

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

MIGRATION INTERNE ET MARCHÉ DU TRAVAIL DES NON-MIGRANTS : UNE
APPROCHE EN PANEL INSTRUMENTÉE POUR L'AFRIQUE DU SUD

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
MAÎTRISE EN ÉCONOMIQUE

PAR
ELIE KENMOGNE SIMO

FÉVRIER 2026

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.12-2023). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

RÉSUMÉ

Ce mémoire analyse l'effet causal de la migration interne sur les résultats du marché du travail des natifs en Afrique du Sud. L'étude s'appuie sur les données longitudinales de l'enquête National Income Dynamics Study (NIDS) couvrant la période 2008–2010/2011 et adopte une approche en longue différence afin d'identifier l'impact des flux migratoires récents.

Pour corriger les biais d'endogénéité liés au tri spatial et à la causalité inverse, la variation du taux de migration rurale–urbaine est instrumentée par un choc climatique exogène construit selon une logique *shift-share*. L'instrument combine des anomalies de température standardisées observées dans les districts d'origine et des parts migratoires historiques, garantissant une variation exogène des flux entrants au niveau des districts de destination.

Les estimations en variables instrumentales indiquent une absence d'effet significatif de la migration rurale–urbaine sur les salaires horaires des natifs. En revanche, les résultats suggèrent un ajustement principalement quantitatif : une augmentation d'un point de pourcentage du taux de migration rurale–urbaine entraîne une hausse modérée mais positive des heures travaillées des natifs. Ces résultats sont cohérents avec un mécanisme d'ajustement par les quantités plutôt que par les prix, dans un contexte sud-africain caractérisé par des rigidités salariales et une segmentation persistante du marché du travail.

Mots-clés : Afrique du Sud, migration interne, migration rurale–urbaine, marché du travail, variables instrumentales, chocs climatiques, identification causale.

TABLE DES MATIÈRES

RÉSUMÉ	i
LISTE DES TABLEAUX	iv
TABLE DES FIGURES	v
LISTE DES ABRÉVIATIONS	vi
REMERCIEMENTS	vii
INTRODUCTION.....	1
CHAPITRE 1 REVUE DE LITTÉRATURE	6
1.1 Fondements théoriques et inégalités territoriales.....	7
1.2 Mécanismes d'impact : salaires, emploi et structure productive	8
1.3 État de la littérature empirique internationale : de la corrélation à l'identification	8
1.4 Spécificités sud-africaines : sélectivité, instruments et lacunes	9
1.5 Approches instrumentales climatiques et cadre d'inférence shift–share	10
1.6 Migration rurale–urbaine et implications	11
1.7 Lacunes de la littérature et positionnement de ce mémoire	11
CHAPITRE 2 PRÉSENTATION DES DONNÉES	13
2.1 Présentation de la base NIDS	14
2.2 Données climatiques et construction de l'instrument de migration	16
2.2.1 Étape 1 : prédiction du taux d'émigration	18
2.2.2 Étape 2 : allocation par destination à l'aide des parts migratoires historiques	19
2.3 Présentation des variables retenues	20
CHAPITRE 3 MÉTHODOLOGIE	27
3.1 Description et justification des variables.....	28
3.2 Spécification du modèle de base.....	29
3.3 Identification et construction de l'instrument	31

3.4	Méthode et spécification économétrique	34
3.4.1	Validité et inférence dans les dispositifs <i>shift–share</i>	36
3.5	Analyse complémentaire	37
CHAPITRE 4 RÉSULTATS EMPIRIQUES		39
4.1	Effets de la migration rurale–urbaine sur les salaires et les heures travaillées	40
4.1.1	Validation de l’instrument pour la migration rurale–urbaine	41
4.1.2	Effets sans contrôles.....	44
4.1.3	Effets avec contrôles.....	46
4.2	Analyse complémentaire : migration interne globale.....	48
4.2.1	Validation de l’instrument pour la migration interne globale	49
4.2.2	Effets sans variables de contrôle.....	52
4.2.3	Effets avec variables de contrôle.....	54
CONCLUSION		58
ANNEXE A CORRÉLATIONS ENTRE LE SALAIRE ET LES VARIABLES		61
ANNEXE B FACTEURS D’INFLATION DE VARIANCE (VIF)		62
ANNEXE C LIEN ENTRE L’ÉMIGRATION PRÉDITE À L’ORIGINE ET L’IMMIGRATION OBSERVÉE À DESTINATION		63
BIBLIOGRAPHIE		64

LISTE DES TABLEAUX

Table 2.1	Statistiques descriptives – niveau individuel et niveau district	23
Table 4.1	Première étape : instrument climatique et variation de la migration rurale–urbaine	42
Table 4.2	Effets de la migration rurale–urbaine sur les salaires et les heures travaillées (sans contrôles)	44
Table 4.3	Effet de la migration rurale–urbaine sur les salaires et les heures travaillées des natifs (avec contrôles)	46
Table 4.4	Première étape : instrument climatique et variation du taux de migration interne .	50
Table 4.5	Effets de la migration interne globale sur les salaires et les heures travaillées (sans contrôles)	53
Table 4.6	Effet de la migration interne globale sur les salaires et les heures travaillées des natifs (avec contrôles)	55
Table A.1	Corrélations de Pearson entre la variation du salaire horaire (log-différence) et les principales variables explicatives	61
Table B.1	Facteurs d’inflation de variance (VIF) pour la régression OLS du salaire.....	62

TABLE DES FIGURES

Figure 0.1	Carte des districts sud-africains.	2
Figure 2.1	Z-scores de température par vague (Vague 1 et Vague 2)	17
Figure C.1	Lien entre l'émigration prédite à l'origine et l'immigration observée à destination.	63

LISTE DES ABRÉVIATIONS

NIDS	National Income Dynamics Study (enquête panel sud-africaine)
SALDRU	Southern Africa Labour and Development Research Unit
LST	Land Surface Temperature (température de surface)
GADM	Global Administrative Areas
MCO/OLS	Moindres Carrés Ordinaires / Ordinary Least Squares
IV	Variabes instrumentales
2SLS	Two-Stage Least Squares (méthode en deux étapes des VI)
VIF	Variance Inflation Factor (facteur d'inflation de variance)
HAC	Heteroskedasticity and Autocorrelation Consistent
PID	Identifiant unique des individus dans NIDS
ZAR	Rand sud-africain (unité monétaire)

REMERCIEMENTS

Je tiens à exprimer ma profonde gratitude envers toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de ce mémoire.

Je souhaite tout d'abord remercier chaleureusement mon directeur de recherche pour la patience et le temps considérable qu'il a bien voulu m'accorder. Je lui suis particulièrement reconnaissant pour l'accompagnement bienveillant qu'il m'a offert tout au long de ce processus, surtout compte tenu du fait qu'il s'agissait de ma première expérience de rédaction académique d'une telle envergure. Ses précieux conseils, sa guidance éclairée et son soutien constant ont été déterminants dans l'aboutissement de ce travail.

Ma profonde gratitude va également à ma fille, qui a été une source inépuisable de motivation et d'inspiration. Sa présence a été un ancrage précieux qui m'a permis de rester concentré et persévérant tout au long de ce parcours exigeant.

Je remercie sincèrement mes camarades pour le soutien émotionnel que nous avons su nous apporter mutuellement. Les échanges stimulants, les conseils partagés et la solidarité dont nous avons fait preuve ont constitué un environnement propice à l'avancement de nos travaux respectifs.

J'exprime aussi ma reconnaissance à la faculté pour l'accompagnement institutionnel, et tout particulièrement à Martine, qui part désormais à la retraite. Son dévouement et son expertise ont été précieux pour de nombreux étudiants dont je fais partie.

Enfin, je tiens à remercier ma famille, mon père, ma mère, et tous ceux qui, de près ou de loin, ont contribué à ce travail. La liste serait trop longue pour tous les citer ici, mais sachez que je garde une pensée reconnaissante pour chacun d'entre vous.

À toutes et à tous, merci pour votre soutien, votre patience et votre encouragement, sans lesquels ce travail n'aurait pu être mené à bien.

INTRODUCTION

Les dynamiques de migration interne représentent aujourd’hui un enjeu central dans de nombreuses économies en développement et émergentes. Elles sont au cœur des transformations sociales et économiques contemporaines, car elles contribuent directement à l’urbanisation rapide et à la recomposition des territoires. Dans de nombreux pays, la mobilité interne détermine l’accès aux opportunités d’emploi, l’adaptation aux inégalités régionales et la résilience face aux chocs économiques et climatiques. Cette ampleur reflète des processus complexes de réallocation de la main-d’œuvre (Harris et Todaro, 1970), de réponses aux inégalités régionales persistantes (Casentini *et al.*, 2024), ainsi que d’adaptation aux chocs économiques et climatiques (Barrios *et al.*, 2006).

Dans plusieurs pays, tels que la Chine, l’Inde ou le Brésil, les mouvements internes ont été fortement influencés par des réformes institutionnelles, des politiques d’urbanisation et l’expansion des infrastructures de transport. En Chine, par exemple, l’assouplissement progressif du système de résidence *hukou* et les réformes économiques entamées à la fin des années 1970 ont déclenché des flux massifs de travailleurs migrants vers les zones urbaines côtières, transformant en profondeur la structure de l’emploi et la géographie économique du pays (Tombe et Zhu, 2019).

Ainsi, l’Afrique du Sud constitue un terrain d’analyse particulièrement pertinent pour étudier la migration interne. L’héritage du système d’apartheid a laissé une géographie profondément marquée par la ségrégation raciale, caractérisée par de fortes disparités territoriales en matière d’emploi, d’infrastructures et d’accès aux services publics. Sous l’apartheid, les politiques de contrôle de la mobilité, notamment les *pass laws*, ont contraint pendant des décennies les populations noires à résider dans des zones spécifiques, marginalisant ainsi leur accès aux centres économiques (Posel, 2006). Ce n’est qu’après l’abolition de ces restrictions, en 1991, que les flux migratoires internes ont pu s’intensifier, redéfinissant les équilibres démographiques et économiques entre districts¹.

1. En Afrique du Sud, un district est une subdivision administrative intermédiaire entre la province et la municipalité locale. Le pays en compte 52, répartis entre 8 municipalités métropolitaines et 44 municipalités de district.

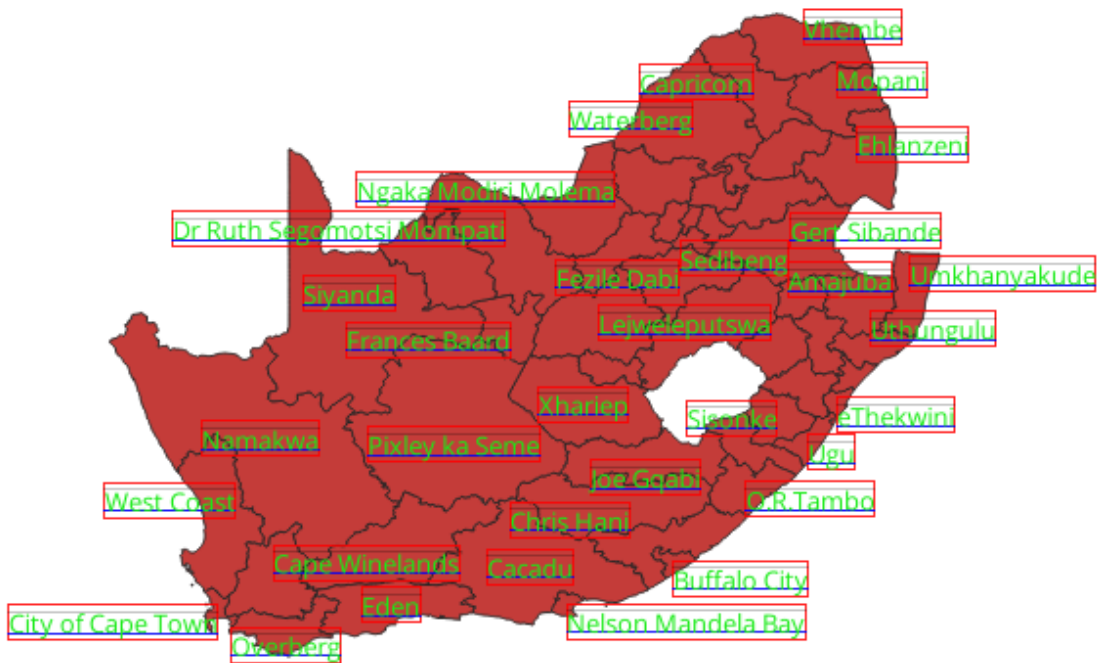


FIGURE 0.1 Carte des districts sud-africains.

Depuis lors, les déplacements inter-districts se sont accélérés (Collinson *et al.*, 2007), portés à la fois par la recherche d'opportunités économiques, les inégalités spatiales persistantes (Noble *et al.*, 2013) et, plus récemment, par des facteurs environnementaux tels que les sécheresses récurrentes et l'instabilité climatique dans certaines régions (Gray et Mueller, 2012). Ces migrations ne relèvent pas uniquement d'une logique d'urbanisation fonctionnelle, mais correspondent aussi à des stratégies de survie ou de mobilité contrainte (Barrios *et al.*, 2006).

Bien que l'attention médiatique se concentre souvent sur les migrations internationales, les flux internes représentent un phénomène plus vaste et aux conséquences plus diffuses. Selon Statistics South Africa (2023), la migration interne en Afrique du Sud est fortement orientée vers les provinces² les plus dynamiques, comme le Gauteng et le Western Cape.

Cette concentration géographique indique que la recherche d'emploi est un déterminant majeur de la mobilité interne. Par ailleurs, les migrants³ sont majoritairement de jeunes adultes âgés de 20 à 34 ans, une tranche d'âge clé sur le marché du travail.

Cependant, les effets économiques de ces mouvements sur les non-migrants à destination restent peu documentés dans la littérature sud-africaine. L'augmentation des flux migratoires internes vers un district peut influencer les salaires et les heures travaillées des résidents locaux, notamment à travers des mécanismes de substitution ou de complémentarité sur le marché du travail.

À partir des données de la National Income Dynamics Study (NIDS), Sharp (2021) analyse les déterminants et les conséquences de la mobilité interne en Afrique du Sud. Pour identifier un effet causal, l'auteur mobilise une stratégie de variables instrumentales (VI) en construisant un instrument de type *shift-share*. Cet instrument s'appuie uniquement sur une variation géographique historique fixe : il exploite les schémas de migration masculine hérités de la période de l'apartheid comme prédicteur exogène des flux contemporains. Si cette approche représente une avancée pour contourner l'endogénéité contemporaine, son pouvoir explicatif repose sur la permanence dans le temps d'un canal migratoire unique et peut être sensible à l'érosion de ce lien historique dans le contexte post-apartheid. De plus, elle ne permet pas de capter la réponse des flux migratoires à des chocs exogènes contemporains, comme les variations climatiques, limitant ainsi sa capacité à isoler l'effet

2. En Afrique du Sud, les provinces constituent le premier niveau de division administrative immédiatement après l'échelon national. Le pays en compte neuf au total.

3. Dans ce mémoire, un migrant est défini comme un individu dont le district de résidence diffère de son district de naissance. Conformément à la littérature empirique standard sur la migration interne, cette définition ne dépend pas de la durée écoulée depuis l'arrivée dans le district de destination. Elle inclut ainsi à la fois les migrations récentes (2008–2011) et les migrations historiques (1931–2006), ces dernières étant utilisées exclusivement comme poids dans la construction de l'instrument. Cette approche est particulièrement pertinente dans le contexte sud-africain, marqué par des migrations internes anciennes et souvent permanentes, même si elle ne permet pas de distinguer entre migrants récents et migrants installés de longue date.

de variations récentes de la migration sur les non-migrants.

De manière similaire, Kollamparambil (2017) étudie les liens entre migration interne et résultats du marché du travail, mais s'appuie sur des stratégies empiriques qui ne permettent pas de lever pleinement les problèmes d'endogénéité liés aux choix de localisation. L'absence d'une stratégie d'identification fondée sur des chocs exogènes limite ainsi la capacité de ces travaux à établir un effet causal strict de la migration interne sur les résultats du marché du travail des natifs.

Dans cette perspective, ce mémoire vise à estimer l'effet causal de la migration interne sur les salaires et les heures travaillées des non-migrants⁴ dans les districts de destination. En mobilisant une stratégie d'instrumentation exogène fondée sur des chocs climatiques (anomalies de température) dans les districts d'origine, pondérés par les parts migratoires historiques de chaque origine vers chaque destination, l'analyse isole l'effet propre des flux migratoires en réduisant les biais d'endogénéité.

Ce travail constitue la première étude appliquée au contexte sud-africain utilisant une stratégie d'instrumentation climatique de type shift-share pour analyser l'impact de la migration rurale-urbaine sur les non-migrants. La méthode, inspirée de Kleemans et Magruder (2018), exploite les données longitudinales des deux premières vagues (2008 pour la vague 1 et octobre 2010 à septembre 2011 pour la vague 2) de l'enquête NIDS (National Income Dynamics Study). L'instrument climatique repose sur des chocs de température calculés pour l'année précédant chacun des épisodes de migration récents observés dans les vagues (2007 pour la vague 1 et octobre 2008 à septembre 2009 pour la vague 2), de manière à garantir que le choc est strictement antérieur aux décisions de migration observées lors des enquêtes. L'étude examine également, de manière complémentaire à la migration rurale-urbaine, une analyse de l'impact de la migration interne globale. Cette extension permet de tester si les migrants pris globalement de façon agrégée (rurale-urbaine, urbaine-urbaine et rurale-rurale) exercent une pression particulière sur le marché du travail, comme le suggèrent Bryceson (2000) et Beauchemin (2011).

Par rapport à Posel et Casale (2003), qui se concentre sur les déterminants de la migration féminine dans une approche descriptive, cette recherche se démarque par une volonté claire d'identification causale. Elle diffère aussi de Sharp (2021), qui utilise des parts migratoires héritées de l'apartheid

4. Un individu est considéré comme non-migrant s'il réside dans son district de naissance.

dans une logique historique, en recourant à des chocs climatiques exogènes pour renforcer la validité de l'instrument. Enfin, contrairement à Kollamparambil (2017), dont les instruments GMM s'appuient sur des valeurs retardées potentiellement endogènes, cette étude repose sur des variations climatiques et géographiques plausiblement exogènes.

L'hypothèse de départ est que les flux migratoires peuvent exercer une pression sur les marchés du travail locaux, affectant soit les salaires, soit la quantité de travail offerte par les natifs. Dans le contexte sud-africain, où les inégalités structurelles et les rigidités sont importantes, il est crucial de tester empiriquement l'ampleur réelle de ces effets. Les résultats attendus anticipent un impact sur le volume horaire (méthode IV) sans effet sur les salaires, confirmant ainsi l'hypothèse d'une absence d'influence sur les indicateurs de l'emploi.

Ce mémoire est structuré en quatre chapitres. Le premier propose une revue de la littérature. Le deuxième décrit les données et la construction des variables. Le troisième détaille la méthodologie économétrique. Le quatrième présente les résultats empiriques. Enfin, une conclusion discute les implications et formule des recommandations pour une meilleure intégration des migrants internes sur le marché du travail sud-africain.

CHAPITRE 1

REVUE DE LITTÉRATURE

L'étude de l'impact de la migration interne sur les résultats du marché du travail est marquée par un débat central autour de l'identification causale. Les corrélations observées entre flux migratoires, salaires et emploi ne permettent pas, en elles-mêmes, d'établir un lien de causalité, en raison de plusieurs sources classiques d'endogénéité. Premièrement, une causalité inverse peut être à l'œuvre : les districts connaissant une expansion de l'activité économique, une création d'emplois ou une hausse des salaires attirent davantage de migrants, de sorte que la migration apparaît comme une conséquence, et non comme une cause, de l'amélioration des conditions du marché du travail. Deuxièmement, les migrants ne se répartissent pas aléatoirement dans l'espace, mais s'auto-sélectionnent (*sorting*) vers des destinations offrant de meilleures opportunités économiques, une structure productive plus dynamique ou un meilleur accès aux infrastructures, ce qui biaise les estimations naïves. Troisièmement, des variables omises, telles que les investissements publics, les politiques locales de développement, les chocs sectoriels ou les dynamiques industrielles régionales, peuvent affecter simultanément la migration interne et les résultats du marché du travail, sans être observées.

Par ailleurs, il est essentiel de distinguer clairement les mécanismes économiques sous-jacents. Les décisions de migration sont principalement motivées par les conditions de *demande de travail* dans les districts de destination, reflétant la disponibilité d'emplois et les perspectives économiques locales. En revanche, l'arrivée de migrants peut ensuite influencer les résultats du marché du travail des résidents en place en modifiant l'*offre de travail* locale, notamment à travers des effets de concurrence ou de complémentarité entre travailleurs.

Ce chapitre poursuit ainsi un double objectif : proposer un cadrage théorique rigoureux des mécanismes reliant migration interne et marché du travail, et présenter un état de la littérature internationale et sud-africaine afin de situer précisément la contribution de ce mémoire. La logique adoptée est de progresser d'un cadre théorique vers les études descriptives, puis vers les approches d'identification causale, en montrant pourquoi une stratégie instrumentale climatique de type *shift-share* constitue un outil pertinent pour corriger les biais d'endogénéité et analyser les effets de la

migration interne sur les non-migrants dans le contexte sud-africain.

1.1 Fondements théoriques et inégalités territoriales

Les fondements analytiques de la migration interne reposent sur un arbitrage revenu-coût, où l'individu compare les revenus espérés à destination et à l'origine, en tenant compte des probabilités d'emploi et des frictions de mobilité. Le modèle de Harris-Todaro formalise cet arbitrage et explique la persistance des migrations vers les centres urbains, même en présence d'un chômage apparemment élevé, dès lors que le revenu espéré urbain excède le revenu rural certain (Harris et Todaro, 1970; Todaro, 1969). Les extensions de ce cadre de base intègrent la dualité formel-informel, les files d'attente pour l'emploi urbain et le rôle des réseaux sociaux qui réduisent les coûts de recherche et d'installation (Bhattacharya, 1993; Munshi, 2003; Lucas, 2005). Ces éléments influencent à la fois le profil des migrants (plus jeunes, plus instruits, disposant de capital social) et l'intensité des flux.

Au niveau agrégé, les effets sur le marché du travail local opèrent par les *prix* (salaires) et les *quantités* (heures travaillées), et dépendent de la substituabilité ou de la complémentarité des compétences entre migrants et natifs, de la vitesse d'ajustement du capital et de la structure sectorielle (Borjas, 2003; Card, 2001; Ottaviano et Peri, 2012; Basso et Peri, 2015; Dustmann *et al.*, 2012). Une forte substituabilité tend à accroître la concurrence sur des segments de qualification similaires, tandis qu'une complémentarité, par exemple entre tâches manuelles et cognitives ou entre niveaux d'expérience, peut soutenir la productivité et amortir les pressions sur les salaires. Dans des marchés segmentés, la concentration des migrants dans certains secteurs ou niches peut limiter la concurrence directe, tout en reconfigurant la spécialisation locale et la productivité globale (Card, 2001; Dustmann *et al.*, 2012). Le rythme d'ajustement du capital est crucial : s'il est lent, les effets à court terme sur les salaires peuvent différer des effets de moyen terme, une expansion de la demande et de l'investissement pouvant ultérieurement neutraliser ou inverser des baisses initiales.

Dans les pays en développement, et singulièrement en Afrique du Sud, la migration interne s'inscrit dans un paysage d'inégalités territoriales persistantes. Outre les différentiels de salaires, l'accès aux services publics (éducation, santé, transport) et les infrastructures urbaines jouent un rôle décisif dans l'attractivité des destinations et dans la capacité d'absorption des flux (Casentini *et al.*, 2024). L'héritage de l'apartheid a institutionnalisé une ségrégation spatiale qui continue de produire de forts

écarts d’opportunités et de capital humain entre districts (Noble *et al.*, 2013). Dans ce contexte, les chocs exogènes, économiques ou climatiques peuvent agir comme déclencheurs de réallocations géographiques de la main-d’œuvre, indépendamment des conditions conjoncturelles spécifiques aux destinations.

1.2 Mécanismes d’impact : salaires, emploi et structure productive

La littérature distingue des canaux *prix* et *quantités*, mais leurs effets empiriques apparaissent souvent modestes une fois l’endogénéité correctement traitée. Les approches par cellules compétences–expérience à l’échelle nationale mettent en avant des effets hétérogènes selon la substituabilité des groupes (Borjas, 2003), tandis que des travaux prenant explicitement en compte les complémentarités (et l’ajustement du capital) documentent des effets nuls à positifs sur les salaires des natifs, particulièrement pour les travailleurs complémentaires (Ottaviano et Peri, 2012; Basso et Peri, 2015). L’ajustement sectoriel et l’expansion de la demande locale peuvent, à moyen terme, absorber l’augmentation de l’offre de travail sans détérioration des salaires moyens (Card, 2001; Dustmann *et al.*, 2012). En parallèle, les heures travaillées réagissent parfois faiblement aux chocs d’offre dans des contextes où les contraintes d’horaires, la réglementation ou l’informalité encadrent la quantité de travail (Dustmann *et al.*, 2012).

Dans le cas des migrations rurales–urbaines, la pression sur les infrastructures, l’urbanisation rapide et la segmentation entre formel et informel constituent des déterminants majeurs de l’impact économique. L’arbitrage des ménages porte autant sur les perspectives de revenus que sur l’accès aux biens publics et aux réseaux, ce qui peut générer des effets différenciés selon le type de flux et le stade d’urbanisation (Tacoli, 2006). Ces considérations motivent une désagrégation empirique par type de migration, particulièrement pertinente dans le contexte sud-africain.

1.3 État de la littérature empirique internationale : de la corrélation à l’identification

Les premières études régionales fondées sur des chocs migratoires exploitent la distribution géographique passée des migrants afin de prédire les arrivées contemporaines, inaugurant ainsi la logique dite *shift–share* (Card, 2001). Cette approche repose sur l’inertie des réseaux migratoires et sur l’auto-renforcement des destinations historiques. D’autres travaux, cherchant à limiter l’endogénéité liée aux choix de localisation, agrègent les données à l’échelle nationale par cellules de qualification

et d'expérience (Borjas, 2003). Dans l'ensemble, cette littérature montre que les estimations naïves tendent à montrer des effets positifs ou nuls de la migration, tandis que des stratégies d'identification plus rigoureuses mettent en évidence des impacts faibles, nuls ou parfois négatifs, selon la structure des complémentarités entre migrants et natifs (Ottaviano et Peri, 2012; Basso et Peri, 2015; Dustmann *et al.*, 2012).

Une contribution méthodologique majeure est proposée par Boustan (2010), qui développe explicitement une approche de type *shift-share* fondée sur des chocs exogènes affectant les régions d'origine des migrants. Dans son cadre, ces chocs sont combinés à des parts migratoires historiques afin de prédire les flux migratoires vers les régions de destination. Cette stratégie permet d'isoler une composante de la migration indépendante des conditions économiques contemporaines à destination, répondant ainsi aux problèmes de causalité inverse et de sélection spatiale qui affectent les estimations corrélationnelles. Des travaux ultérieurs mobilisent des chocs institutionnels ou environnementaux de grande ampleur selon une logique comparable (Naidu et Hornbeck, 2014; Henderson *et al.*, 2017).

Le dispositif empirique retenu dans ce mémoire s'inscrit directement dans cette lignée. Il repose sur la combinaison d'un choc climatique exogène mesuré dans les districts d'origine et de parts migratoires historiques reliant origines et destinations, afin de construire un instrument de type *shift-share*. Cette approche vise à générer une variation de la migration interne corrélée aux flux migratoires, tout en étant orthogonale aux chocs économiques locaux dans les districts de destination, condition essentielle à l'identification causale des effets de la migration sur les résultats du marché du travail des non-migrants.

1.4 Spécificités sud-africaines : sélectivité, instruments et lacunes

En Afrique du Sud, les déterminants microéconomiques de la migration interne incluent l'âge, le sexe, le niveau d'éducation et les réseaux familiaux (Posel, 2006; Bouare, 2000). Les flux circulaires entre zones rurales et urbaines sont anciens et s'inscrivent dans une géographie héritée de l'apartheid, avec des différentiels persistants d'accès aux services et aux opportunités. Sur le plan causal, les travaux mesurant l'effet de la migration sur les salaires ou l'emploi des résidents non-migrants restent rares. Certaines études recourent à des instruments fondés sur des parts migratoires passées,

mais leur validité est discutée lorsque ces parts capturent des structures économiques persistantes potentiellement corrélées aux résultats contemporains (Sharp, 2021). D’autres contributions mobilisent des variables retardées dans un cadre GMM, sans garantir la stricte exogénéité vis-à-vis des chocs locaux affectant simultanément la migration et le marché du travail (Kollamparambil, 2017).

Au niveau régional africain, des chocs climatiques ont été utilisés pour expliquer des réallocations de population et des dynamiques d’urbanisation, suggérant que les aléas environnementaux peuvent fournir une source d’exogénéité exploitable pour l’identification (Barrios *et al.*, 2006; Henderson *et al.*, 2017). Dans le contexte sud-africain, où les districts les plus exposés sont souvent les plus pauvres, ces chocs constituent des candidats naturels pour une stratégie instrumentale fondée sur l’origine des flux. La littérature récente discute aussi des migrations dites forcées par des aléas, relativement indépendantes des conditions de demande dans les zones d’accueil (Nyoni et Kollamparambil, 2022). Toutefois, la combinaison explicite de chocs climatiques mesurés à l’origine et de parts migratoires historiques vers les destinations, dans l’esprit des instruments *shift-share*, reste peu exploitée pour estimer l’impact *causal* sur les non-migrants en Afrique du Sud.

1.5 Approches instrumentales climatiques et cadre d’inférence *shift-share*

Ce mémoire s’inscrit dans la tradition des instruments *shift-share*, en combinant un choc climatique mesuré à l’origine et des parts migratoires historiques reliant origines et destinations. Le choc est construit comme une anomalie standardisée (z-score) par district d’origine, à partir d’une moyenne et d’un écart-type de long terme, et mesuré antérieurement aux mouvements migratoires afin d’assurer l’exogénéité temporelle. La pondération par les parts historiques projette ensuite ce choc sur les destinations et génère une composante « push » des arrivées, conformément à la logique inaugurée par Card (2001) et aux pratiques de désagrégation par compétences et territoires (Basso et Peri, 2015). L’objectif est de maximiser la pertinence de l’instrument tout en respectant la condition d’exclusion, en minimisant tout canal direct entre les chocs à l’origine et les résultats du marché du travail à destination.

Sur le plan de l’inférence, les dispositifs *shift-share* exigent des précautions lorsque les parts historiques créent des dépendances croisées : des erreurs standards usuelles peuvent sous-estimer l’incertitude. Des contributions récentes proposent des conditions de dispersion OU granularité de

l'exposition et des correctifs d'inférence adaptés (Adao *et al.*, 2018; Jaeger *et al.*, 2018). Dans la pratique, il convient de documenter la distribution des poids, d'examiner la concentration de l'exposition et d'adopter des corrections de variance appropriées. La validité empirique repose ainsi sur la pertinence (les chocs climatiques d'origine affectent l'émigration (Barrios *et al.*, 2006)), la condition d'exclusion et une inférence robuste. Les choix opérationnels (différences longues, tests de première étape de Kleibergen–Paap, tests d'Anderson–Rubin, variantes d'agrégation des poids) sont détaillés au chapitre 3.

1.6 Migration rurale–urbaine et implications

La distinction entre migrations rurales–urbaines et autres formes de mobilité est essentielle pour comprendre ses effets. Les flux ruraux–urbains s'accompagnent souvent d'une pression accrue sur les infrastructures urbaines, d'une concurrence dans les segments informels et d'un accès inégal aux services de base (Tacoli, 2006). Cette configuration peut moduler la substituabilité des tâches et la vitesse d'ajustement du capital, rendant plausible un profil d'effets différent de celui observé dans des contextes d'immigration internationale. Dans ce mémoire, la désagrégation par type de flux répond à la nécessité d'isoler ce canal et de produire des résultats interprétables pour les politiques publiques d'urbanisation et d'investissement dans les infrastructures.

1.7 Lacunes de la littérature et positionnement de ce mémoire

Malgré un corpus fourni sur les déterminants individuels de la migration interne en Afrique du Sud, les estimations causales de l'effet sur les *non-migrants* restent limitées. Les approches descriptives ou corrélatives n'isolent pas l'exogénéité nécessaire pour conclure sur les effets de marché ; les instruments fondés sur des parts historiques non couplées à des chocs exogènes, ou sur des variables retardées endogènes, soulèvent des interrogations en matière d'identification (Sharp, 2021; Kollamparambil, 2017). Ce mémoire apporte une contribution spécifique en adaptant au cas sud-africain une stratégie instrumentale *climatique shift-share*, dans laquelle un choc climatique mesuré à l'origine est projeté sur les destinations via des parts migratoires historiques pour identifier l'effet de la migration sur les salaires et l'emploi des *non-migrants*. Cette combinaison, migration interne, instrument climatique d'origine, pondération par parts historiques origine–destination, identification des effets sur les non-migrants, reste à notre connaissance peu explorée dans la littérature appliquée à l'Afrique du Sud, et s'inscrit dans la continuité méthodologique des travaux internationaux (Card,

2001; Basso et Peri, 2015; Adao *et al.*, 2018; Jaeger *et al.*, 2018).

L’articulation proposée entre cadre théorique, spécificités sud-africaines et inférence *shift–share* prépare la présentation des données (sources, construction des variables climatiques et migratoires, structure du panel) et les spécifications économétriques en deux étapes détaillées au chapitre 3. Les résultats empiriques, rapportés au chapitre 4, mettent à l’épreuve les mécanismes discutés ici, en distinguant les effets sur les salaires et les heures travaillées des natifs et en documentant l’hétérogénéité selon le type de flux. L’ensemble vise à éclairer la question de savoir si, et dans quelles conditions, la migration interne induite par des chocs d’origine reconfigure le marché du travail local sans détériorer les opportunités des résidents en place, ou au contraire produit des pressions mesurables sur les prix et les quantités de travail.

Ce chapitre a montré que les effets de la migration interne dépendent d’une combinatoire de mécanismes micro et macro, et que les estimations causales requièrent une source d’exogénéité crédible. Le contexte sud-africain, marqué par des inégalités territoriales héritées et des flux ruraux–urbains anciens, se prête à une identification instrumentale où des chocs climatiques mesurés à l’origine, projetés par des parts historiques vers les destinations, génèrent une variation pertinente de la migration. Les précautions d’inférence propres aux dispositifs *shift–share* sont intégrées à notre stratégie empirique, de manière à fournir des estimations robustes des effets sur les non-migrants.

CHAPITRE 2

PRÉSENTATION DES DONNÉES

Ce chapitre présente les sources de données mobilisées, les traitements appliqués et les principales variables construites pour évaluer l'effet de la migration interne sur les salaires et les heures travaillées des natifs sud-africains. L'analyse repose sur deux sources principales. D'une part, l'enquête National Income Dynamics Study (NIDS), accessible via la Microdata Library de la Banque mondiale, fournit des données individuelles longitudinales permettant de suivre les trajectoires migratoires et professionnelles d'un échantillon représentatif d'individus sud-africains. D'autre part, des données climatiques satellitaires mensuelles sur les températures, extraites de la plateforme NASA Giovanni¹, sont utilisées pour capter des chocs exogènes affectant les zones d'origine des migrants.

Comme le soulignent Hu (2023), les données satellitaires de température de surface (LST – Land Surface Temperature) sont particulièrement pertinentes pour caractériser les épisodes de chaleur extrême. Fondées sur des observations spatiales régulières, elles offrent une haute résolution spatiale et permettent de détecter des anomalies thermiques locales, même dans les zones dépourvues de stations météorologiques au sol. Les satellites géostationnaires, grâce à leur capacité à fournir des relevés thermiques mensuels, permettent ainsi d'identifier de manière fiable les chocs climatiques extrêmes qui affectent les conditions de vie locales (Williams et Others, 2024).

Ces deux sources sont combinées avec des fichiers cartographiques géoréférencés du projet Global Administrative Areas (GADM)², permettant une agrégation cohérente au niveau des 52 districts sud-africains. L'intégration de ces données permet la construction d'un instrument exogène de la variation du taux de migration basé sur les chocs climatiques, suivant une logique shift-share : les anomalies thermiques dans les zones d'origine sont pondérées par les parts migratoires historiques afin d'estimer un flux migratoire instrumenté vers chaque district de destination.

Cette stratégie vise à identifier de manière causale l'effet de la migration interne sur les salaires et les heures travaillées des natifs, tout en surmontant les biais d'endogénéité liés aux choix migratoires

1. <https://giovanni.gsfc.nasa.gov/giovanni/>

2. <https://gadm.org/data.html>

individuels.

2.1 Présentation de la base NIDS

L'enquête National Income Dynamics Study (NIDS) constitue à ce jour la principale source de données longitudinales représentative à l'échelle nationale pour étudier les dynamiques de revenu, d'emploi et de migration en Afrique du Sud. Mise en œuvre par le Southern Africa Labour and Development Research Unit (SALDRU) à l'Université du Cap, elle suit depuis 2008 un échantillon initial d'environ 28 000 individus répartis dans 7 296 ménages. L'échantillon couvre l'ensemble des 9 provinces sud-africaines et les 52 districts administratifs, avec une stratification explicite selon le milieu de résidence (rural et urbain), ce qui garantit la représentativité au niveau national, provincial et par type de lieu de résidence.

La distinction entre zones rurales et urbaines repose sur la classification officielle de Statistics South Africa, telle qu'elle est intégrée dans l'enquête NIDS. Cette classification est définie au niveau du lieu de résidence des individus, à partir de critères géospatiaux et fonctionnels (densité, type d'habitat, infrastructures et activités économiques), et non sur la base de seuils démographiques agrégés. Dans ce mémoire, les flux migratoires sont qualifiés de ruraux–urbains lorsque les individus migrent d'une zone classée comme rurale vers une zone classée comme urbaine selon cette nomenclature, et inversement pour les flux urbains–ruraux.

Les données sont recueillies à travers une structure de panel longitudinal³, qui permet de suivre les mêmes individus d'une vague à l'autre grâce à un identifiant unique `pid`. Cette architecture rend possible l'étude des trajectoires individuelles : changements de résidence (migrations inter-districts), différences salariales ou d'heures travaillées, niveau d'éducation ou encore transformations dans la composition des ménages.

La base est organisée en plusieurs fichiers complémentaires. Les fichiers principaux comprennent les modules individuels pour adultes (âgés de 15 ans ou plus), les informations sur les enfants, les caractéristiques des ménages, ainsi que les fichiers de liaison permettant de relier les observations entre les différentes vagues. Ce mémoire exploite exclusivement les données des adultes âgés de 15

3. Un panel longitudinal suit les mêmes individus à travers plusieurs périodes, permettant d'observer l'évolution individuelle dans le temps.

à 64 ans de la vague 1 et 2, population considérée comme en âge de travailler, pour un échantillon de 14 572 observations.

Un individu est défini comme migrant s’il réside, au moment de l’enquête, dans un district différent de celui de sa naissance, conformément à la définition retenue dans Posel (2006) pour une approche similaire dans le contexte sud-africain. Cette définition permet de distinguer les migrants internes des non-migrants à chaque vague. Deux catégories sont considérées. Les migrants historiques regroupent les individus installés dans un district de destination différent de leur district de naissance entre 1931 et 2006. Ces flux servent uniquement à la construction des parts migratoires historiques utilisées dans la stratégie d’instrumentation *shift-share*. Les migrants récents⁴, quant à eux, varient selon la vague : à la vague 1, ils correspondent aux individus s’étant installés dans un district de destination différent de leur district de naissance entre janvier 2008 et décembre 2008, tandis qu’à la vague 2, il s’agit des individus ayant migré entre janvier 2010 et septembre 2011. Cette distinction temporelle et spatiale assure la cohérence dans la construction des instruments et l’identification des effets propres de la migration sur les résultats du marché du travail.

L’analyse économétrique repose sur un échantillon d’individus présents à la fois à la vague 1 et à la vague 2 de l’enquête NIDS, condition nécessaire au calcul des longues différences des variables individuelles. Les régressions sont estimées sur les individus natifs.

Les migrants récents ne constituent pas des unités d’observation dans les estimations de seconde étape. Ils sont utilisés exclusivement pour mesurer les flux migratoires entrants au niveau des districts, lesquels servent à construire les variables explicatives de migration et l’instrument climatique de type *shift-share*.

Grâce à sa richesse et à sa structure longitudinale, la NIDS permet de construire des variables exprimées en log-différences, comme les salaires et les heures travaillées, ainsi que la construction de la différence des taux de migration entre les deux vagues. Elle permet également d’identifier les épisodes de mobilité résidentielle, de suivre les effets de la migration sur les trajectoires professionnelles et de contrôler l’hétérogénéité inobservable individuelle.

4. De 2008 à septembre 2011, en excluant 2009 pour éviter le chevauchement avec les périodes d’épisodes climatiques susceptibles d’affecter la migration observée lors de la vague 2.

La NIDS présente toutefois certaines limites méthodologiques. Les salaires et les heures travaillées sont autodéclarés, ce qui peut induire un biais de mesure. L’attrition entre les vagues, bien que relativement modérée (environ 17% entre les vagues 1 et 2), peut affecter la représentativité longitudinale.

Malgré ces limites, la NIDS reste une base incontournable pour l’étude des mobilités internes sud-africaines. Posel (2006) explore les déterminants de la mobilité résidentielle, tandis que Kollamparambil (2017) analyse l’impact de la migration sur l’emploi, l’informalisation et les inégalités de genre. Ces études, bien qu’informatives, ne mobilisent pas pleinement la richesse du panel pour estimer des effets causaux, ce que ce mémoire se propose précisément de faire.

2.2 Données climatiques et construction de l’instrument de migration

En complément des données individuelles de l’enquête NIDS, ce mémoire mobilise des données climatiques satellitaires issues de la plateforme NASA Giovanni⁵. Les données de température de surface (T2M) sont utilisées afin de mesurer des anomalies climatiques persistantes susceptibles d’affecter durablement les conditions de vie, les activités économiques locales et, par conséquent, les décisions de migration interne.

Les données sont disponibles au format NetCDF, à une fréquence mensuelle, avec une résolution spatiale de 0,5 degré. Elles couvrent la période 1996–2011, ce qui permet de construire des indicateurs de température antérieurs aux vagues de l’enquête NIDS. Chaque mois, une grille raster de température moyenne est générée pour l’ensemble du territoire sud-africain.

Ces rasters sont croisés avec les limites administratives au niveau du district, issues du projet Global Administrative Areas (GADM)⁶. À l’aide d’un script automatisé dans PyQGIS, les températures mensuelles sont agrégées à l’échelle des 52 districts sud-africains. Le résultat est une base de données climatiques structurée au niveau district-mois, mesurée en degrés Celsius.

À partir de cette série temporelle, un indicateur de choc climatique est construit sous la forme d’un

5. <https://giovanni.gsfc.nasa.gov/giovanni/>

6. <https://gadm.org/data.html>

z-score standardisé. Pour chaque district o et précédant chaque vague t (octobre 2008 à septembre 2009 pour la vague 2 et janvier 2007 à décembre 2007 pour la vague 1) de l'enquête, le choc thermique est défini comme suit :

$$\text{Choc}_o = \frac{\bar{T}_o^{\text{court}} - \bar{T}_o^{\text{long}}}{\sigma_o^{\text{long}}} \quad (2.1)$$

où \bar{T}_o^{court} est la moyenne des températures sur les 12 mois précédant la vague 1 et 12 mois précédant la vague 2, \bar{T}_o^{long} est la moyenne de long terme (calculée entre 1996 et 2011) pour le district o , et σ_o^{long} est l'écart-type correspondant. Cette méthode, inspirée de Maccini et Yang (2009), permet d'identifier des anomalies locales, positives ou négatives, reflétant un écart inhabituel par rapport aux normes climatiques historiques.

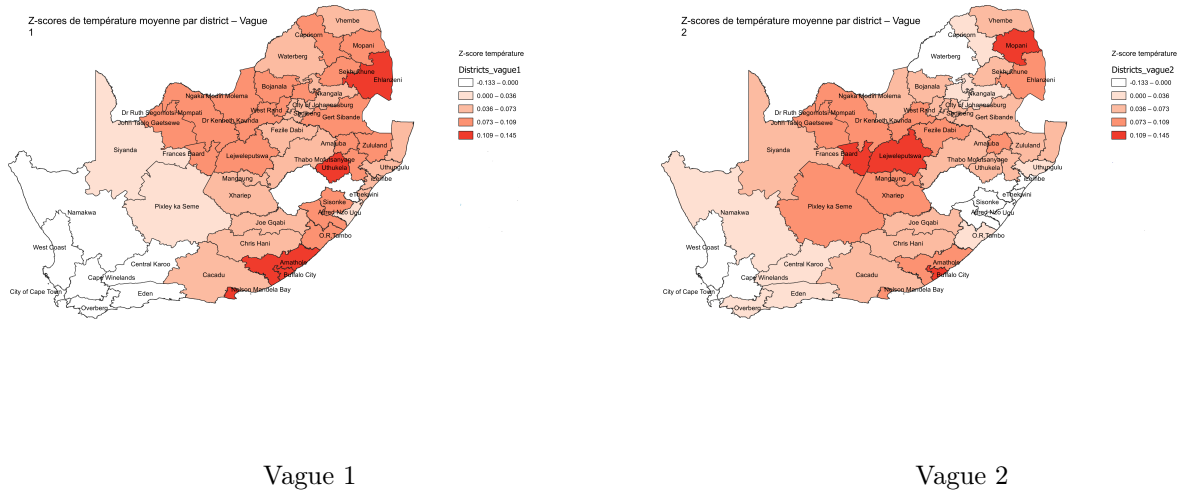


FIGURE 2.1 Z-scores de température par vague (Vague 1 et Vague 2)

Les deux cartes comparatives mettent en évidence une intensification des anomalies de chaleur entre la vague 1 et la vague 2.

Ces chocs climatiques sont ensuite utilisés pour construire un instrument exogène de migration interne, en suivant une logique shift-share. L'hypothèse sous-jacente est que des températures anormalement élevées ou basses dans les zones d'origine peuvent détériorer les conditions de vie (ressources agricoles, habitat, santé), poussant les individus à migrer vers d'autres districts (Barrios *et al.*, 2006; Kelley *et al.*, 2015).

La mesure et le contrôle de la situation économique initiale à destination rendent les chocs comparables entre districts et garantissent leur caractère exogène vis-à-vis des conditions économiques contemporaines des destinations (Jaeger *et al.*, 2018).

La construction de l'instrument repose sur une stratégie en deux étapes. Dans un premier temps, les chocs climatiques sont utilisés pour prédire le taux d'émigration sortant de chaque district d'origine. Dans un second temps, ces taux prédits sont pondérés par des parts migratoires historiques vers chaque destination, de manière à obtenir une mesure instrumentale de la variation du taux d'immigration reçue à destination.

2.2.1 Étape 1 : prédiction du taux d'émigration

Pour chaque district d'origine o , on estime la relation entre les anomalies climatiques précédant les vagues et la variation du taux d'émigration sortante entre les deux vagues (janvier 2008 à décembre 2008 pour la vague 1 et mai 2010 à septembre 2011 pour la vague 2). L'équation de cette étape s'écrit :

$$\Delta\text{EmigRate}_o = \pi_0 + \pi_1 \text{Choc}_o + u_o \quad (2.2)$$

où $\Delta\text{EmigRate}_o$ désigne la variation du taux moyen d'émigration sortante depuis le district d'origine o entre la vague 1 et la vague 2. Cette variable mesure l'évolution de la propension des résidents du district o à quitter leur zone d'origine au cours de la période considérée et constitue la variable dépendante d'une équation de prédiction, et non d'une équation structurelle d'intérêt.

Le terme π_0 correspond à une constante, capturant le niveau moyen de variation du taux d'émigration commun à l'ensemble des districts, indépendamment des conditions climatiques. Il absorbe les facteurs agrégés invariants ou communs à tous les districts sur la période étudiée.

Le choc climatique Choc_o correspond à un choc exogène au niveau du district d'origine, défini comme une anomalie thermique standardisée (z-score). Il est calculé comme l'écart entre la température moyenne de court terme observée avant chaque vague et la moyenne de long terme propre au

district o , normalisé par l'écart-type historique des températures du district. Cette construction permet d'identifier des déviations climatiques exceptionnelles relativement aux conditions climatiques habituelles locales.

Le coefficient π_1 mesure la sensibilité du taux d'émigration sortante aux variations exogènes des conditions climatiques dans le district d'origine. Ce paramètre n'est pas interprété comme un effet causal final de la température sur la migration, mais comme un coefficient de prédiction permettant d'identifier la composante exogène des flux migratoires sortants, utilisée ultérieurement dans la construction de l'instrument de migration interne de type *shift-share*.

Le terme u_o désigne un terme d'erreur idiosyncratique, capturant l'ensemble des facteurs non observés propres au district d'origine susceptibles d'affecter la variation du taux d'émigration et supposés orthogonaux au choc climatique compte tenu de la construction exogène de celui-ci.

À partir de cette estimation, on obtient la composante prédite de la variation du taux d'émigration, notée $\widehat{\Delta\text{EmigRate}}_o$. Afin d'exprimer la migration en nombre d'individus, cette variation prédite est multipliée par la population totale du district d'origine, ce qui permet d'obtenir le nombre de migrants sortants prédits :

$$\widehat{\text{Migrants}}_o = \widehat{\Delta\text{EmigRate}}_o \times \text{Population}_o \quad (2.3)$$

2.2.2 Étape 2 : allocation par destination à l'aide des parts migratoires historiques

Les migrants sortants prédits sont ensuite répartis entre les districts de destination à l'aide de parts migratoires historiques w_{od} , qui mesurent la proportion des migrants originaires du district o s'étant installés historiquement dans le district de destination d au cours de la période 1931–2006.

La construction des parts migratoires repose sur les données du NIDS et s'écrit :

$$w_{od} = \frac{M_{od}}{\sum_{o' \neq d} M_{o'd}} \quad (2.4)$$

où M_{od} désigne le nombre de migrants ayant quitté le district d'origine o pour s'installer dans le district de destination d avant la période d'étude. Ainsi, w_{od} correspond à la fraction des arrivées historiques dans d en provenance de o .

Les flux migratoires prédits entre chaque paire origine–destination sont alors obtenus comme suit :

$$\widehat{\text{Migrants}}_{o \rightarrow d} = \widehat{\text{Migrants}}_o \times w_{od} \quad (2.5)$$

Ces flux prédits sont ensuite agrégés au niveau du district de destination et normalisés par la population totale de ce district afin d'obtenir un taux de migration prédit :

$$Z_d = \frac{\sum_{o \neq d} \widehat{\text{Migrants}}_{o \rightarrow d}}{\text{Population}_d} \quad (2.6)$$

L'instrument Z_d capture ainsi la composante exogène de la variation du taux de migration vers le district d , déterminée uniquement par des chocs climatiques exogènes dans les districts d'origine et par la structure historique des réseaux migratoires.

Cette stratégie permet de surmonter l'endogénéité potentielle entre migration et conditions économiques locales, tout en garantissant la cohérence dimensionnelle entre l'instrument et la variable endogène de migration interne. Elle s'inscrit dans la lignée des approches *shift–share* introduites par Card (2001) et développées par la suite par Munshi (2003) et Tabellini (2022).

2.3 Présentation des variables retenues

Ce mémoire mobilise les deux premières vagues (2008 et 2010/2011) de l'enquête longitudinale sud-africaine NIDS pour analyser l'impact de la migration interne sur les résultats du marché du travail. Trois catégories de variables sont distinguées : (i) les variables dépendantes, (ii) la variable explicative principale et (iii) les variables de contrôle.

Les indicateurs de performance économique retenus sont : (1) la variation du logarithme du salaire horaire et (2) la variation du logarithme des heures hebdomadaires travaillées. L'utilisation de la

forme logarithmique permet une interprétation des coefficients en pourcentage, et la construction en différences longues (vague 2 – vague 1) neutralise les caractéristiques inobservables fixes des individus. Ces variables sont calculées uniquement pour les natifs, définis comme les individus résidant dans leur district de naissance au moment de l’enquête.

Le taux de migration utilisé dans l’analyse est défini comme un *taux de migration récente*, et non comme un stock migratoire. Pour chaque district de destination, le taux de migration correspond au nombre de migrants récents entrants rapporté à la population adulte totale résidente du district, et la variable utilisée dans l’analyse est la différence de ce taux entre les deux vagues, exprimée en points de pourcentage. Plus précisément, le numérateur correspond au nombre d’individus ayant changé de district de résidence entre la vague 1 et la vague 2 de l’enquête NIDS, tandis que le dénominateur est la population adulte résidente totale du district considéré.

L’analyse empirique principale du mémoire porte sur la variation du taux de migration *rurale–urbaine* entre les deux vagues. Ce taux mesure la variation de la proportion d’individus ayant migré d’un district rural vers un district urbain sur la période 2008–2011. Il constitue la variable de migration centrale dans les régressions causales estimées par la méthode des variables instrumentales.

Le choix de travailler sur la variation du taux de migration récente, plutôt que sur un niveau ou un stock de migrants, s’inscrit dans une logique d’identification causale. Les stocks migratoires antérieurs à la période d’étude sont susceptibles d’avoir déjà été absorbés par les marchés du travail locaux, notamment via des ajustements du capital, de la structure productive et de l’organisation du travail. En revanche, les flux récents constituent un choc marginal plus directement comparable à une variation exogène de l’offre de travail locale, ce qui est cohérent avec l’objectif d’identification poursuivi.

Cette approche est conforme à la littérature récente sur l’impact des migrations internes, qui privilégie l’analyse des flux plutôt que des stocks afin d’identifier les mécanismes d’ajustement à court et moyen terme (Jaeger *et al.*, 2018). La faible proportion moyenne de migrants récents, et en particulier de migrants ruraux–urbains, reflète ainsi la nature marginale des flux observés sur la période 2008–2011, plutôt qu’un problème de mesure ou de représentativité.

À titre complémentaire, une analyse de la migration interne globale est également menée afin de

situer les résultats principaux dans une perspective plus générale. Toutefois, comme discuté ultérieurement, la faiblesse instrumentale associée à la migration globale limite la portée causale de ces estimations, ce qui justifie que la migration rurale–urbaine demeure le cœur empirique du mémoire. Le taux de migration utilisé dans l’analyse est défini comme un *taux de migration récente*, et non comme un stock migratoire. Plus précisément, le numérateur correspond au nombre d’individus ayant changé de district de résidence entre la vague 1 et la vague 2 de l’enquête NIDS, tandis que le dénominateur est la population adulte résidente totale du district considéré.

Afin de tenir compte des conditions initiales du marché du travail et de la situation socio-économique locale, comme le préconise Jaeger *et al.* (2018) pour limiter le biais lié à la condition de départ, plusieurs variables mesurées en baseline (2008) sont introduites : le taux d’emploi moyen dans le district, le salaire initial, les heures initiales travaillées et le taux d’emploi formel.

Ces contrôles reflètent l’hétérogénéité initiale entre districts et permettent de réduire le biais de confusion lié à des trajectoires de croissance locales distinctes.

Ainsi, la stratégie empirique repose sur l’analyse des différences longues des résultats économiques (salaires et heures travaillées) en fonction des variations du taux d’immigration, tout en contrôlant les conditions initiales observées en 2008 au niveau du district de destination.

TABLE 2.1 Statistiques descriptives – niveau individuel et niveau district

Variables	Moyenne (écart-type)			Baseline (2008)	Min	Max
	Tous	Natifs	Migrants			
A. Variables au niveau individuel						
Migrant récent (indicateur)	–	–	0.016 (0.013)	0.023 (0.022)	0	1
Migrant rural–urbain (indicateur)	–	–	0.006 (0.008)	0.008 (0.013)	0	1
Migrant historique (avant 2008)	–	–	0.440 (0.071)	–	–	–
Δ Log heures travaillées	–	-0.2216 (0.9047)	–	–	-2.5903	2.8134
Δ Log salaire horaire (ZAR)	–	0.467 (1.769)	–	–	-9.6159	7.7836
B. Variables au niveau du district (agrégées)						
Stock migratoire à destination	–	0.482 (0.158)	0.517 (0.158)	–	–	–
Δ Taux d’immigration totale (%)	-1.4148 (2.6275)	–	–	–	-14.3794	5.9762
Δ Taux de migration rurale–urbaine (%)	-0.4382 (1.2700)	–	–	–	-5.4050	4.3470
Instrument climatique shift–share	-0.1212 (0.0734)	–	–	–	-0.3944	-0.03584
B.1 Variables en baseline du district de destination						
Taux d’emploi initiaux	0.7357 (0.0971)	0.7408 (0.0922)	0.7479 (0.0876)	–	0.5440	0.9206
Salaire initial (log, ZAR)	2.256 (1.801)	2.2779 (1.803)	2.5767 (0.537)	–	-5.8483	8.842
Heures initiales travaillées (log)	3.5411 (0.2362)	3.5445 (0.2388)	3.5551 (0.2636)	–	2.737	3.797
Taux d’emploi formel initiaux	0.2648 (0.1262)	0.2698 (0.1210)	0.2888 (0.1240)	–	0.0488	0.6064

Notes : 1. Les variables de variation (préfixées par Δ) sont calculées en différences longues entre la vague 2 (2010–2011) et la vague 1 (2008), 2008 la vague initiale, pour les individus observés dans les deux vagues de l’enquête. 2. Les variables de salaire et d’heures travaillées sont exprimées en logarithme. 3. Les observations correspondent à la population adulte (14 572). Le total d’observations s’explique par l’attrition du panel entre les vagues et par la restriction à la population en âge de travailler. Parmi ces individus, 7 667 sont des natifs et 7 856 sont des migrants internes, dont 6 662 migrants historiques. 4. Les statistiques descriptives de référence (baseline) sont calculées à partir de l’échantillon individuel de la vague 1, puis agrégées au niveau des districts administratifs (52 districts).

Le tableau 2.1 présente les statistiques descriptives des principales variables pour l'échantillon adulte de l'enquête NIDS, en distinguant les valeurs pour l'ensemble des individus, les natifs, les migrants internes récents ainsi que les conditions initiales observées en 2008.

L'analyse met en évidence plusieurs tendances structurantes concernant la dynamique migratoire et les conditions initiales du marché du travail.

Concernant les flux, la variation moyenne du taux de migration est négative (-1.4148 points de pourcentage, écart-type 2.6275), ce qui traduit une contraction nette des mobilités internes entre la vague 1 et la vague 2. Le cas particulier de la migration rurale-urbaine confirme ce constat : sa variation nette moyenne est de -0.4382 (écart-type 1.27), avec un minimum de -5.405 et un maximum de 4.347 , ce qui révèle une forte hétérogénéité territoriale dans la dynamique des flux. Ces résultats rejoignent les observations de Posel (2006) qui soulignent la discontinuité des mobilités internes en Afrique du Sud.

Les stocks migratoires révèlent l'importance des dynamiques historiques : les migrants installés avant 2008 constituent près de 44 % de l'échantillon (écart-type 0.071), tandis que les migrants récents représentent seulement 1.6 % (écart-type 0.013). La migration rurale-urbaine, historiquement dominante, reste marginale dans la période récente (0.6 % en moyenne). Ces chiffres confirment que la période 2008-2010 reflète davantage des ajustements marginaux que de nouveaux flux massifs, comme le souligne Kollamparambil (2017).

L'instrument climatique, construit à partir des anomalies de température observées dans les districts d'origine et agrégées selon une logique *shift-share*, présente une moyenne négative (-0.1212 , écart-type 0.0734). Il varie entre -0.3944 et -0.03584 . Ce signe négatif moyen reflète les conditions climatiques effectivement observées sur la période considérée (2007-2009), durant laquelle les températures de court terme dans les districts d'origine sont, en moyenne, inférieures à leurs moyennes historiques de long terme.

La dispersion relativement modérée de l'instrument indique une variation significative mais concentrée autour de valeurs négatives, cohérente avec la nature standardisée du choc climatique.

Étant donné que l'instrument est défini comme une somme pondérée de ces anomalies par des parts

migratoires historiques strictement positives, son signe moyen négatif résulte mécaniquement de la distribution des chocs climatiques sous-jacents. Le signe de l'instrument n'a pas d'interprétation causale en soi ; seule sa variation transversale entre districts de destination est exploitée pour l'identification. Cette construction s'inscrit dans la tradition des instruments *shift-share* introduits par Boustan (2010) et développés ultérieurement dans la littérature empirique, adaptés ici au contexte de la migration interne en Afrique du Sud.

Les variables dépendantes montrent des contrastes marqués. La variation des heures travaillées est en moyenne négative pour les natifs (-0.2216 , écart-type 0.9047), avec des extrêmes allant de -2.5903 à 2.8134 , ce qui traduit une réduction modérée mais hétérogène du volume horaire. En revanche, la variation du salaire est en moyenne positive pour les natifs (0.3262 , écart-type 1.6802), bien que caractérisée par une forte dispersion, avec des valeurs allant de -9.6159 à 7.7836 .

Les conditions initiales observées en 2008, agrégées au niveau des districts, confirment ce diagnostic. Le taux d'emploi moyen s'élève à 0.736 (écart-type 0.097), avec des niveaux comparables entre natifs (0.741) et migrants (0.748). Le salaire initial moyen est de 2.256 en log (écart-type 1.8), légèrement plus élevé pour les migrants (2.5767) que pour les natifs (2.2779). Les heures initiales travaillées se situent autour de 3.54 en log (écart-type 0.24), avec des valeurs très proches entre natifs (3.54) et migrants (3.56). Enfin, le taux d'emploi formel reste faible (0.265 , écart-type 0.126), mais légèrement plus élevé pour les migrants (0.289) que pour les natifs (0.270).

Ces écarts modestes suggèrent que, dès la baseline de 2008, les migrants disposent d'une insertion légèrement plus favorable sur certains indicateurs, en particulier les salaires et l'emploi formel, même si l'ensemble du marché du travail reste marqué par une forte informalité.

Dans l'ensemble, ces résultats décrivent un marché du travail caractérisé par une contraction des flux migratoires récents, une hétérogénéité salariale élevée et une informalité persistante. Ils renforcent la pertinence d'intégrer les conditions initiales comme variables de contrôle dans l'analyse empirique et justifient le recours à un instrument climatique pour identifier les effets causaux de la migration sur les résultats économiques des natifs, conformément aux approches de Basso et Peri (2015).

Enfin, la baisse relative des flux récents, en particulier des migrations rurales-urbaines, laisse supposer que ce type de mobilité peut générer des ajustements spécifiques sur les salaires et les heures

travaillées des natifs, ce qui justifie l'analyse différenciée.

Le chapitre suivant présentera la méthodologie économétrique retenue pour estimer l'impact causal de la migration interne sur les conditions de travail des natifs, en exploitant la variation exogène induite par l'instrument climatique et en tenant compte des caractéristiques structurelles du marché du travail sud-africain mises en évidence par ces statistiques descriptives.

CHAPITRE 3

MÉTHODOLOGIE

Ce chapitre expose le cadre méthodologique retenu pour estimer l'effet causal de la migration interne sur les résultats du marché du travail des natifs sud-africains. Il s'inscrit dans la continuité des travaux pionniers en économie régionale (Harris et Todaro, 1970) et s'appuie sur les avancées récentes en économétrie des chocs exogènes et des effets spatiaux (Adao *et al.*, 2019).

La stratégie empirique repose sur une approche en longue différence combinée à une identification en variables instrumentales. Les estimations principales se concentrent sur les flux de migration rurale-urbaine, configuration pour laquelle l'instrument climatique de type *shift-share* est le plus pertinent du point de vue de l'identification causale. Une analyse complémentaire fondée sur la migration interne globale est proposée dans un second temps afin d'examiner l'hétérogénéité des effets.

Premièrement, l'analyse est menée au niveau individuel en longue différence entre la vague 1 et la vague 2 de l'enquête NIDS. Cette approche neutralise les caractéristiques inobservables fixes propres à chaque individu. La variable explicative principale est la variation du taux de migration interne entrant dans le district de destination. Elle est instrumentée par un choc climatique exogène construit selon une approche de type *shift-share*. Les effets de la migration sont estimés sur deux variables dépendantes : la variation du logarithme du salaire horaire et celle du logarithme du nombre d'heures travaillées par semaine. Conformément aux recommandations de Jaeger *et al.* (2018), seules les caractéristiques économiques initiales mesurées à la vague 1 (taux d'emploi, salaire initial, heures initiales travaillées et taux d'emploi formel) sont incluses comme contrôles, afin de limiter les biais liés aux conditions de départ.

Deuxièmement, nous estimons un modèle de base par moindres carrés ordinaires (MCO), qui documente les corrélations entre la variation des flux migratoires internes et celle des indicateurs du marché du travail. Les deux ensembles de variables sont exprimés en longues différences, conformément à une approche descriptive inspirée de Kleemans et Magruder (2018). Cette étape fournit une première évaluation non causale des associations entre migration et performance économique locale.

Troisièmement, pour surmonter les biais d’endogénéité largement documentés dans la littérature sur les migrations (Borjas, 2003), nous mobilisons une stratégie d’identification instrumentale. L’instrument repose sur deux sources exogènes combinées : (i) des anomalies climatiques mesurées dans les districts d’origine, construites à partir de données satellitaires selon la méthodologie de Dell *et al.* (2014), et (ii) la structure spatiale des flux migratoires historiques observés dans les données NIDS, conformément à une logique shift–share formalisée par Bazzi *et al.* (2017). Cette combinaison permet de générer une mesure prédite du changement des flux migratoires entrants dans chaque district entre les vagues 1 et 2, indépendante des conditions économiques locales contemporaines.

Enfin, la validité statistique de l’instrument est évaluée à l’aide de tests de première étape avec erreurs-types clusterisées par district de destination, notamment la statistique de Kleibergen–Paap, recommandée pour les modèles hétéroscédastiques (Angrist et Pischke, 2009), ainsi que le test d’Anderson–Rubin.

3.1 Description et justification des variables

Cette étude mobilise un ensemble cohérent de variables, dont le choix repose sur des fondements théoriques et empiriques établis dans la littérature. Nous distinguons les variables de contrôle, la variable explicative principale ainsi que les variables dépendantes, en justifiant leur inclusion dans le modèle.

En réponse aux limites potentielles des instruments de type shift–share mises en évidence par Jaeger *et al.* (2018), nous contrôlons exclusivement les caractéristiques économiques initiales du district de destination, mesurées en 2008 (période initiale). L’objectif est de neutraliser les biais liés aux conditions de départ sur le marché du travail. Ces conditions initiales incluent le logarithme du salaire initial, le logarithme des heures initiales travaillées, le taux d’emploi et le taux d’emploi formel. Leur inclusion tient compte de l’hétérogénéité économique des districts et des écarts structurels d’accès à l’emploi formel, particulièrement marqués en Afrique du Sud (Wittenberg et Collinson, 2007).

La variable explicative centrale est la variation du taux de migration interne récente entrant dans le district de résidence, calculée en longue différence entre la vague 2 et la vague 1. Ce choix s’appuie sur une littérature abondante en économie du travail et en économie régionale qui analyse les effets de la mobilité de population sur les économies locales (Borjas, 2003; Card, 2001; Basso et Peri, 2015).

Une distinction est opérée entre migration totale, migration rurale–urbaine et migration récente, afin de documenter l’impact spécifique de ces flux sur le marché du travail des natifs, en cohérence avec les travaux récents sur les pays en développement (Henderson *et al.*, 2017).

La variation du taux de migration est instrumentée par un choc climatique exogène, combinant anomalies de température dans les districts d’origine et schémas migratoires historiques, selon une logique shift–share inspirée de Boustan (2010); Dell *et al.* (2014); Bazzi *et al.* (2017). Ce cadre d’identification s’inscrit dans une littérature croissante exploitant des variations naturelles (climat, géographie) pour analyser les déterminants et les effets de la migration. Par exemple, Naidu et Hornbeck (2014) montrent qu’un choc environnemental majeur a provoqué une migration interne durable aux États-Unis, tandis que Henderson *et al.* (2017) documentent le rôle de la variabilité climatique dans l’urbanisation africaine.

Deux variables dépendantes sont analysées : le logarithme du salaire horaire et le logarithme des heures travaillées hebdomadaires. Elles sont exprimées en log puis transformées en longues différences, ce qui permet d’interpréter les coefficients comme des variations en pourcentage. Le choix de ces indicateurs s’appuie sur la distinction théorique entre ajustements par les prix et par les quantités (Basso et Peri, 2015). Une variation du salaire reflète des pressions concurrentielles ou des ajustements de productivité, tandis qu’une variation des heures travaillées traduit des ajustements comportementaux ou institutionnels. Leur combinaison offre une compréhension complémentaire de la manière dont les marchés locaux absorbent les chocs migratoires.

Enfin, à titre descriptif, nous considérons également le taux d’émigration, afin de suivre l’évolution des sorties de population depuis chaque district. Cet indicateur n’entre pas directement dans les régressions principales, mais il éclaire la dynamique migratoire sud-africaine.

3.2 Spécification du modèle de base

Le modèle économétrique de base vise à estimer l’effet de la migration interne sur les résultats du marché du travail des natifs sud-africains. L’analyse repose sur une approche en longue différence entre la vague 2 et la vague 1 de l’enquête NIDS, méthode adaptée à l’étude des dynamiques migratoires et de leurs effets sur les trajectoires économiques individuelles (Wooldridge, 2010).

Le modèle de base est estimé par moindres carrés ordinaires (MCO), dans la lignée des travaux fondateurs de Card (2001) et Borjas (2003). Dans les estimations principales, la variable de migration correspond à la variation du taux de migration rurale–urbaine entrant dans le district de destination, et dans les estimations complémentaires, il s’agit de la migration totale, peu importe l’origine et la destination. Le modèle de base est estimé par moindres carrés ordinaires (MCO) selon la spécification suivante :

$$\Delta \log(y_{id}) = \beta_1 \cdot \Delta migration_d + \mathbf{X}_d \gamma + \varepsilon_{id} \quad (3.1)$$

où $\Delta \log(y_{id})$ désigne la variation logarithmique du salaire horaire ou du nombre d’heures travaillées pour l’individu i résidant dans le district d . La variable $\Delta migration_d$ correspond à la variation du taux de migration rurale–urbaine entre la vague 2 et la vague 1. Le vecteur \mathbf{X}_d regroupe les conditions économiques initiales mesurées à la vague 1.

Les erreurs-types sont clusterisées au niveau du district de destination afin de tenir compte de la corrélation intra-district et de produire des inférences robustes à l’hétéroscédasticité.

L’interprétation du coefficient β_1 est directe : il mesure la variation en pourcentage du salaire ou des heures travaillées associée à une augmentation d’un point de pourcentage du taux de migration interne. Ce type de formulation est largement mobilisé dans la littérature (Dustmann *et al.*, 2012; Basso et Peri, 2015).

Cependant, l’estimation de l’équation (3.1) par moindres carrés ordinaires (MCO) est susceptible de souffrir de plusieurs sources d’endogénéité, qui menacent l’interprétation causale de β_1 .

Premièrement, comme le soulignent Borjas (2003) et Kleemans et Magruder (2018), la migration interne n’est pas aléatoire : les individus choisissent leur destination en fonction des opportunités économiques anticipées. Cela crée un problème de causalité inverse (*reverse causality*) : le taux de migration observé dans un district est corrélé non seulement à la migration, mais aussi à des facteurs inobservés (comme l’expansion sectorielle, les investissements publics ou les politiques locales d’emploi) qui influencent simultanément les salaires et l’emploi des natifs.

Deuxièmement, des facteurs de confusion omis (*omitted variable bias*) peuvent biaiser l'estimation. Des chocs de demande locaux, des changements technologiques spécifiques à un secteur, ou des améliorations des infrastructures dans le district de destination peuvent à la fois attirer des migrants et augmenter la productivité et les salaires des natifs. Inversement, des chocs négatifs (déclin industriel, désinvestissement) peuvent simultanément réduire l'attractivité migratoire et détériorer les conditions du marché du travail.

Troisièmement, des erreurs de mesure dans la construction du taux de migration ΔM_d , notamment liées à la sous-déclaration des flux informels ou aux limites des frontières administratives des districts, peuvent introduire un biais d'atténuation (*attenuation bias*) vers zéro.

Conséquemment, une estimation OLS de β_1 tend à être biaisée et inconsistante. La direction du biais est a priori ambiguë. Si les migrants sont attirés sélectivement par les districts connaissant une expansion économique dynamique (migration pro-cyclique), alors le coefficient OLS sera vraisemblablement surestimé (biais positif). À l'inverse, si les migrants s'installent préférentiellement dans des zones en déclin structurel pour des raisons de coût du logement réduit, de présence de réseaux familiaux ou d'opportunités dans le secteur informel, le biais peut être négatif.

Pour corriger ces biais d'endogénéité et identifier un effet causal, nous utilisons une stratégie en variables instrumentales (VI) en deux étapes (*2SLS*). L'instrument mobilisé exploite la variation exogène générée par la combinaison de deux sources : (i) des chocs climatiques dans les districts d'origine des migrants, et (ii) les profils migratoires historiques observés avant la période d'étude. Cette approche de type *shift-share* (ou « variation partagée »), formalisée par Bazzi *et al.* (2017) et appliquée dans des contextes migratoires par Boustan (2010) et ?, permet de construire une prédiction exogène des flux migratoires entrants, indépendante des conditions économiques contemporaines dans les districts de destination.

3.3 Identification et construction de l'instrument

L'identification de l'effet causal de la migration interne repose sur une stratégie en variables instrumentales (IV). Nous mobilisons un instrument climatique de type *shift-share*, dont la construction technique est détaillée au chapitre 2. Nous insistons ici sur la logique d'identification, la structure temporelle des chocs climatiques et les hypothèses de validité sous-jacentes.

Le principe de l'instrument repose sur l'idée que des chocs climatiques anormaux dans les districts d'origine modifient la propension des individus à émigrer. Ces chocs sont mesurés exclusivement *avant* les décisions de migration et sont exogènes aux conditions économiques des districts de destination. Ils influencent principalement la migration sortante depuis les zones d'origine, sans affecter directement les résultats du marché du travail à destination (Maccini et Yang, 2009). Combinés avec les parts migratoires historiques origine–destination, ils permettent de prédire la répartition géographique des flux migratoires contemporains (Munshi, 2003; Bazzi *et al.*, 2017).

Sur le plan temporel, les chocs climatiques sont définis séparément pour chaque vague de l'enquête. Pour la vague 1 (2008), le choc est construit à partir des températures moyennes observées entre janvier et décembre 2007. Pour la vague 2 (2010–2011), il repose sur les températures observées entre octobre 2008 et septembre 2009. Cette construction garantit que les conditions climatiques utilisées précèdent systématiquement les mouvements migratoires observés à chaque vague.

Ce mécanisme assure la pertinence de l'instrument : les chocs climatiques affectent bien la variation des taux de migration. L'exclusion repose sur l'hypothèse que les anomalies climatiques dans les districts d'origine n'ont pas d'effet direct sur les salaires et les heures travaillées des natifs dans les districts de destination, en dehors de leur impact sur la migration reçue. Cette hypothèse est renforcée par la segmentation des marchés du travail sud-africains et le caractère localisé des chocs environnementaux (Barrios *et al.*, 2006; Kelley *et al.*, 2015).

Bien que la définition retenue du statut de migrant ne distingue pas explicitement la durée de résidence dans le district de destination, cette approche demeure cohérente avec la stratégie d'identification adoptée. En effet, l'instrument climatique mobilisé dans ce mémoire capte exclusivement des variations exogènes récentes des flux migratoires, induites par des chocs climatiques antérieurs aux vagues d'enquête. Dans le cadre de l'estimation en deux étapes, seule la composante de la migration expliquée par ces chocs est exploitée pour l'identification causale, ce qui implique que les migrations anciennes, non corrélées aux chocs climatiques récents, ne contribuent pas à l'effet estimé. L'effet identifié doit ainsi être interprété comme l'impact causal d'un choc de migration récente sur les résultats du marché du travail des natifs, indépendamment de la présence de migrants installés de longue date.

L'instrument climatique shift–share présente plusieurs avantages. Premièrement, il repose sur une source de variation clairement exogène, à savoir les anomalies de température ou de précipitation, qui ne sont pas influencées par les conditions économiques locales (Dell *et al.*, 2014; Maccini et Yang, 2009). Deuxièmement, l'utilisation de parts migratoires historiques permet de capter la dimension structurelle des réseaux migratoires et de refléter la manière dont les flux passés conditionnent les destinations futures (Munshi, 2003; Imbert *et al.*, 2022). Troisièmement, cet instrument a été largement validé dans la littérature empirique, tant pour l'étude de la migration internationale (Bazzi *et al.*, 2017; Jaeger *et al.*, 2018) que pour l'analyse de la migration interne dans les pays en développement (Henderson *et al.*, 2017). Enfin, la combinaison climat–réseaux migratoires permet de contourner les biais de simultanéité entre migration et performance économique des districts de destination, fréquents dans les modèles corrélatifs (Borjas, 2003).

L'instrument présente néanmoins certaines limites. Tout d'abord, l'hypothèse d'exclusion peut être remise en question si les chocs climatiques dans les zones d'origine affectent indirectement les districts de destination par d'autres canaux (par exemple, via les prix agricoles ou des transferts interrégionaux) (Tabellini, 2022). Ensuite, l'utilisation de parts migratoires historiques suppose une relative stabilité des structures migratoires dans le temps, ce qui peut être critiqué si les réseaux évoluent rapidement en réponse à des changements institutionnels ou économiques (Jaeger *et al.*, 2018). Par ailleurs, certains auteurs soulignent que les instruments shift–share peuvent être sensibles au choix de la période de référence et à la granularité spatiale retenue (Adao *et al.*, 2019).

Une autre limite potentielle concerne le biais d'attrition : si les chocs climatiques influencent à la fois les décisions migratoires et la probabilité qu'un individu reste observé dans le panel (par exemple, en augmentant les non-réponses ou les sorties d'échantillon dans les zones affectées), alors l'instrument pourrait refléter à la fois un effet migratoire et un biais de sélection. Dans ce cas, l'exogénéité conditionnelle de l'instrument serait affaiblie.

Enfin, la puissance de l'instrument dépend de l'ampleur et de la fréquence des chocs climatiques : dans des contextes où ces chocs sont rares ou très localisés, l'identification peut manquer de force statistique.

Ce canal d'identification est particulièrement pertinent pour les migrations rurales–urbaines, dans

la mesure où les chocs climatiques affectent prioritairement les zones rurales et influencent principalement les décisions de départ vers les centres urbains. Cette spécificité s’explique par la plus grande dépendance des économies rurales aux activités sensibles aux conditions climatiques, notamment l’agriculture et l’élevage, qui représentent une part importante des moyens de subsistance dans ces régions. Les anomalies climatiques (sécheresses, variations de température) perturbent directement la production agricole et réduisent les revenus des ménages ruraux, créant ainsi une incitation accrue à migrer vers les centres urbains à la recherche d’opportunités économiques alternatives.

En revanche, les zones urbaines, caractérisées par une plus grande diversification économique et une moindre dépendance aux activités primaires, sont moins directement affectées par ces chocs climatiques dans leurs décisions de mobilité interne. L’instrument climatique capture donc plus précisément les variations exogènes de la migration rurale–urbaine, offrant une identification plus propre des effets causaux sur les marchés du travail urbains.

Pour cette raison, les régressions principales de ce mémoire sont estimées sur un échantillon restreint aux flux ruraux–urbains, tandis que la migration interne globale est analysée dans un cadre complémentaire.

3.4 Méthode et spécification économétrique

L’effet causal de la migration est estimé à l’aide d’un modèle en deux étapes (2SLS). La première étape prédit la variation du taux de migration à partir de l’instrument climatique, tandis que la seconde étape estime l’effet de la migration instrumentée sur les résultats du marché du travail des natifs.

3.4.0.0.1 Première étape.

La première étape consiste à estimer la variation du taux de migration interne au niveau du district de destination à partir de l’instrument climatique :

$$\Delta\text{Migration}_d = \pi_0 + \pi_1 Z_d + \gamma' \mathbf{X}_d + \eta_d \tag{3.2}$$

où $\Delta \text{Migration}_d$ désigne la variation en points de pourcentage du taux de migration observé dans le district d entre la vague 1 et la vague 2. L'instrument Z_d correspond au terme *shift-share*, construit à partir des anomalies climatiques observées dans les districts d'origine, pondérées par les parts migratoires historiques. Ainsi, un district recevant historiquement des migrants en provenance de zones ayant subi des anomalies climatiques plus marquées connaîtra une variation plus forte des flux entrants.

Le vecteur \mathbf{X}_d regroupe les conditions économiques initiales mesurées à la vague 1 (salaire, heures travaillées, taux d'emploi et taux d'emploi formel). L'erreur η_d capte les facteurs inobservés susceptibles d'influencer la variation du taux de migration.

Le coefficient π_1 mesure l'effet des chocs climatiques instrumentés sur la variation du taux de migration interne. La pertinence de l'instrument est évaluée à l'aide de la statistique de Kleibergen–Paap et du test d'Anderson–Rubin. Conformément aux tables de Stock et Yogo (2005), le seuil critique associé à un biais maximal de 20 % est de 8,7 dans le cas d'un modèle avec une seule variable endogène et un seul instrument. Ainsi, une statistique de première étape supérieure à ce seuil garantit que le biais de l'estimateur IV ne dépasse pas 20 % de celui de l'estimateur OLS.

3.4.0.0.2 Deuxième étape.

La deuxième étape exploite la composante instrumentée du flux migratoire prédite à la première étape (au niveau du district) pour estimer son effet sur les résultats économiques des individus résidant dans ce district :

$$\Delta \log(y_{id}) = \beta \cdot \widehat{\text{migration}}_d + \boldsymbol{\theta} \cdot \mathbf{X}_{id} + \varepsilon_d \quad (3.3)$$

où $\Delta \log(y_{id})$ désigne la variation logarithmique du salaire horaire mensuel ou du nombre d'heures travaillées de l'individu i dans le district d . La variable $\widehat{\text{migration}}_d$ représente la composante instrumentée de la migration entrante prédite à la première étape.

L'interprétation du coefficient β est la suivante : il mesure le pourcentage de variation du salaire ou des heures travaillées des natifs associé à une augmentation d'un point de pourcentage du taux de

migration interne. Les variables de contrôle \mathbf{X}_d sont les conditions initiales au niveau du district.

Ce dispositif économétrique permet de capter une variation exogène de la migration et d’identifier rigoureusement son effet causal sur les performances économiques des natifs. La section suivante évalue empiriquement la validité de cette stratégie à l’aide des tests de robustesse (Kleibergen–Paap et Anderson–Rubin), essentiels pour garantir la solidité de l’approche IV (Andrews *et al.*, 2019).

3.4.1 Validité et inférence dans les dispositifs *shift–share*

L’usage d’un instrument *shift–share* peut induire des dépendances croisées lorsque plusieurs destinations partagent les mêmes origines. Nous adoptons une inférence clusterisée au niveau du district de destination et documentons la distribution des parts migratoires historiques afin de garantir la validité de l’inférence (Adao *et al.*, 2018; Jaeger *et al.*, 2018).

Le test d’Anderson–Rubin permet de tester la significativité du coefficient de seconde étape sans conditionner à la force de l’instrument. Une interprétation alternative de ce test est toutefois que la significativité observée pourrait refléter une violation de la condition d’exclusion, si les chocs climatiques mesurés dans les districts d’origine affectaient directement les résultats du marché du travail dans les districts de destination, indépendamment des flux migratoires.

Plusieurs éléments du dispositif empirique visent à limiter ces canaux directs. Premièrement, les chocs climatiques sont mesurés exclusivement dans les districts d’origine et antérieurement aux décisions de migration, ce qui réduit la probabilité d’un effet simultané sur les conditions économiques à destination. Deuxièmement, l’ensemble des spécifications inclut des contrôles pour les conditions initiales du marché du travail dans les districts de destination (salaires, emploi, heures travaillées et emploi formel mesurés à la vague 1), afin de neutraliser les trajectoires économiques locales préexistantes. Troisièmement, l’analyse principale se concentre sur les flux ruraux–urbains, pour lesquels les chocs climatiques constituent un déterminant bien documenté de la migration sortante, alors que les mécanismes de transmission directe vers les zones urbaines de destination sont plus limités.

Dans ce cadre, le test d’Anderson–Rubin est interprété comme une condition nécessaire à l’inférence causale, sous l’hypothèse que l’instrument climatique n’affecte les résultats à destination qu’à travers son impact sur la migration. Les résultats sont donc interprétés avec prudence, en cohérence avec

les hypothèses structurelles du modèle et la littérature sur les instruments climatiques de type *shift–share*.

3.5 Analyse complémentaire

La stratégie empirique principale de ce mémoire repose sur une spécification économétrique unique, présentée aux équations (3.2) et (3.3). L’analyse causale se concentre délibérément sur la migration rurale–urbaine, qui constitue non seulement la forme historiquement dominante de mobilité interne en Afrique du Sud, mais aussi le canal théoriquement le plus direct par lequel les chocs d’offre de travail devraient affecter les marchés du travail urbains des districts de destination.

Cette focalisation sur les flux ruraux–urbains, bien que méthodologiquement justifiée, soulève néanmoins trois questions auxquelles une analyse d’hétérogénéité basée sur la migration interne globale permet de répondre.

Premièrement, il s’agit d’évaluer la validité externe et la généralisation des résultats. Les effets estimés sur le sous-ensemble des flux ruraux–urbains sont-ils généralisables à l’ensemble des mouvements migratoires internes, ou capturent-ils des mécanismes spécifiques à ce type de flux ? Une divergence significative des résultats mettrait en lumière l’hétérogénéité des effets de la migration selon sa nature, tandis qu’une convergence suggérerait un effet agrégé plus uniforme.

Deuxièmement, cette analyse constitue un test de robustesse et d’interprétation causale exigeant. L’utilisation d’un instrument climatique *shift–share* pour les flux globaux permet de vérifier si l’instrument, conçu à l’origine pour les flux ruraux–urbains motivés par des chocs agricoles, conserve sa puissance et produit des effets similaires. Si c’est le cas, cela renforce la crédibilité d’un canal causal commun. À l’inverse, un instrument faible ou des résultats divergents confirmeraient que les moteurs et les effets de la migration diffèrent fondamentalement selon les types de flux, validant ainsi le choix de les étudier séparément.

Troisièmement, l’analyse offre un apport politique et facilite la lecture de la littérature. Les décideurs publics s’intéressent souvent à l’impact agrégé de la mobilité interne. En fournissant une estimation pour la migration globale, cette analyse permet un dialogue direct avec les études existantes qui ne font pas cette distinction (Bryceson, 2000; Beauchemin, 2011) et informe des politiques qui ciblent

la migration dans son ensemble.

Sur le plan méthodologique, cette analyse complémentaire est menée en maintenant strictement identiques tous les éléments du cadre principal : les équations de première et de seconde étapes (3.2) et (3.3), la stratégie d'identification par variables instrumentales, la construction de l'instrument climatique de type *shift-share*, et l'ensemble des variables de contrôle (incluant les conditions initiales de 2008). La seule modification opérée est la substitution de la variable de traitement : la variation du taux de migration globale entrant dans un district remplace la variation du taux de migration rurale-urbaine.

Cette démarche produit ainsi un contrefactuel méthodologique précieux. Elle permet d'isoler l'effet de la définition du traitement en maintenant constants l'identification, l'estimation et les contrôles. L'interprétation des résultats est donc double. Des effets similaires entre les deux spécifications indiqueraient que l'impact sur le marché du travail est principalement porté par l'effet de volume total des nouveaux arrivants, indépendamment de leur origine précise. À l'inverse, des effets divergents souligneraient l'importance de la composition des flux migratoires (rurale-urbaine vs autres) et des mécanismes sous-jacents spécifiques, tels que les différentiels de qualifications ou les secteurs d'insertion.

En définitive, cette analyse complémentaire ne remet pas en cause la primauté analytique accordée aux flux ruraux-urbains, dont la cohérence avec l'instrument climatique reste le socle de l'identification causale. Elle enrichit plutôt le mémoire en situant ces résultats dans un spectre plus large, en testant la robustesse de l'interprétation, et en fournissant des estimations directement utiles pour le débat politique sur la mobilité interne dans son ensemble.

CHAPITRE 4

RÉSULTATS EMPIRIQUES

Ce chapitre présente les résultats empiriques de l'analyse économétrique visant à estimer l'effet causal de la migration interne sur les résultats du marché du travail des natifs sud-africains. Les biais potentiels d'endogénéité sont corrigés à l'aide d'une stratégie de variables instrumentales en deux étapes (2SLS) (Angrist et Pischke, 2009). Par convention, la variable « salaire (log) » désigne le logarithme du salaire horaire, tandis que la variable « heures travaillées (log) » correspond au logarithme du nombre d'heures hebdomadaires, toutes deux exprimées en différences longues entre la vague 1 (2008) et la vague 2 (2010-2011).

Les résultats reposent sur une série de spécifications économétriques estimées en différences longues entre ces deux vagues de l'enquête NIDS. La variable dépendante est successivement le salaire horaire et le nombre d'heures travaillées hebdomadaires des natifs, exprimés en logarithme afin d'interpréter les coefficients comme des variations relatives.

La variable explicative principale est la variation du taux de migration entrante au niveau du district de destination. Conformément à la stratégie empirique du mémoire, l'analyse causale principale se concentre sur la migration *rurale-urbaine*, qui constitue la forme historiquement dominante de la migration interne en Afrique du Sud. Ces flux sont particulièrement susceptibles d'exercer une pression spécifique sur les marchés du travail urbains et constituent donc la variable de migration principale dans les régressions causales, tandis que la migration interne globale est examinée dans une section distincte à titre complémentaire.

La stratégie d'identification repose sur un instrument climatique de type *shift-share* (Adao *et al.*, 2018), construit à partir des anomalies climatiques observées dans les districts d'origine et pondérées par les parts migratoires historiques (Card, 2001; Boustan, 2010; Bazzi *et al.*, 2017), conformément à la méthodologie détaillée au chapitre 2. L'ensemble des modèles inclut comme variables de contrôle les conditions économiques initiales mesurées à la vague 1, conformément aux recommandations de Jaeger *et al.* (2018).

Les estimations sont réalisées selon deux approches. Dans un premier temps, des régressions par

moindres carrés ordinaires (MCO) sont estimées afin de documenter les corrélations brutes entre migration et résultats du marché du travail. Dans un second temps, la migration est instrumentée à l'aide de l'instrument climatique de type *shift-share*, afin de corriger les biais d'endogénéité potentiels liés aux choix de localisation des migrants.

Les spécifications empiriques sont présentées successivement sans variables de contrôle, puis avec l'introduction de conditions économiques initiales mesurées à la vague 1, incluant le salaire, les heures travaillées, le taux d'emploi et le taux d'emploi formel au niveau du district. Ce choix vise à neutraliser l'hétérogénéité préexistante entre territoires avant la migration.

L'ensemble des estimations utilise des erreurs types robustes et clusterisées au niveau du district de destination. Les tableaux de résultats sont organisés de manière cohérente avec cette progression, en distinguant clairement les estimations MCO et VI, ainsi que les spécifications avec et sans contrôles.

L'organisation de ce chapitre suit une démarche progressive, allant de la validation de la stratégie d'identification à l'analyse des effets causaux principaux, puis à leur mise en perspective comparative. La section 4.1 présente d'abord les résultats centraux du mémoire, en se concentrant sur les flux de migration *rurale-urbaine* pour lesquels l'identification causale est la plus robuste. Cette section s'ouvre par la validation de la puissance instrumentale, puis expose les effets estimés successivement sans contrôles et avec l'introduction des variables de contrôle économiques initiales mesurées à la vague 1. La section 4.2 élargit ensuite l'analyse en examinant la migration interne globale, afin d'évaluer dans quelle mesure les effets observés pour les flux ruraux-urbains s'inscrivent dans des mécanismes plus généraux ou reflètent au contraire des dynamiques spécifiques à ce type particulier de migration.

4.1 Effets de la migration rurale-urbaine sur les salaires et les heures travaillées

Cette section présente les résultats centraux du mémoire. Conformément à la stratégie méthodologique exposée au chapitre 3, l'analyse causale principale porte sur les flux de migration rurale-urbaine, pour lesquels l'instrument climatique de type *shift-share* présente une puissance statistique suffisante et une interprétation causale crédible.

Les estimations sont conduites en deux temps. Nous présentons d'abord la validation de l'instrument

pour la migration rurale–urbaine, puis les effets estimés de cette migration sur les résultats du marché du travail des natifs, successivement sans variables de contrôle et avec l’introduction des conditions économiques initiales mesurées à la vague 1 (salaire, emploi, heures travaillées et emploi formel).

4.1.1 Validation de l’instrument pour la migration rurale–urbaine

Avant d’interpréter les résultats issus de la méthode des variables instrumentales (2SLS), il est nécessaire d’évaluer la validité statistique de l’instrument climatique mobilisé.

Cette sous-section vise à évaluer la pertinence de l’instrument climatique pour l’analyse causale des effets de la migration rurale–urbaine. Celui-ci est construit à partir des anomalies standardisées de température observées dans les districts d’origine, pondérées par les parts migratoires historiques origine–destination, selon une logique de type *shift–share* inspirée de Boustan (2010); Adao *et al.* (2018); Jaeger *et al.* (2018). L’objectif est de vérifier que cet instrument est suffisamment corrélé avec la variation observée du taux de migration rurale–urbaine, sans dépendre de canaux directs vers les résultats du marché du travail. Conformément à la méthodologie des variables instrumentales, nous examinons la relation entre l’instrument (anomalies climatiques dans les districts d’origine) et la variable endogène (variation du taux de migration rurale–urbaine), qui constitue la variable explicative principale dans nos régressions de seconde étape. La variable dépendante de cette première étape est donc la variation du taux de migration rurale–urbaine, tandis que la variable indépendante principale est l’instrument climatique de type *shift–share*.

Nous procédons par une série de régressions successives. Dans un premier temps, la variation du taux de migration rurale–urbaine est expliquée uniquement par l’instrument climatique, sans aucun contrôle. Dans un second temps, nous introduisons les conditions économiques initiales mesurées à la vague 1, ajoutées progressivement dans l’équation de première étape : d’abord le salaire initial, puis les heures travaillées initiales, ensuite le taux d’emploi initial et enfin le taux d’emploi formel. Les résultats sont reportés dans le tableau 4.1.

TABLE 4.1 Première étape : instrument climatique et variation de la migration rurale–urbaine

	(1)	(2)	(3)	(4)	(5)
Instrument climatique	8,759** (3,935)	9,070** (4,300)	9,892** (4,843)	9,828** (4,787)	9,760** (4,713)
Emploi initial (%)		-0,439 (1,769)	-1,058 (1,806)	-0,832 (1,763)	-0,828 (1,774)
Salaire initial (log)			0,138 (0,125)	0,137 (0,128)	0,132 (0,122)
Heures initiales (log)				-0,283 (0,224)	-0,293 (0,232)
Emploi formel (%)					1,643 (2,594)
Kleibergen–Paap F	8.3	8.2	9.5	9	11.8
Anderson–Rubin F	11,85***	14,11**	7,93**	6,32**	13,24**
N	330	330	330	330	330
$N(\text{clusters})$	42	42	42	42	42

Notes : *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$. La variable dépendante est la variation du taux de migration rurale–urbaine entre 2008 et 2011. Les erreurs types sont robustes et clusterisées au niveau du district de destination.

Le tableau 4.1 montre que, initialement, l'échantillon comprenait 52 districts, mais que 10 d'entre eux ont été exclus en raison d'informations manquantes sur certaines variables clés. L'échantillon final est ainsi composé de 330 observations regroupées en 42 clusters, soit en moyenne 7,9 observations par district de destination.

Les résultats indiquent que l'instrument climatique est positivement et significativement corrélé à la variation du taux de migration rurale–urbaine dans toutes les spécifications. Les coefficients estimés sont compris entre 8,76 et 9,89, et sont statistiquement significatifs au seuil de 5%. Cela signifie qu'une augmentation d'une unité du choc climatique (mesuré par un z-score de température dans les districts d'origine) est associée à une hausse d'environ 9 points de pourcentage du taux de migration rurale vers les zones urbaines. Autrement dit, une anomalie climatique équivalente à un écart-type

accroît substantiellement la part des migrants quittant les zones rurales au profit des villes, ce qui est cohérent avec l'hypothèse selon laquelle les chocs climatiques constituent un puissant facteur de mobilité interne.

Les statistiques de Kleibergen–Paap F se situent entre 8,2 et 11,8, ce qui excède les seuils critiques associés à une taille maximale de biais IV de 20 % (6.66), tout en restant inférieur au seuil plus exigeant de 16,38 correspondant à 10 % Stock2005. L'instrument peut ainsi être qualifié de modérément fort, sans atteindre la puissance optimale. Cette évaluation est néanmoins renforcée par les tests d'Anderson–Rubin, qui sont significatifs dans l'ensemble des spécifications (F compris entre 7,24 et 15,11, $p < 0,05$), indiquant que l'instrument demeure statistiquement valide même en présence d'une identification imparfaite (Anderson et Rubin, 1949; Andrews *et al.*, 2019).

Dans l'ensemble, ces résultats confirment la pertinence empirique de l'instrument climatique pour expliquer la variation de la migration rurale–urbaine. Malgré une force modérée de la première étape, les tests robustes à l'instrument faible soutiennent l'utilisation d'une stratégie par variables instrumentales pour identifier l'effet causal de la migration interne sur les résultats du marché du travail des natifs dans les estimations de seconde étape présentées ci-après.

4.1.2 Effets sans contrôles

TABLE 4.2 Effets de la migration rurale–urbaine sur les salaires et les heures travaillées (sans contrôles)

	Salaire (log)		Heures travaillées (log)	
	(1) MCO	(2) VI	(3) MCO	(4) VI
Taux de migration rurale–urbaine	-0,035 (0,118)	-0,241 (0,203)	0,119*** (0,037)	0,258* (0,150)
Constante	0,457** (0,175)	0,302 (0,262)	-0,141 (0,084)	-0,047 (0,106)
N	246	246	330	330
R^2	0,001	-0,025	0,027	-0,010
Kleibergen–Paap F		8,72		11,82
Anderson–Rubin F		7,49		12,84*

Notes : *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$. Les erreurs types sont robustes et clusterisées au niveau du district de destination. Les estimations VI utilisent l’anomalie de température comme instrument du taux de migration rurale–urbaine. Les variables dépendantes sont exprimées en log-différences entre 2008 et 2011. L’échantillon pour les salaires ($N=246$) est plus petit que celui pour les heures travaillées ($N=330$) en raison de la non-réponse sur les salaires.

Pour les salaires, les estimations MCO et VI ne révèlent aucun effet statistiquement significatif : une hausse d’un point de pourcentage du taux de migration rurale–urbaine est associée à des variations de salaire non significatives de $-0,03\%$ en MCO et de $-0,24\%$ en VI. Autrement dit, les données ne permettent pas de rejeter l’hypothèse d’un effet nul de la migration rurale–urbaine sur les salaires des natifs. Ces résultats sont cohérents avec une partie de la littérature internationale, qui met en évidence une absence d’effet salarial robuste des migrations sur les travailleurs non migrants (Card, 2001; Dustmann *et al.*, 2012).

En revanche, les résultats concernant les heures travaillées indiquent un effet positif et statistiquement significatif. En l’absence de contrôles, une augmentation d’un point de pourcentage du taux de migration rurale–urbaine accroît le volume d’heures travaillées des natifs de $0,12\%$ en MCO (significatif au seuil de 1%) et de $0,26\%$ en VI (significatif au seuil de 10%). L’approche par variables instrumentales confirme ainsi que l’ajustement au choc migratoire s’opère principalement par les quantités de travail plutôt que par les prix, en ligne avec les prédictions du modèle de Harris–

Todaro (Harris et Todaro, 1970). Au-delà de la significativité statistique, l'ampleur économique de cet effet mérite d'être soulignée. Le coefficient estimé par variables instrumentales indique qu'une augmentation d'un point de pourcentage du taux de migration rurale-urbaine est associée à une hausse d'environ 0,26 % du nombre d'heures travaillées des natifs. Rapporté à un travailleur natif moyen travaillant environ 40 heures par semaine, cet effet correspond à une augmentation d'environ 0,1 heure par semaine, soit approximativement 0,4 heure supplémentaire par mois. Bien que modeste à l'échelle individuelle, cet ajustement devient économiquement significatif lorsqu'il est agrégé à l'ensemble du marché du travail local, en particulier dans un contexte urbain caractérisé par une forte intensité de main-d'œuvre et un chômage structurel élevé.

Le fait que l'ajustement s'opère par les heures travaillées plutôt que par les salaires n'est pas surprenant au regard de la littérature, en particulier dans les économies caractérisées par des rigidités salariales, un chômage structurel élevé et une forte segmentation du marché du travail, comme c'est le cas en Afrique du Sud (Harris et Todaro, 1970; Jaeger *et al.*, 2018). Une explication complémentaire de cet ajustement par les quantités de travail réside dans la segmentation sectorielle du marché du travail sud-africain, largement documentée dans la littérature. Plusieurs études montrent que les migrants ruraux-urbains sont surreprésentés dans des secteurs urbains à forte intensité de travail peu qualifié, tels que la construction, le commerce informel, les services personnels et certaines activités manufacturières légères, tandis que les natifs occupent plus fréquemment des emplois formels, relativement stables et protégés, notamment dans le secteur public et les services formels (Posel et Casale, 2003; Leibbrandt *et al.*, 2010; Turok et Borel-Saladin, 2014; Kollamparambil, 2017).

Dans ce contexte de segmentation persistante, l'arrivée de migrants internes tend à accroître l'activité économique locale et la demande de travail dans les secteurs intensifs en main-d'œuvre, sans exercer de pression directe sur les salaires des natifs. Les ajustements observés prennent alors la forme d'une augmentation des heures travaillées ou de l'intensité du travail des natifs, plutôt que d'une modification des rémunérations. Cette complémentarité sectorielle constitue ainsi un mécanisme plausible permettant d'interpréter le signe positif estimé sur les heures travaillées en l'absence de contrôles, et s'inscrit pleinement dans les analyses existantes du fonctionnement du marché du travail sud-africain.

Dans l'ensemble, ces résultats suggèrent qu'en l'absence de contrôles, l'ajustement du marché du

travail local au choc migratoire s’effectue principalement par une intensification de l’offre de travail des natifs, mesurée en heures travaillées, plutôt que par une modification des salaires.

4.1.3 Effets avec contrôles

TABLE 4.3 Effet de la migration rurale–urbaine sur les salaires et les heures travaillées des natifs (avec contrôles)

	Salaire (log)		Heures travaillées (log)	
	MCO	VI (2SLS)	MCO	VI (2SLS)
Taux de migration rurale–urbaine	0,083 (0,112)	-0,111 (0,200)	0,058 (0,042)	0,100 (0,104)
Salaire initial (log)	-0,511*** (0,128)	-0,480*** (0,111)	0,036 (0,037)	0,030 (0,038)
Taux d’emploi initial (%)	-1,825 (1,521)	-1,763 (1,577)	0,173 (0,589)	0,149 (0,577)
Heures initiales (log)	0,634 (0,423)	0,566 (0,415)	-1,210*** (0,136)	-1,196*** (0,136)
Taux d’emploi formel initial (%)	1,325 (2,302)	1,847 (2,466)	0,569 (0,628)	0,478 (0,646)
Constante	0,827 (2,048)	0,789 (2,007)	4,043*** (0,499)	4,060*** (0,488)
Kleibergen–Paap F	–	9.6	–	11
Anderson–Rubin p -val	–	7,24**	–	11,5***
R^2	0,203	0,182	0,405	0,402
N	246	246	330	330

Notes : *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$. Les erreurs types sont robustes et clusterisées au niveau du district de destination. Les estimations VI utilisent l’anomalie de température comme instrument du taux de migration rurale–urbaine. Les variables dépendantes sont exprimées en log-différences entre 2008 et 2011.

Les résultats présentés dans le tableau 4.3 confirment l’absence d’effet salarial robuste de la migration rurale–urbaine sur les natifs, y compris après l’introduction de contrôles capturant les conditions initiales du marché du travail. Les estimations MCO et VI produisent des coefficients de signe opposé pour les salaires (respectivement 0,083 et $-0,111$), mais ces effets ne sont pas statistiquement

significatifs ($p > 0,10$ dans les deux cas) pour une augmentation d'un point de pourcentage du taux de migration rurale-urbaine. Ces résultats indiquent que, conditionnellement aux caractéristiques initiales, la migration interne ne se traduit pas par une modification mesurable des salaires des natifs.

La statistique de Kleibergen-Paap associée à l'estimation VI pour les salaires ($F = 9,6$) dépasse les seuils critiques usuels de Stock-Yogo, suggérant que l'instrument fondé sur l'anomalie de température ne souffre pas d'un problème d'instrument faible. Le test d'Anderson-Rubin est significatif au seuil de 5%, ce qui indique que l'équation est identifiée, même si le coefficient estimé du taux de migration reste statistiquement indiscernable de zéro. Cette combinaison de résultats est cohérente avec une absence d'effet salarial détectable, malgré une identification satisfaisante de la relation causale.

Les résultats relatifs aux heures travaillées présentent un profil différent. L'estimation MCO suggère un effet positif mais non significatif du taux de migration rurale-urbaine sur les heures travaillées des natifs ($0,058$, $p > 0,10$). L'estimation par variables instrumentales conduit également à un coefficient positif ($0,100$), mais celui-ci demeure statistiquement non significatif. Bien que les statistiques de première étape indiquent une identification solide (statistique de Kleibergen-Paap égale à 11 et test d'Anderson-Rubin significatif), les estimations ne permettent pas de conclure à un effet causal précis de la migration rurale-urbaine sur le volume d'heures travaillées des natifs lorsque l'on contrôle pour les conditions initiales du marché du travail. Du point de vue de la signification économique, l'ampleur du coefficient estimé par variables instrumentales pour les heures travaillées ($\hat{\beta} = 0,100$) suggère qu'une augmentation d'un point de pourcentage du taux de migration rurale-urbaine serait associée à une hausse d'environ 0,10% des heures travaillées des natifs. Pour un travailleur moyen effectuant environ 40 heures par semaine, cet effet correspond à moins de 0,05 heure supplémentaire par semaine, soit un ajustement économiquement très limité. Cette faible magnitude, combinée à l'absence de significativité statistique une fois les contrôles introduits, indique que la migration rurale-urbaine n'engendre pas d'ajustement quantitativement important du temps de travail des natifs lorsque les différences structurelles initiales entre districts sont prises en compte.

Dans l'ensemble, ces résultats suggèrent que l'introduction de contrôles absorbant l'hétérogénéité initiale des districts atténue fortement les effets estimés de la migration rurale-urbaine sur les

résultats du marché du travail des natifs. L'absence d'effet salarial persiste, tandis que l'effet positif observé sur les heures travaillées dans les spécifications sans contrôles disparaît une fois les conditions initiales prises en compte. Cette évolution des coefficients indique que les ajustements observés précédemment reflétaient en partie des différences structurelles préexistantes entre les districts plutôt qu'un effet causal direct de la migration.

Ces résultats sont cohérents avec les prédictions du modèle de Harris–Todaro (Harris et Todaro, 1970), selon lesquelles, dans des économies caractérisées par des rigidités salariales et un chômage structurel élevé, l'ajustement à un choc d'offre de travail ne se traduit pas nécessairement par une variation des salaires. Ils rejoignent également les conclusions de la littérature empirique montrant que, une fois contrôlées les conditions initiales, les effets de la migration sur les salaires et les heures des travailleurs natifs tendent à être faibles ou nuls (Card, 2001; Dustmann *et al.*, 2012; Ottaviano et Peri, 2012; Jaeger *et al.*, 2018). Dans le contexte sud-africain, marqué par des rigidités institutionnelles importantes et une segmentation persistante du marché du travail, ces résultats suggèrent que la migration rurale–urbaine ne modifie pas substantiellement les conditions de travail des natifs à court et moyen terme.

4.2 Analyse complémentaire : migration interne globale

Cette section propose une analyse complémentaire des effets de la migration interne globale sur les résultats du marché du travail des natifs. Contrairement à l'analyse principale qui se concentre sur les flux de migration rurale–urbaine, nous examinons ici l'ensemble des mouvements migratoires internes. Ceux-ci incluent également les mouvements entre zones urbaines–urbaines, entre zones rurales–rurales et rurales–urbaines, agrégés en une variable unique de migration interne, captant l'intensité globale des flux migratoires à destination de chaque district.

Cette approche agrégée répond à trois objectifs principaux. Premièrement, elle permet de situer nos résultats dans la littérature existante, qui a traditionnellement étudié la migration interne sans toujours distinguer les sous-flux (Lucas, 1997; Lall *et al.*, 2006). Deuxièmement, elle offre un point de comparaison pour évaluer l'hétérogénéité des effets selon le type de migration. Troisièmement, elle permet d'éclairer les mécanismes d'ajustement du marché du travail dans un cadre plus large.

Cependant, cette analyse comporte d'importantes limitations méthodologiques. L'identification cau-

sale des effets de la migration interne globale est particulièrement complexe en raison de l'hétérogénéité des motifs migratoires et des sources d'endogénéité. Comme discuté précédemment, les décisions de migration sont influencées par des facteurs économiques locaux qui affectent simultanément les résultats du marché du travail des natifs, créant un problème de causalité inverse (Borjas, 2003). De plus, l'agrégation de flux migratoires aux déterminants variés peut masquer des mécanismes spécifiques et atténuer les effets estimés (Jaeger *et al.*, 2018).

La stratégie d'identification instrumentale, bien que théoriquement applicable à la migration interne globale, rencontre ici des limites pratiques importantes. L'instrument climatique, construit à partir d'anomalies de température dans les districts d'origine et de structures migratoires historiques, montre une puissance statistique réduite pour expliquer la variation agrégée des flux migratoires. Cette faiblesse instrumentale reflète la diversité des déterminants de la migration interne globale, dont seule une partie est sensible aux chocs climatiques.

Les résultats présentés dans cette section doivent donc être interprétés avec une prudence accrue. Ils documentent principalement des associations corrélacionnelles et fournissent des ordres de grandeur indicatifs, plutôt qu'une identification causale robuste. Cette analyse complémentaire sert ainsi de contrepoint à l'approche principale, mettant en lumière les défis méthodologiques spécifiques à l'étude de la migration interne globale et justifiant le choix de se concentrer sur les flux ruraux-urbains pour l'analyse causale.

4.2.1 Validation de l'instrument pour la migration interne globale

Cette sous-section examine la puissance de l'instrument climatique pour expliquer la variation de la migration interne globale. Dans cette analyse de première étape, la variable dépendante est la variation du taux de migration interne globale (ensemble des flux migratoires, incluant les mouvements ruraux-urbains, urbains-urbains et ruraux-ruraux). La variable explicative principale est l'instrument climatique de type *shift-share*, construit à partir des anomalies standardisées de température dans les districts d'origine et pondéré par les structures migratoires historiques.

L'objectif est de vérifier si cet instrument présente une corrélation suffisante avec le flux migratoire agrégé pour être utilisé dans une stratégie de variables instrumentales. Nous procédons par une série de régressions successives, en introduisant progressivement comme contrôles les conditions

économiques initiales mesurées à la vague 1 : le taux d'emploi initial, le salaire initial (log), les heures travaillées initiales (log) et le taux d'emploi formel initial. Les résultats de cette validation sont présentés dans le tableau 4.4.

TABLE 4.4 Première étape : instrument climatique et variation du taux de migration interne

	(1)	(2)	(3)	(4)	(5)
Instrument climatique	1,881 (6,551)	5,766 (6,875)	11,337* (5,924)	11,188* (5,902)	11,575* (6,037)
Taux d'emploi init.		-5,484 (5,494)	-8,391 (6,020)	-7,867 (5,992)	-7,890 (5,924)
Salaire init. (log)			-0,028 (0,142)	-0,029 (0,158)	0,001 (0,148)
Heures init. (log)				-0,658* (0,386)	-0,600 (0,367)
Taux emploi formel init.					-9,397 (7,211)
Kleibergen–Paap F	3	3	3,4	3,7	5,1
$N(\text{cluster})$	42	42	42	42	42
N	330	330	330	330	330
Anderson–Rubin F	8,85***	7,11**	5,23**	8,62** (0,015)	5,84**

Notes : La variable dépendante est la variation du taux de migration interne entre 2008 et 2011. L'instrument climatique est une anomalie de température normalisée mesurée dans les districts d'origine et agrégée au niveau des districts de destination selon une approche shift–share. Les erreurs types sont robustes et clusterisées au niveau du district de destination.

Les résultats présentés dans le tableau 4.4 indiquent que le coefficient associé à l'instrument climatique est positif dans l'ensemble des spécifications, avec des valeurs comprises entre 1,88 et 11,58. Toutefois, cet effet n'atteint la significativité statistique conventionnelle qu'à partir des spécifications les plus complètes, et uniquement au seuil de 10 %, tandis que les erreurs types demeurent relativement élevées. L'échantillon final comprend 330 individus regroupés en 42 districts de destination, soit en moyenne environ 7,9 individus par cluster. Ce nombre relativement limité de clusters est susceptible de réduire la précision statistique des estimations et d'accroître l'incertitude entourant l'évaluation de la force instrumentale.

Les statistiques de Kleibergen–Paap F , comprises entre 3,0 et 5,1 selon les spécifications, demeurent

inférieures au seuil critique de 8,7 correspondant à un biais maximal de 20 %, tel que proposé par Stock et Yogo (2005). Ces résultats indiquent que l’instrument climatique est faible dans le cas de la migration interne globale, ce qui limite la fiabilité des estimations par variables instrumentales fondées sur cette mesure agrégée et justifie une interprétation prudente des coefficients de deuxième étape.

En revanche, les tests d’Anderson–Rubin fournissent des statistiques F systématiquement significatives, comprises entre 5,23 et 8,85 selon les spécifications ($p < 0,05$). Ces résultats conduisent au rejet de l’hypothèse nulle d’un effet nul de la variable endogène dans un cadre d’inférence robuste aux instruments faibles. Ils suggèrent que, malgré une première étape peu puissante, l’instrument climatique conserve un pouvoir d’identification suffisant pour détecter l’existence d’une relation statistiquement significative, sans toutefois permettre une estimation précise de son ampleur.

Dans l’ensemble, l’instrument climatique apparaît insuffisamment puissant pour garantir une identification pleinement robuste de la migration interne globale. Ce constat motive le choix méthodologique central de ce mémoire, qui consiste à concentrer l’analyse causale principale sur les flux de migration rurale–urbaine, pour lesquels le lien entre chocs climatiques et mobilité est plus direct et la puissance instrumentale sensiblement plus élevée. L’analyse de la migration interne globale est dès lors interprétée comme une extension complémentaire permettant de situer les résultats principaux dans une perspective plus générale.

L’analyse de la première étape montre ainsi que l’instrument climatique fondé sur les anomalies de température, bien qu’exogène par construction, présente une capacité explicative limitée pour la variation de la migration interne totale. Cette faiblesse instrumentale restreint la portée causale des estimations en variables instrumentales et impose un recours privilégié à des outils d’inférence robustes aux instruments faibles (Staiger et Stock, 1997; Stock et Yogo, 2005; Andrews *et al.*, 2019).

Néanmoins, l’inclusion de la migration interne globale répond à un objectif méthodologique et empirique important. D’une part, elle permet de documenter les corrélations agrégées entre mobilité interne et ajustements du marché du travail, dans la lignée d’une littérature descriptive et corrélative abondante portant sur la migration interne dans les pays en développement et en Afrique subsaharienne (Lucas, 1997; Lall *et al.*, 2006; Gagnon *et al.*, 2021). D’autre part, elle offre un point

de comparaison essentiel pour analyser l'hétérogénéité des mécanismes d'ajustement selon le type de migration, en particulier entre la migration interne globale et les flux ruraux-urbains, pour lesquels le lien entre chocs climatiques et mobilité est théoriquement et empiriquement plus direct (Harris et Todaro, 1970; Barrios *et al.*, 2006; Henderson *et al.*, 2017).

Cette analyse complémentaire permet enfin de mettre en évidence les limites empiriques d'une approche agrégée de la migration interne. Plusieurs travaux montrent en effet que l'agrégation de flux migratoires hétérogènes — combinant migrations de proximité, migrations circulaires et déplacements temporaires — tend à atténuer les effets estimés sur les salaires et l'emploi, en masquant les canaux d'ajustement spécifiques (Jaeger *et al.*, 2018; Adao *et al.*, 2019). En ce sens, l'étude de la migration interne globale renforce, par contraste, la pertinence du choix méthodologique central du mémoire, fondé sur l'identification causale des flux ruraux-urbains.

4.2.2 Effets sans variables de contrôle

Le tableau 4.5 présente les effets estimés de la variation du taux de migration interne globale sur les salaires et les heures travaillées des natifs. Deux approches sont mobilisées : les moindres carrés ordinaires (MCO) et les variables instrumentales (VI), cette dernière utilisant l'instrument climatique décrit au chapitre 3.

TABLE 4.5 Effets de la migration interne globale sur les salaires et les heures travaillées (sans contrôles)

	Salaire (log)		Heures travaillées (log)	
	(1) MCO	(2) VI	(3) MCO	(4) VI
Taux de migration interne	0,098 (0,074)	-0,511 (0,758)	0,013 (0,022)	1,199 (4,337)
N	246	246	330	330
R^2	0,12	-0,033	0,04	-13,34
Kleibergen–Paap F		3,9		4
Anderson–Rubin F (p-val)		7,45**		8,76**

Notes : *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$. Les erreurs types sont robustes et clusterisées au niveau du district de destination. Les estimations VI utilisent l’anomalie de température comme instrument du taux de migration interne globale. Les variables dépendantes sont exprimées en log-différences entre 2008 et 2011. L’échantillon pour les salaires ($N=246$) est plus petit que celui pour les heures travaillées ($N=330$) en raison de la non-réponse partielle sur les salaires.

Le tableau 4.5 présente les effets estimés de la variation du taux de migration interne globale sur les salaires et les heures travaillées des natifs, en l’absence de variables de contrôle. Deux approches sont mobilisées : les moindres carrés ordinaires (MCO) et les variables instrumentales (VI), cette dernière utilisant l’anomalie de température comme instrument du taux de migration interne globale.

S’agissant des salaires, les résultats ne mettent pas en évidence d’effet statistiquement significatif de la migration interne globale. L’estimation MCO indique un coefficient positif mais non significatif ($\hat{\beta} = 0,098$), tandis que l’estimation VI produit un coefficient négatif et également non significatif ($\hat{\beta} = -0,511$). Le changement de signe entre les deux estimateurs suggère la présence potentielle de biais d’endogénéité dans l’estimation MCO. Toutefois, la faiblesse de l’instrument, comme l’indique la statistique de Kleibergen–Paap ($F = 3,9$), inférieure aux seuils critiques usuels, empêche toute interprétation causale robuste du coefficient VI dans cette spécification. Bien que le test d’Anderson–Rubin soit significatif au seuil de 5 %, cette évidence statistique doit être interprétée avec prudence compte tenu de la faible puissance de la première étape.

Concernant les heures travaillées, l’estimation MCO indique un effet très proche de zéro et non significatif ($\hat{\beta} = 0,013$). En revanche, l’estimation VI suggère un effet positif de grande ampleur

($\hat{\beta} = 1,199$), mais associé à une incertitude extrêmement élevée. Cette instabilité des estimations reflète directement la faiblesse instrumentale mise en évidence par la statistique de Kleibergen–Paap ($F = 4$), largement inférieure aux seuils recommandés. Néanmoins, le test d’Anderson–Rubin, robuste à la présence d’instruments faibles, est significatif au seuil de 5 %, ce qui permet de rejeter l’hypothèse nulle d’absence totale d’effet. Ce résultat suggère l’existence possible d’un ajustement du marché du travail par les heures travaillées, sans toutefois permettre d’identifier de manière fiable l’ampleur de cet ajustement. Du point de vue de la signification économique, l’ampleur du coefficient estimé par variables instrumentales pour les heures travaillées ($\hat{\beta} = 1,199$) impliquerait qu’une augmentation d’un point de pourcentage du taux de migration interne globale serait associée à une hausse d’environ 1,2 % des heures travaillées des natifs. Pour un travailleur moyen effectuant environ 40 heures par semaine, cet effet correspondrait à près de 0,5 heure supplémentaire par semaine. Toutefois, compte tenu de l’incertitude extrêmement élevée associée à cette estimation et de la faiblesse manifeste de l’instrument, cette magnitude doit être interprétée comme un ordre de grandeur purement illustratif plutôt que comme une estimation causale précise. Elle souligne surtout l’instabilité des résultats lorsque la migration interne est considérée dans son ensemble, par opposition aux flux plus homogènes de migration rurale–urbaine.

Dans l’ensemble, ces résultats indiquent que, en l’absence de contrôles, la migration interne globale n’exerce pas d’effet clairement identifiable sur les salaires des natifs, tandis que les effets potentiels sur les heures travaillées demeurent difficilement quantifiables en raison de la faiblesse de l’instrument. Cette fragilité empirique justifie une interprétation prudente des estimations VI et renforce la pertinence de concentrer l’analyse causale principale sur des formes plus spécifiques de migration interne, pour lesquelles le lien entre chocs exogènes et mobilité est plus direct et la puissance instrumentale plus élevée.

4.2.3 Effets avec variables de contrôle

Le tableau 4.6 introduit les conditions économiques initiales mesurées à la vague 1 (salaire, emploi, heures travaillées et emploi formel) comme test de robustesse, conformément aux recommandations de Jaeger *et al.* (2018).

TABLE 4.6 Effet de la migration interne globale sur les salaires et les heures travaillées des natifs (avec contrôles)

	Salaire (log)		Heures travaillées (log)	
	(1) MCO	(2) VI	(3) MCO	(4) VI
Taux de migration interne	0,140*	-0,126	-0,004	0,084
	(0,079)	(0,250)	(0,017)	(0,107)
Salaire initial (log)	-0,529***	-0,470***	0,043	0,044
	(0,131)	(0,117)	(0,028)	(0,039)
Emploi initial (%)	-0,955	-2,558	0,181	0,729
	(0,951)	(1,706)	(0,413)	(0,931)
Heures initiales (log)	0,706*	0,514	-1,231***	-1,175***
	(0,415)	(0,463)	(0,088)	(0,158)
Emploi formel (%)	2,394	0,784	0,659	1,431
	(1,711)	(2,061)	(1,061)	(1,159)
Constante	0,092	1,459	4,042***	3,567***
	(1,858)	(2,190)	(0,406)	(0,818)
Kleibergen–Paap F	–	3	–	4,5
Anderson–Rubin F (p-val)	–	4.43	–	8.09**
R^2	0.387	0.34	0.399	0.18
N	246	246	330	330

Notes : 1. Les estimations sont réalisées en différences longues entre la vague 2 (2010–2011) et la vague 1 (2008). Les variables dépendantes (salaire horaire et heures travaillées) sont exprimées en logarithme.

2. L'unité d'observation du tableau est l'individu natif présent aux deux vagues de l'enquête NIDS et ayant déclaré de manière valide ses revenus et ses heures travaillées aux deux périodes. Les migrants récents ne constituent pas des unités d'observation dans les régressions de seconde étape.

3. Les valeurs de N (246 ou 330 selon la spécification) correspondent au nombre d'individus natifs disposant d'informations complètes et exploitables sur les variables dépendantes considérées. La différence entre ces effectifs s'explique par la non-réponse partielle sur les salaires ou les heures travaillées, fréquente dans les enquêtes longitudinales.

4. Les flux migratoires utilisés comme variables explicatives sont mesurés au niveau agrégé des districts de destination. Plusieurs migrants peuvent ainsi contribuer à une même observation agrégée, sans affecter le nombre d'unités individuelles utilisées dans les estimations de seconde étape.

5. Ce tableau n'est donc pas directement comparable au Tableau 2.1, qui repose sur un échantillon individuel descriptif plus large. Le Tableau 4.6 se concentre exclusivement sur l'échantillon restreint nécessaire à l'identification causale en longue différence.

Le tableau 4.6 présente les estimations de l'effet de la migration interne globale sur les salaires et les heures travaillées des natifs, en incluant les conditions économiques initiales mesurées à la vague 1 (salaire, emploi, heures travaillées et emploi formel), conformément aux recommandations méthodologiques de Jaeger *et al.* (2018).

Les résultats confirment l'absence d'un effet salarial robuste de la migration interne globale sur les natifs. Le coefficient associé au taux de migration interne est faiblement positif et marginalement significatif en MCO ($\hat{\beta} = 0,140$, significatif au seuil de 10 %), mais devient négatif et non significatif en variables instrumentales ($\hat{\beta} = -0,126$). Cette instabilité entre les estimateurs, combinée à l'absence de significativité statistique en VI et à la faiblesse de l'instrument (statistique de Kleibergen–Paap $F = 3$), suggère que les salaires des natifs réagissent peu aux variations agrégées de la migration interne. Ce constat est cohérent avec une large partie de la littérature empirique internationale, qui met en évidence des effets salariaux faibles ou nuls lorsque les ajustements du marché du travail s'opèrent principalement par des marges autres que les prix (Card, 2001; Dustmann *et al.*, 2012; Peri, 2016).

Concernant les heures travaillées, l'estimation MCO demeure proche de zéro et non significative ($\hat{\beta} = -0,004$), tandis que l'estimation VI suggère un effet positif ($\hat{\beta} = 0,084$), bien que celui-ci ne soit pas statistiquement significatif selon les tests standards. En revanche, le test d'Anderson–Rubin, robuste à la présence d'instruments faibles, est significatif pour les heures travaillées ($F = 8,09$, $p < 0,05$), indiquant qu'une relation statistiquement détectable subsiste entre la migration interne globale et les quantités de travail, malgré la faiblesse de la première étape (Kleibergen–Paap $F = 4,5$). Conformément aux recommandations de la littérature sur l'inférence en présence d'instruments faibles, ce résultat suggère l'existence d'un effet potentiel sur les heures travaillées, sans permettre une identification précise de son ampleur (Staiger et Stock, 1997; Andrews *et al.*, 2019; Keane et Neal, 2022). Du point de vue de la signification économique, le coefficient estimé par variables instrumentales pour les heures travaillées ($\hat{\beta} = 0,084$) suggère qu'une augmentation d'un point de pourcentage du taux de migration interne globale serait associée à une hausse d'environ 0,08 % des heures travaillées des natifs. Pour un travailleur moyen effectuant environ 160 heures par mois, cet effet correspondrait à un accroissement d'environ 0,13 heure par mois. Bien que cette magnitude soit modeste et statistiquement imprécise, elle est cohérente avec l'idée que les ajustements induits par la migration interne globale s'opèrent principalement par des marges quantitatives faibles, plutôt

que par des variations salariales marquées.

Dans l'ensemble, ces résultats appellent à une interprétation prudente. Les statistiques de Kleibergen–Paap, égales à 3 pour les salaires et à 4,5 pour les heures travaillées, demeurent nettement inférieures aux seuils critiques proposés par Stock et Yogo (2005), indiquant que l'instrument climatique est peu informatif pour expliquer la variation de la migration interne globale. Cette faiblesse limite la portée causale des estimations par variables instrumentales et justifie le recours à des tests d'inférence robustes, tels que celui d'Anderson–Rubin. Ces limites sont cohérentes avec des travaux récents soulignant les difficultés d'identification causale lorsque les flux migratoires agrégés résultent de mécanismes hétérogènes et de décisions individuelles complexes (Adao *et al.*, 2019; Jaeger *et al.*, 2018).

La comparaison avec les résultats obtenus pour la migration rurale–urbaine est néanmoins instructive. Alors que la migration interne globale ne génère pas d'effets clairement identifiables sur les salaires des natifs et seulement des effets faiblement détectables sur les heures travaillées, les flux ruraux–urbains produisent des effets plus nets et statistiquement robustes, en particulier sur les quantités de travail. Cette hétérogénéité suggère que l'agrégation de l'ensemble des flux internes masque des mécanismes d'ajustement spécifiques, plus directement liés aux dynamiques de réallocation spatiale de la main-d'œuvre entre zones rurales et urbaines, comme le soulignent les modèles de migration avec rigidités urbaines et ajustement quantitatif (Todaro, 1969; Harris et Todaro, 1970; Bryan et Morten, 2019).

En définitive, l'analyse empirique met en évidence que, dans le contexte sud-africain, la migration interne globale n'affecte pas significativement les salaires des natifs, tandis que les ajustements potentiels s'opèrent, lorsque l'identification le permet, principalement par les quantités de travail. Les heures travaillées apparaissent ainsi comme la marge d'ajustement dominante face aux chocs migratoires internes, un résultat pleinement cohérent avec les prédictions théoriques et les évidences empiriques récentes pour les pays en développement (Kleemans et Magruder, 2018; Gagnon *et al.*, 2021).

CONCLUSION

Ce mémoire a étudié l'impact causal de la migration interne sur le marché du travail des natifs sud-africains, en se concentrant sur deux marges essentielles d'ajustement : les salaires et les heures travaillées. Dans un contexte post-apartheid marqué par des disparités spatiales persistantes, une urbanisation rapide et une forte segmentation du marché du travail, comprendre les effets de la mobilité interne constitue un enjeu central pour l'analyse économique et l'élaboration des politiques publiques (Posel, 2006; Turok et Borel-Saladin, 2014; Kollamparambil, 2017).

Sur le plan empirique, l'analyse repose sur les deux premières vagues de l'enquête longitudinale NIDS (2008–2011) et mobilise une stratégie en différences longues, permettant de neutraliser l'hétérogénéité individuelle invariante dans le temps. L'endogénéité des choix migratoires est traitée à l'aide d'une approche de variables instrumentales fondée sur un instrument climatique de type *shift-share*, combinant des anomalies de température observées dans les districts d'origine et des parts migratoires historiques origine–destination (Boustan, 2010; Adao *et al.*, 2018; Jaeger *et al.*, 2018).

La validité statistique de cet instrument a été évaluée à l'aide des tests de Kleibergen–Paap et d'Anderson–Rubin. Les résultats montrent que, pour la migration interne globale, l'instrument présente une puissance limitée, avec des statistiques de Kleibergen–Paap comprises entre 3 et 5, inférieures aux seuils critiques usuels (Stock et Yogo, 2005). En revanche, la puissance instrumentale est sensiblement plus élevée lorsqu'on se concentre sur les flux de migration rurale–urbaine, pour lesquels les statistiques de Kleibergen–Paap se situent entre 8 et 11. Cette différence reflète la spécificité du contexte sud-africain : les zones rurales, fortement dépendantes de l'agriculture, sont particulièrement sensibles aux chocs climatiques, qui fragilisent les moyens de subsistance et accroissent la propension à migrer vers les centres urbains.

Trois résultats principaux émergent de l'analyse empirique.

Premièrement, les salaires des natifs ne semblent pas affectés de manière robuste par la migration interne. Que ce soit pour la migration rurale–urbaine ou pour la migration interne globale, les estimations par moindres carrés ordinaires produisent des coefficients faibles et instables, tandis

que les estimations en variables instrumentales sont statistiquement indiscernables de zéro. Même lorsque l'instrument est suffisamment puissant, comme dans le cas des flux ruraux–urbains, aucune variation significative des salaires des natifs n'est détectée. Ce résultat est cohérent avec une large littérature internationale documentant des effets salariaux faibles ou nuls des migrations locales (Card, 2001; Dustmann *et al.*, 2012; Peri, 2016).

Deuxièmement, les heures travaillées des natifs apparaissent comme une marge d'ajustement plus sensible aux flux migratoires, en particulier dans le cas de la migration rurale–urbaine. Dans les spécifications sans variables de contrôle, une augmentation d'un point de pourcentage du taux de migration rurale–urbaine est associée à une hausse significative des heures travaillées des natifs. Cet effet positif, confirmé par l'approche en variables instrumentales et par les tests d'Anderson–Rubin, indique que l'ajustement au choc migratoire s'opère principalement par les quantités de travail plutôt que par les prix.

Ce résultat n'est pas fondamentalement surprenant au regard de la littérature existante. Dans de nombreux contextes caractérisés par des rigidités salariales, un chômage structurel élevé et une forte segmentation du marché du travail, l'arrivée de migrants tend à modifier l'intensité du travail des natifs plutôt que leur rémunération (Harris et Todaro, 1970; Ottaviano et Peri, 2012; Jaeger *et al.*, 2018). Dans le contexte sud-africain, où les salaires sont souvent rigides et où l'accès à l'emploi formel est contraint, l'augmentation des heures travaillées peut refléter une intensification de l'activité des natifs, une réallocation des tâches ou une réponse adaptative des entreprises face à une offre de travail accrue.

Il convient toutefois de souligner que cet effet positif sur les heures travaillées s'atténue et perd sa significativité statistique lorsque l'on introduit des contrôles pour les conditions économiques initiales des districts. Cette évolution suggère que les ajustements observés dans les spécifications sans contrôles reflètent en partie des différences structurelles préexistantes entre territoires, plutôt qu'un effet causal direct et universel de la migration sur les heures travaillées des natifs. L'effet positif identifié doit donc être interprété comme un mécanisme d'ajustement local et contextuel, plutôt que comme un résultat généralisable à l'ensemble des marchés du travail.

Troisièmement, l'analyse d'hétérogénéité met en évidence que les effets observés sont spécifiques

aux flux ruraux–urbains, tandis que la migration interne globale ne produit pas d’effets clairement identifiables sur les salaires ou les heures travaillées. Cette distinction est cohérente avec la stratégie d’identification retenue et avec les prédictions du modèle de Harris–Todaro (Harris et Todaro, 1970), selon lesquelles les migrations rurales–urbaines, déclenchées par des chocs négatifs dans les zones d’origine, exercent une pression particulière sur les marchés du travail urbains.

Au-delà des résultats empiriques, ce mémoire souligne la pertinence d’une stratégie d’identification climatique pour analyser les migrations internes dans les pays en développement. La combinaison de données climatiques satellitaires, d’informations migratoires historiques et de modèles en variables instrumentales permet d’identifier des effets causaux à une échelle infranationale encore peu explorée dans le contexte sud-africain. Sur le plan méthodologique, cette contribution montre que les instruments climatiques de type *shift–share* peuvent être mobilisés de manière crédible, à condition de restreindre l’analyse aux flux pour lesquels la relation entre chocs exogènes et mobilité est suffisamment directe — ici, les flux ruraux–urbains.

Plusieurs limites doivent toutefois être soulignées. L’analyse se limite aux deux premières vagues de l’enquête NIDS, repose sur des données autodéclarées et ne permet pas d’observer finement la dynamique temporelle des migrations ni la qualité des emplois occupés, notamment dans le secteur informel. Par ailleurs, l’attrition du panel pourrait ne pas être aléatoire. Ces limites invitent à interpréter les résultats avec prudence et à envisager des extensions futures.

En définitive, ce mémoire met en évidence une dynamique spécifique du marché du travail sud-africain : la migration interne, et plus particulièrement la migration rurale–urbaine, n’affecte pas les salaires des natifs mais peut conduire, dans certains contextes, à une intensification de leur activité mesurée en heures travaillées. Ce résultat, loin d’être paradoxal, s’inscrit dans une littérature soulignant que, dans les économies caractérisées par des rigidités institutionnelles et un chômage élevé, les ajustements au choc migratoire s’opèrent principalement par les quantités de travail plutôt que par les prix. Ces enseignements soulignent l’importance de distinguer les différentes formes de migration interne afin de mieux comprendre leurs implications économiques et d’orienter des politiques publiques adaptées aux réalités territoriales sud-africaines.

ANNEXE A
CORRÉLATIONS ENTRE LE SALAIRE ET LES VARIABLES

TABLE A.1 Corrélations de Pearson entre la variation du salaire horaire (log-différence) et les principales variables explicatives

Variable	Corrélation de Pearson
Migration rurale–urbaine (Δ)	-0.027
Heures travaillées (log, initiales)	0.162
Migration totale (Δ)	0.099
Instrument climatique (température, Δ)	-0.064
Taux d’emploi initial	-0.100
Salaire horaire (log, initial)	-0.394

L’examen des corrélations de Pearson entre la variation du salaire (en log-différences) et les principales variables explicatives met en évidence plusieurs résultats saillants. La variation du salaire présente une corrélation positive mais modérée avec les heures travaillées retardées, suggérant que les ajustements salariaux observés sont partiellement liés aux dynamiques de l’offre de travail. Le salaire initial est en revanche fortement et négativement corrélé à la variation du salaire, ce qui est cohérent avec un mécanisme de convergence salariale et de réversion à la moyenne. À l’inverse, la migration rurale–urbaine prise isolément, la migration totale, l’emploi initial ainsi que l’instrument climatique ne présentent que des corrélations faibles avec la variation du salaire, ce qui suggère l’absence de relation linéaire bivariée marquée et justifie le recours à une stratégie d’identification en variables instrumentales pour l’analyse causale.

ANNEXE B
FACTEURS D'INFLATION DE VARIANCE (VIF)

TABLE B.1 Facteurs d'inflation de variance (VIF) pour la régression OLS du salaire

Variable	VIF	1/VIF
Heures initiales (log)	1.96	0.510
Variation des heures travaillées (Δ)	1.90	0.525
Migration rurale-urbaine (Δ)	1.89	0.530
Instrument climatique (température, Δ)	1.49	0.670
Migration totale (Δ)	1.49	0.672
Taux d'emploi initial	1.16	0.860
Salaire initial (log)	1.05	0.952
Moyenne VIF	1.56	

Note : Le VIF mesure la multicolinéarité entre les variables explicatives. Les valeurs observées sont nettement inférieures aux seuils critiques usuels (5 et 10), indiquant l'absence de multicolinéarité problématique dans les spécifications estimées Wooldridge2010.

Les résultats indiquent une multicolinéarité modérée (moyenne ≈ 3), compatible avec une interprétation fiable des coefficients.

ANNEXE C

LIEN ENTRE L'ÉMIGRATION PRÉDITE À L'ORIGINE ET L'IMMIGRATION OBSERVÉE À DESTINATION

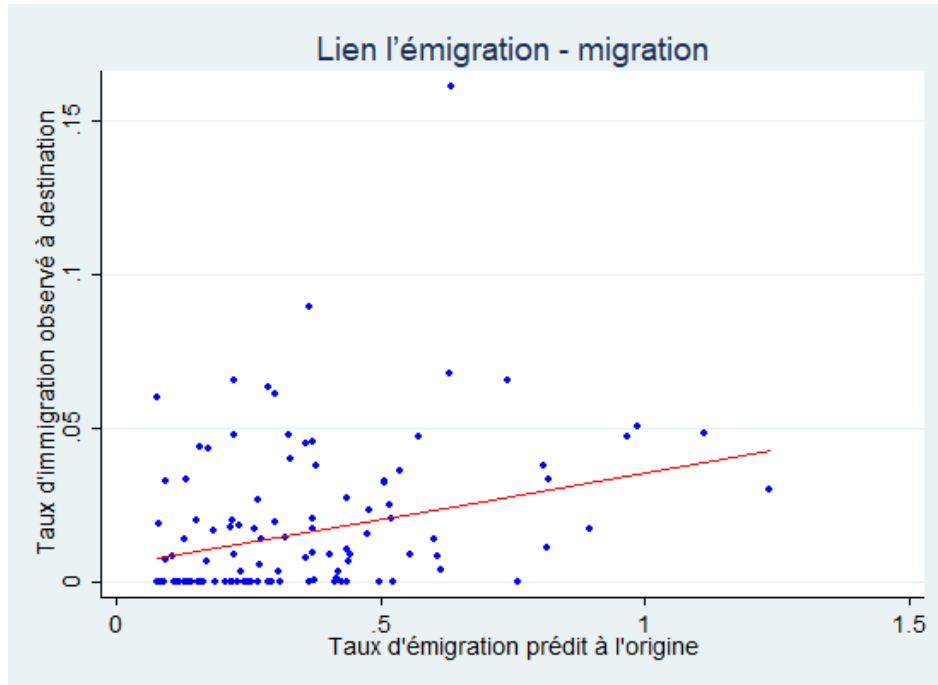


FIGURE C.1 Lien entre l'émigration prédite à l'origine et l'immigration observée à destination

La figure C.1 montre une relation positive entre l'émigration prédite à l'origine (via chocs climatiques) et l'immigration observée à destination, cohérente avec une propagation des chocs d'origine vers les flux entrants.

BIBLIOGRAPHIE

- Adao, R., Arkolakis, C. et Esposito, F. (2019). *General Equilibrium Effects in Space : Theory and Measurement*. Working Paper 25544, National Bureau of Economic Research
- Adao, R., Kolesar, M. et Morales, E. (2018). *Shift-Share Designs : Theory and Inference*. Working Paper 24944, National Bureau of Economic Research
- Anderson, T. W. et Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, 20(1), 46–63.
<http://dx.doi.org/10.1214/aoms/1177730090>
- Andrews, I., Stock, J. H. et Sun, L. (2019). Weak instruments in instrumental variables regression : Theory and practice. *Annual Review of Economics*, 11, 727–753.
<http://dx.doi.org/10.1146/annurev-economics-080218-025643>
- Angrist, J. D. et Pischke, J.-S. (2009). *Mostly Harmless Econometrics : An Empiricist's Companion*. Princeton, NJ : Princeton University Press.
- Barrios, S., Bertinelli, L. et Strobl, E. (2006). Climate change and rural-urban migration : The case of sub-saharan africa. *Journal of Urban Economics*, 60(3), 357–371.
<http://dx.doi.org/10.1016/j.jue.2006.04.005>
- Basso, G. et Peri, G. (2015). The association between immigration and labor market outcomes in the united states. *IZA Journal of Migration*, 4(1), 1–30.
- Bazzi, S., Gaduh, A., Rothenberg, A. D. et Wong, M. (2017). Shifting ground : The changing dynamics of intergenerational mobility in indonesia. *American Economic Review*, 107(2), 434–439.
- Beauchemin, C. (2011). Rural-urban migration in west africa : Towards a reversal? *Population, Space and Place*, 17(1), 47–72.
- Bhattacharya, P. C. (1993). Rural-urban migration in economic development. *Journal of Economic Surveys*, 7(3), 243–281. <http://dx.doi.org/10.1111/j.1467-6419.1993.tb00167.x>
- Borjas, G. J. (2003). The labor demand curve is downward sloping : Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4), 1335–1374.
- Bouare, O. (2000). Determinants of internal migration. *Southern African Journal of Demography*, 8(1), 23–28.
- Boustan, L. P. (2010). Was postwar suburbanization “white flight” ? evidence from the black migration. *The Quarterly Journal of Economics*, 125(1), 417–443.
<http://dx.doi.org/10.1162/qjec.2010.125.1.417>
- Bryan, G. et Morten, M. (2019). The aggregate productivity effects of internal migration : Evidence from indonesia. *Journal of Political Economy*, 127(5), 2229–2268.
<http://dx.doi.org/10.1086/701810>

- Bryceson, D. F. (2000). Disappearing peasantries ? rural labour redundancy in the neo-liberal era and beyond. In D. F. Bryceson, C. Kay, et J. Mooij (dir.), *Disappearing Peasantries ? Rural Labour in Africa, Asia and Latin America* 299–326. London : Intermediate Technology Publications.
- Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1), 22–64.
<http://dx.doi.org/10.1086/209979>
- Casentini, G., Hammond, L. et Bakewell, O. (2024). Internal migration, informality and inequality in the global south. In O. Bakewell, D. Jolivet, et G. Singh (dir.), *The Palgrave Handbook of South-South Migration and Inequality* 371–392. Springer International Publishing.
- Collinson, M. A., Tollman, S. M. et Kahn, K. (2007). Migration, settlement change and health in post-apartheid south africa : Triangulating health and demographic surveillance with national census data. *Scandinavian Journal of Public Health*, 35(69 Suppl), 77–84.
<http://dx.doi.org/10.1080/14034950701356401>
- Dell, M., Jones, B. F. et Olken, B. A. (2014). What do we learn from the weather ? the new climate-economy literature. *Journal of Economic Literature*, 52(3), 740–798.
<http://dx.doi.org/10.1257/jel.52.3.740>
- Dustmann, C., Frattini, T. et Preston, I. P. (2012). The effect of immigration on the distribution of wages. *Review of Economic Studies*, 80(1), 145–173.
- Gagnon, J., Hall, J. et Pouchol, M. (2021). *Internal migration and labor market adjustment in developing countries*. Working Paper 367, OECD Development Centre
- Gray, C. et Mueller, V. (2012). Drought and population mobility in rural ethiopia. *World Development*, 40(1), 134–145.
- Harris, J. R. et Todaro, M. P. (1970). Migration, unemployment and development : A two-sector analysis. *American Economic Review*, 60(1), 126–142.
- Henderson, J. V., Storeygard, A. et Deichmann, U. (2017). Has climate change driven urbanization in africa ? *Journal of Development Economics*, 124, 60–82.
<http://dx.doi.org/10.1016/j.jdeveco.2016.09.001>
- Hu, Y. (2023). Global monitoring of extreme heat using satellite-based land surface temperature data. *Ecological Indicators*, 154, 110857.
- Imbert, C., Seror, M. A. R., Zhang, Y. et Zylberberg, Y. (2022). Migrants and firms : Evidence from china. *American Economic Review*, 112(6), 1885–1914.
- Jaeger, D. A., Ruist, J. et Stuhler, J. (2018). Shift-share instruments and the impact of immigration. *American Economic Journal : Applied Economics*, 10(2), 1–38.
- Keane, M. P. et Neal, T. (2022). A practical guide to weak instruments in applied work. *Journal of Econometrics*, 231(2), 393–419. <http://dx.doi.org/10.1016/j.jeconom.2022.07.007>
- Kelley, C. P., Mohtadi, S., Cane, M. A., Seager, R. et Kushnir, Y. (2015). Climate change in the fertile crescent and implications of the recent syrian drought. *Proceedings of the National Academy of Sciences*, 112(11), 3241–3246. <http://dx.doi.org/10.1073/pnas.1421533112>

- Kleemans, M. et Magruder, J. (2018). Labour market responses to immigration : Evidence from internal migration driven by weather shocks. *The Economic Journal*, 128(613), 2032–2065.
- Kollamparambil, U. (2017). Impact of internal in-migration on income inequality in receiving areas : A district level study of south africa. *The Journal of Development Studies*, 53(12), 2145–2163.
- Lall, S. V., Selod, H. et Shalizi, Z. (2006). *Rural-urban migration in developing countries : A survey of theoretical predictions and empirical findings*. Policy Research Working Paper 3915, The World Bank.
- Leibbrandt, M., Woolard, I., McEwen, H. et Koep, C. (2010). *Employment and Inequality Outcomes in South Africa*. Working Paper 54, Southern Africa Labour and Development Research Unit (SALDRU), University of Cape Town, Cape Town
- Lucas, R. E. B. (1997). Internal migration in developing countries. In *Handbook of Population and Family Economics*, volume 1 721–798. Elsevier.
- Lucas, R. E. B. (2005). *International Migration and Economic Development : Lessons from Low-Income Countries*. Edward Elgar.
- Maccini, S. et Yang, D. (2009). Under the weather : Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3), 1006–1026.
<http://dx.doi.org/10.1257/aer.99.3.1006>
- Munshi, K. (2003). Networks in the modern economy : Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2), 549–599.
- Naidu, S. et Hornbeck, R. (2014). When the levee breaks : Black migration and economic development in the american south. *American Economic Review*, 104(3), 963–990.
<http://dx.doi.org/10.1257/aer.104.3.963>
- Noble, M., Zembe, W., Wright, G. et Avenell, D. (2013). *Multiple deprivation and income poverty at small area level in South Africa in 2011*. Report, Southern African Social Policy Research Institute (SASPRI). Report based on the South African Index of Multiple Deprivation (SAIMD).
- Nyoni, G. et Kollamparambil, U. (2022). Rural-urban migration and the well-being of the migrant-sending households : An impact evaluation study. *South African Journal of Economic and Management Sciences*, 25(1), 1–12.
<http://dx.doi.org/10.4102/sajems.v25i1.4120>
- Ottaviano, G. I. P. et Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1), 152–197.
- Peri, G. (2016). Immigrants, productivity, and labor markets. *Journal of Economic Perspectives*, 30(4), 3–30.
- Posel, D. (2006). Moving on : Patterns of labour migration in post-apartheid south africa. In C. Cross, D. Gelderblom, N. Roux, et J. Mafukidze (dir.), *Africa on the Move : African Migration and Urbanisation in Comparative Perspective* 217–231. Cape Town : HSRC Press.

- Posel, D. et Casale, D. (2003). What has been happening to internal labour migration in south africa, 1993-1999? *South African Journal of Economics*, 71(3), 455–479.
- Sharp, M. (2021). The labour market impacts of female internal migration : Evidence from the end of apartheid. *Regional Science and Urban Economics*, 91, 103624.
- Staiger, D. et Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557–586. <http://dx.doi.org/10.2307/2171753>
- Statistics South Africa (2023). Annual report 2022/2023. Consulté en juillet 2024.
- Stock, J. H. et Yogo, M. (2005). Testing for weak instruments in linear iv regression. In *Identification and Inference for Econometric Models : Essays in Honor of Thomas Rothenberg* 80–108. Cambridge University Press.
- Tabellini, M. (2022). The economic and political effects of immigration : Evidence from the age of mass migration. In *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press
- Tacoli, C. (2006). *The Earthscan Reader in Rural-Urban Linkages*. Earthscan.
- Todaro, M. P. (1969). A model of labor migration and urban unemployment in less developed countries. *The American Economic Review*, 59(1), 138–148.
- Tombe, T. et Zhu, X. (2019). Trade, migration, and productivity : A quantitative analysis of china. *American Economic Review*, 109(5), 1843–1872.
- Turok, I. et Borel-Saladin, J. (2014). Is urbanisation in south africa on a sustainable trajectory? *Development Southern Africa*, 31(5), 675–691.
<http://dx.doi.org/10.1080/0376835X.2014.937524>
- Williams, C. A. et Others (2024). Global geostationary satellite record of surface temperature anomalies. *Scientific Data*, 11(1), 123.
<http://dx.doi.org/10.1038/s41597-024-03074-w>
- Wittenberg, M. et Collinson, M. (2007). Restructuring of households in rural south africa : Reflections on average household size in the agincourt sub-district 1992-2003. *Scandinavian Journal of Public Health*, 35(69 Suppl), 85–93.
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2 éd.). Cambridge, MA : MIT Press.