Embora grande parte da literatura tenha se concentrado nos efeitos de curto prazo das melhorias no transporte — como o impacto das ferrovias sobre o valor da terra e a urbanização no século XIX — há razões teóricas para argumentar que seu impacto é relevante no longo prazo. Esse impacto ocorre por meio da persistência das escolhas locacionais: se uma população se estabelece em determinada área devido a alguma vantagem natural ou artificial no passado, essa população pode permanecer nesse local mesmo quando a vantagem original se torna irrelevante.

Barsanetti (2023), por exemplo, mostra que o tempo de permanência como terminal ferroviário no passado é um forte preditor do tamanho das cidades em 2010: municípios que permaneceram por mais tempo como terminais ferroviários apresentaram maior crescimento populacional e do PIB urbano. Esses resultados são consistentes com Américo (2024), que identifica efeitos positivos do acesso ferroviário na transformação estrutural e urbanização do Brasil. Utilizando o caminho de menor custo como variável instrumental e uma análise de event-study, Américo (2024) demonstra que uma redução de 10% na distância de um município até a ferrovia reduz a proporção de trabalhadores na agricultura em 0,6 pontos percentuais, enquanto a participação na indústria aumenta em 0,4 pontos percentuais em 2010. Além disso, ele destaca que as ferrovias foram determinantes para a adoção de novas tecnologias de fiação e tecelagem, impulsionando a modernização do setor. Esses resultados desafiam a visão tradicional da história econômica do Brasil, na qual as ferrovias têm um papel secundário na industrialização (Furtado, 1959; Dean, 1969). No entanto, são consistentes com evidências internacionais, incluindo Argentina (Fajgelbaum e Redding, 2022), EUA (Atack et al., 2021), Suécia (Berger, 2019) e Japão (Tang, 2014).

Além disso, evidências recentes sugerem que mudanças nos custos do comércio interno e internacional ajudaram a moldar a concentração espacial da manufatura no Brasil no período entre guerras. Badia-Miró et al. (2023) mostram que, nas décadas após a Primeira Guerra Mundial, os custos do comércio internacional entre o Brasil e seus principais parceiros aumentaram significativamente, enquanto os custos do comércio interno diminuíram — um fenômeno que os autores denominam reversão dos custos de comércio. O encarecimento do comércio transatlântico enfraqueceu os laços comerciais entre portos brasileiros e estrangeiros, ao passo que as conexões internas entre o núcleo industrial do Sudeste e as regiões do Norte se fortaleceram. Como resultado, a integração dos mercados ocorreu em paralelo à aglomeração da atividade econômica e ao desenvolvimento de uma estrutura centro-periferia centrada em São Paulo.

Instituições e o Desenvolvimento dos Transportes. As evidências apresentadas para o contexto brasileiro sugerem que investimentos em infraestrutura podem mitigar os efeitos da chamada “maldição da geografia”. Exemplos de outros países reforçam esse argumento. Hornbeck e Rotemberg (2021), por exemplo, demonstram que o acesso ferroviário impulsionou significativamente a produção manufatureira nos Estados Unidos, estimando que, sem as ferrovias, a produtividade agregada do país teria sido 25% menor em 1890. No contexto da Revolução Industrial na Grã-Bretanha, Bogart et al. (2017) mostram que a melhoria no acesso aos mercados teve um impacto direto no crescimento populacional. Sem os avanços na infraestrutura de transporte, o crescimento urbano entre 1680 e 1831 teria sido 25 pontos percentuais inferior, e 97 pontos percentuais menor entre 1831 e 1851. Esses números destacam a importância das melhorias no transporte para o desenvolvimento das economias britânica e norte-americana.

No entanto, se investimentos em infraestrutura podem superar barreiras geográficas desfavoráveis, por que alguns países conseguem avançar nesse aspecto enquanto outros não? Um fator essencial é que superar uma geografia adversa exige instituições fortes e eficientes, capazes de alocar recursos de maneira adequada para a construção e manutenção da infraestrutura. A literatura sugere que países com instituições democráticas robustas e maior capacidade estatal tendem a desenvolver redes de transporte mais extensas e eficientes.

Na América Latina, por exemplo, a construção ferroviária frequentemente dependia do apoio financeiro do governo, cujas receitas, por sua vez, eram muitas vezes obtidas por meio de tributos sobre o comércio — um setor que também se beneficiava da expansão das ferrovias. Bignon et al. (2015) argumentam que existia um mecanismo de retroalimentação entre as receitas gover-

namentais e o desenvolvimento ferroviário, criando a possibilidade de múltiplos equilíbrios. Isso significa que alguns países poderiam ficar presos a um “círculo vicioso” de baixa arrecadação e infraestrutura limitada, enquanto outros conseguiriam romper essa barreira e expandir suas redes de transporte.

Bignon et al. (2015) analisam esse fenômeno na América Latina, documentando diferenças significativas entre países. Em nações como Colômbia e Equador, tanto o desenvolvimento ferroviário quanto o crescimento fiscal foram limitados, enquanto Argentina e Uruguai apresentaram trajetórias de expansão mais expressivas. Os autores demonstram empiricamente a relação bidirecional entre ferrovias e arrecadação governamental: o aumento da malha ferroviária impulsiona a receita fiscal, enquanto o crescimento da arrecadação permite maiores investimentos na infraestrutura ferroviária. Historicamente, esse equilíbrio pode ser alterado por choques exógenos que impactam qualquer um dos lados da equação. Por exemplo, uma reforma política que melhore a eficiência legislativa pode aumentar a arrecadação estatal, possibilitando maiores investimentos em infraestrutura.

De forma mais ampla, essa análise ilustra como instituições mais eficazes podem ajudar países a escapar de armadilhas de subdesenvolvimento por meio de melhorias no transporte e na infraestrutura econômica.

4 Conclusões

Nas últimas duas décadas, a pesquisa empírica avançou consideravelmente na análise dos impactos de longo prazo dos eventos históricos no desenvolvimento econômico. As evidências demonstram consistentemente que a história importa, mas ainda não está completamente claro o porque dessa importância. O próximo grande desafio da literatura, portanto, é aprofundar a compreensão dos mecanismos de causalidade. Uma linha promissora de pesquisa é explorar o papel da cultura, um fator capaz de estabelecer normas ou crenças sobre o comportamento, promovendo mudanças institucionais ou sua persistência (Mokyr, 2017; Greif, 2006; Greif e Tabellini, 2010).